2015

Flavour studies with LHCb:

b-meson mixing, lepton-flavour violation and the velo upgrade

Thomas Bird

http://thomasbird.com

The University of Manchester
Faculty of Engineering and Physical Sciences
School of Physics and Astronomy

CERN-THESIS-2015-393
19/11/2015

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy (PhD) in the Faculty of Engineering and Physical Sciences

Contents

Abstract

Semileptonic B decays of the type $B_q^0 \to D_q^- \mu^+ \nu$ (where $D^- \to K^- K^+ \pi^-$) are selected and their lifetimes are corrected using a statistical simulation-based correction called the k-factor. Using $1\,\text{fb}^{-1}$ of LHCb data the B_d^0 and B_s^0 mixing frequencies are measured to be

$$\Delta m_d = (0.503 \pm 0.011\,(\text{stat}) \pm 0.013\,(\text{syst}))\,\text{ps}^{-1} \text{ and}$$
$$\Delta m_s = (17.93 \pm 0.22\,(\text{stat}) \pm 0.15\,(\text{syst}))\,\text{ps}^{-1}\,.$$

We exclude the null hypothesis of no mixing for the B_d^0 and B_s^0 by 5.8 and 13.0 standard deviations respectively. This is the first observation of $B_s^0 - \bar{B}_s^0$ mixing using only semileptonic B decays.

The lepton flavour violating decay $D^0 \to e^\pm \mu^\mp$ is searched for, using tagged D^0 decays from $D^{*+} \to D^0 \pi^+$, and the measurement is normalised using $D^0 \to K^- \pi^+$ decays. No evidence is seen of an excess over the expected background and so a limit is placed $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) < 1.3 \times 10^{-8}$ at a 90 % confidence level using $3\,\text{fb}^{-1}$ of LHCb data. This improves the previous measurement by a factor of 20 and is the world's best measurement.

Possible upgrades to the LHCb VELO detector are simulated and aspects of the upgraded detector are optimised to ensure that all tracks within the angular acceptance can be detected with high precision. Finally the simulated performance of the current and upgraded VELO detectors are compared.

Lay Abstract

This thesis contains two pieces of research. The first studies two similar particles: the B_d^0 and the B_s^0. Both of these particles change between their normal state and their antimatter state many times per second. We measured that the B_d^0 changes to its antimatter state and back again about 80 thousand million times per second, while the B_s^0 changes 3 million million times per second. This was measured when the B_d^0 and the B_s^0 decayed into several other particles, one of which we couldn't detect. This is the first time that the B_s^0 has been seen to change back and forth between its normal and antimatter states when some particles it decayed into couldn't be detected.

Next we searched for D^0 particles decaying into an electron and a muon. Our theory of particle physics says that this shouldn't happen and indeed we were unable to see it happening. We were able to say that if it does happen, there is less than a one in 77 million chance of it happening, which makes it the worlds best measurement of this and improves upon the previous measurement by 20 times.

In addition, I contributed to the design of a new particle detector that precisely measures the origin of particles created in the collisions of the LHC. In 2020, the new detector will replace the current detector. I simulated the new detector and used these simulations to optimise it.

Declaration

This work represents the combined efforts of the LHCb collaboration. Some of the content has been published elsewhere and/or presented to several audiences. No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

Thanks to all the people who helped me during my PhD: in the first few years Sajan and Rob who watched over me while working on the Δm measurements along with Adam and the rest of the semileptonic working group; Tim, Paula, Chris, Malcolm, Heinrich and the VELO group for the help and guidance on all things VELO; and finally Marco, Eduardo, Sneha, Jon and the charm working group for the support while working on $D^0 \to e^{\pm}\mu^{\mp}$. Thanks to all in the Manchester LHCb group for the many useful ideas and fruitful discussions. A special thanks to George who taught me grammar and supervised me throughout my PhD.

Thanks to everyone in the Manchester HEP group for the coffee-time discussions, especially Fred and Sabah, who often livened them up. Thanks to the RAL-LHCb and LHCb-RICH groups for many lunchtime coffees, especially Antonis. Apologies to everyone I had to share an office with, both at CERN and in Manchester. Thanks to the LTA at CERN for the many entertaining Friday nights. Well done to Duncan, Mark and Jamie for successfully living with me. I'd like to thank Sam and Ali for getting me jumping again and the BCPA and MUSC for many fun weekends, as well as Noggin for the smooth reserve ride. Thanks to Mum, Dad, Rob and Nick!

Finally, thanks to Grandad and George for reading the whole of this thesis, along with my two examiners Andy and Francesca.

CHAPTER 1

Theoretical background

The standard model of particle physics (SM) is a powerful theory which has proven robust time and time again, despite its known issues. But, it is because of these issues that we know that the SM is not the ultimate theory which governs our world. The SM is widely discussed in literature and so it will not be covered here, but a review of the SM is given in Ref. [1]. In this chapter, a brief theoretical overview will be given along with the motivation for the two distinct physics analyses which this thesis contains: B^0_d and B^0_s mixing frequency measurements followed by a search for the lepton flavour violating decay $D^0 \to e^{\pm}\mu^{\mp}$.

1.1 Flavour violation

In both analyses flavour violation is searched for. In the SM (with massless neutrinos) lepton flavour violation is forbidden as there are no couplings

through which this could occur. However, in the case of B_d^0 and B_s^0 mixing, the weak interaction allows for flavour changing processes to occur in the quark sector.

All quark flavour changing processes are mediated by the $W^{\pm}$ boson in the SM and so they are known as weak charged-current interactions. The six quark mass eigenstates (u, c, t, d, s and b) are eigenstates of the electromagnetic and strong interactions. The flavour eigenstates are eigenstates of the weak interaction. By convention the flavour and mass eigenstates are chosen to be equal for the up-type quarks (u, c and t), while for the down-type quarks the flavour eigenstates (d', s' and b') are a superposition of the down-type mass eigenstates. In a simple model with only two quark generations, these can be written as

$$d' = \cos(\theta_c)\, d + \sin(\theta_c)\, s$$
$$s' = -\sin(\theta_c)\, d + \cos(\theta_c)\, s\,,$$

which can also be expressed in matrix form

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos(\theta_c) & \sin(\theta_c) \\ -\sin(\theta_c) & \cos(\theta_c) \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}\,.$$

This is a 2×2 unitary matrix known as the Cabibbo matrix and θ_c is the Cabibbo angle which describes the mixing between the first two generations. Expanding this to three generations one obtains the Cabbibo-Kobayashi-

Maskawa (CKM) matrix [2],

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.1}$$

where the elements in this matrix, V_{ij}, are the coupling strength of a $q_i q_j W^+$ vertex, in which q_i represents a down-type quark and q_j represents an up-type quark. The charge and parity ($\mathcal{CP}$) conjugate of this interaction vertex is $\overline{q}_i \overline{q}_j W^-$ and it's coupling strength is V_{ij}^*. For $\mathcal{CP}$-violation (CPV) to occur the coupling strengths of the two $\mathcal{CP}$-conjugates need to be different and so $V_{ij} \neq V_{ij}^*$. This only happens if there is an imaginary component in the matrix element.

In the SM the CKM matrix is a complex 3×3 unitary matrix. Using the fact that this matrix must be unitary, the degrees of freedom can be reduced from 18 to only four. Of the four free parameters, three are mixing angles (θ_{ij}, where $ij = 12, 13, 23$), which are similar to the Cabibbo angle. The remaining free parameter is the complex phase (δ_{13}), which adds an imaginary component to the matrix elements; this is the sole source of CPV in the charged-current weak interaction. In this parametrisation, the CKM

matrix becomes

$$V_{\text{CKM}} = \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta_{13}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}$$
$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.2}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$. Another common way to parametrise the matrix is with the Wolfenstein parametrisation, which uses the knowledge from experiments that the angles are hierarchical ($s_{13} \ll s_{23} \ll s_{12} \ll 1$) to find a set of four parameters (λ, ρ, η and A) to describe the matrix that are all of order unity. This parametrisation is an approximate one and so usually only terms below $\mathcal{O}(\lambda^4)$ are considered. Using this approximate parametrisation, the CKM matrix becomes

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \,. \tag{1.3}$$

Here it can be seen that the two terms with imaginary components $\mathcal{O}(\lambda^3)$ are V_{ub} and V_{td}.

Again using the fact that the matrix must be unitary, 12 equations can be constructed where three complex terms sum to zero. Plotting the cumulative sum of these terms on the complex plane would create a closed triangle if

the matrix was indeed unitary. Of these twelve equations, only two involve terms which are the same order of λ and one of these creates the triangle known as The Unitarity Triangle; the equation that describes this triangle is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \ ,$$

where all of these terms are $\mathcal{O}\left(\lambda^3\right)$. The triangle is shown in Fig. 1.1, where the lengths of all sides have been divided by $|V_{cd}V_{cb}^*|$ (which is the length of the best measured side) and the triangle is rotated; this results in the side representing $|V_{cd}V_{cb}^*|$ running from $(0, 0)$ to $(1, 0)$ and the apex being placed at $(\bar{\rho}, \bar{\eta})$, where $\bar{\rho} = \rho\bar{c}$, $\bar{\eta} = \eta\bar{c}$ and $\bar{c} = \sqrt{1-\lambda^2}$.

In order to generate the observed matter-antimatter asymmetry of the Universe, a large amount of CPV is necessary. The single phase, δ_{13}, provided by the SM fails to provide enough CPV by several orders of magnitude. In most new physics models, additional sources of CPV are created. In the SM neutrinos do not have mass; however, there is now strong evidence that

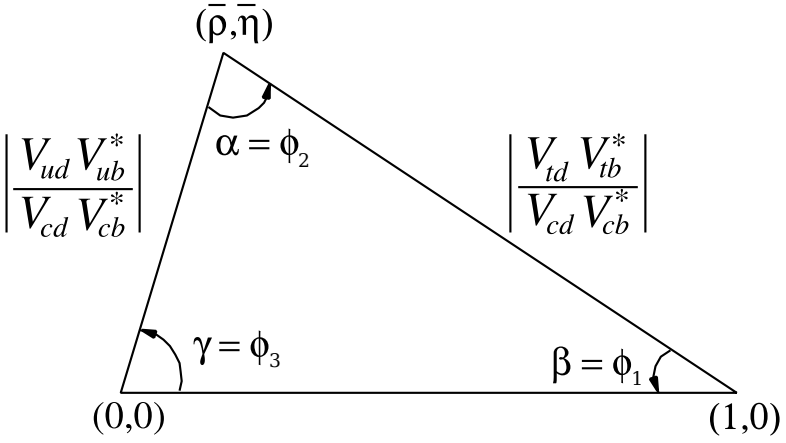

Figure 1.1: The Unitarity Triangle plotted on the complex plane. It is constructed from the unitarity requirements of the CKM matrix. The three angles of this triangle are also defined in this diagram. Figure taken from Ref. [1].

the three flavour states of the neutrinos (ν_e, ν_μ and ν_τ) oscillate between one another and so the mass eigenstates (ν_1, ν_2 and ν_3) must have non-zero masses. The SM can be extended to a standard model with non-degenerate massive neutrinos (νSM). In the νSM neutrino mixing is governed by a matrix analogous to the CKM matrix, called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [3–5], which is given by

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} . \tag{1.4}$$

If the neutrinos in the νSM are Majorana in nature [6], then the PMNS matrix would have three phases through which CPV could occur. The addition of these three phases is a very interesting and attractive solution to the lack of CPV in the SM. By convention we commonly refer to the flavour eigenstates of the neutrinos (ν_e, ν_μ and ν_τ), unlike the quark sector where the mass eigenstates are commonly referred to. The hierarchy ($s_{13} \ll s_{23} \ll s_{12} \ll 1$) of the CKM matrix does not exist in the PMNS matrix. Figure 1.2 illustrates the difference between the magnitudes of the elements in the CKM matrix and the PMNS matrix. Since flavour violation is possible in the neutral leptons, the charged leptons could also violate lepton flavour conservation and so the charged lepton sector should also be studied in detail. A brief overview of charged lepton-flavour violation (CLFV) is given in Section 1.3, but first B mixing is discussed.

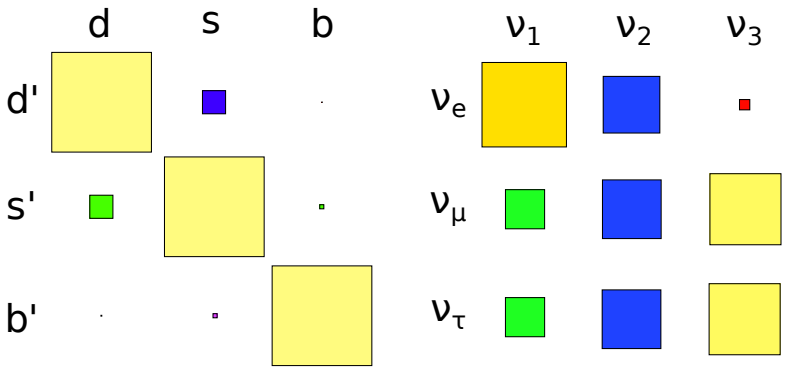

Figure 1.2: Illustration of the relative sizes of the elements in the CKM (left) and PMNS (right) matrices; the area is the square of the magnitude of the matrix element. The mass eigenstates are shown along the top row and the flavour eigenstates are shown down the side. The figure was adapted from Ref. [7].

1.2 *B* mixing

There are two neutral B-meson systems that display particle-antiparticle mixing, the B^0_d and the B^0_s mesons. This means that these two mesons have a sinusoidally oscillating probability of being detected at a later time as their antiparticle. This process is very similar for both the B^0_d and B^0_s mesons, and so in the formalism of this process they are often denoted together as B^0_q, where $q = d, s$. In the SM, the transitions $B^0_q \to \bar{B}^0_q$ and $\bar{B}^0_q \to B^0_q$, which allow mixing to occur, happen via weak interactions. Figure 1.3 shows the box diagrams which are the lowest order interactions of this process. The box diagrams are mediated by two virtual $W^\pm$ bosons and two virtual up-type quarks. Because the particles in the loop are virtual, their mass can be much larger than the energy of the collider that created the B^0_q meson, and so measurements of mixing are sensitive to new physics at energy scales

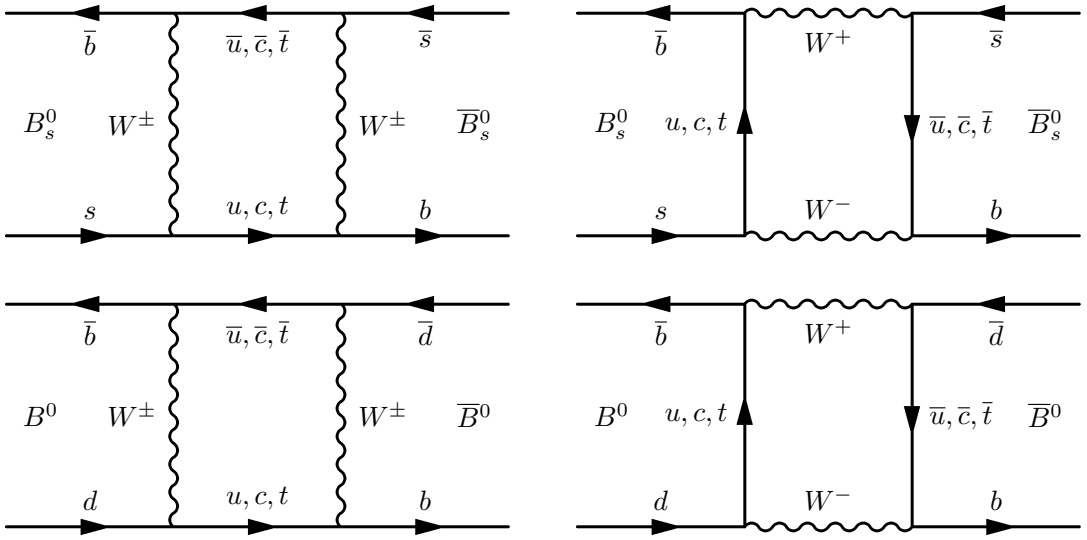

Figure 1.3: Four Feynman box diagrams showing two different ways of exchanging virtual particles which allow both the B_d^0 and B_s^0 mesons to mix. The diagrams are taken from Ref. [8].

much larger than can be directly probed. In the following section combined global charge, parity and time ($\mathcal{CPT}$) symmetry is assumed.

The time dependence of the B_q^0 system is described by the Schrödinger equation,

$$i\frac{d}{dt}\begin{pmatrix}|B_q^0(t)\rangle\\ |\bar{B}_q^0(t)\rangle\end{pmatrix} = \left(M^q - \tfrac{i}{2}\Gamma^q\right)\begin{pmatrix}|B_q^0(t)\rangle\\ |\bar{B}_q^0(t)\rangle\end{pmatrix}, \tag{1.5}$$

where $|B_q^0(t)\rangle$ and $|\bar{B}_q^0(t)\rangle$ are the flavour eigenstates at proper time t, while M^q and Γ^q are 2×2 Hermitian matrices containing the real and imaginary (or equivalently the mass and decay) components of the effective Hamiltonian. The diagonal components of M^q are M_{11}^q and M_{22}^q; these represent the masses of the of the light (m_L^q) and heavy (m_H^q) mass eigenstates. Similarly, the diagonal components of Γ^q are Γ_{11}^q and Γ_{22}^q, which represent the decay

widths of the light ($\Gamma_q{}^L$) and heavy ($\Gamma_q{}^H$) mass eigenstates. The off-diagonal components of M^q and Γ^q, M^q_{12} and Γ^q_{12}, are associated with the flavour changing transitions.

The two mass eigenstates of the B^0_q system can be written as a superposition of the flavour eigenstates,

$$\begin{aligned} |B^L_q\rangle &= p|B^0_q\rangle + q|\bar{B}^0_q\rangle \\ |B^H_q\rangle &= p|B^0_q\rangle - q|\bar{B}^0_q\rangle \ , \end{aligned}$$

where $|B^L_q\rangle$ is the light mass eigenstate, $|B^H_q\rangle$ is the heavy mass eigenstate and q and p are the coefficients of the two flavour eigenstates $|B^0_q\rangle$ and $|\bar{B}^0_q\rangle$. In the case of no CPV in B^0_q mixing, the coefficients satisfy the condition $|p/q| = 1$; currently it seems as if there is no CPV in B^0_q mixing despite previous measurements that were in tension with SM predictions [9–11]. Calculating theoretical predictions of the mixing parameters is often performed in terms of $|M^q_{12}|$, $|\Gamma^q_{12}|$ and $\phi^q = \alpha\mathrm{rg}(-M^q_{12}/\Gamma^q_{12})$. The physically observable quantities of the system are the mass difference (Δm_q) and decay width difference ($\Delta\Gamma_q$) between the two mass eigenstates; these are related to the theoretically calculable quantities through the relations

$$\begin{aligned} \Delta m_q &\equiv m^H_q - m^L_q = 2|M^q_{12}|\left(1 + \frac{1}{8}\frac{|\Gamma^q_{12}|^2}{|M^q_{12}|^2}\sin^2\phi_q + \ldots\right) > 0 \ , \\ \Delta\Gamma_q &\equiv \Gamma^L_q - \Gamma^H_q = 2|\Gamma^q_{12}|\cos\phi\left(1 + \frac{1}{8}\frac{|\Gamma^q_{12}|^2}{|M^q_{12}|^2}\sin^2\phi_q + \ldots\right) \ , \end{aligned}$$

where Δm_q is defined to be positive and the sign of $\Delta\Gamma_q$ needs to be experimentally determined. In the limit of no CPV, these equations become

$$\Delta m_q \backsimeq 2|M^q_{12}| \quad \Delta\Gamma_q \backsimeq 2|\Gamma^q_{12}| \ .$$

In the SM, the value of M_{12}^q can be calculated [1] with

$$M_{12}^q = -\frac{G_F^2}{12\pi^2} m_{B_q^0} m_{W^\pm}^2 S_0\left(m_t^2/m_{W^\pm}^2\right) \eta_B B_{B_q^0} f_{B_q^0}^2 \left|V_{tq}^* V_{tb}\right|^2, \tag{1.6}$$

where G_F is the Fermi constant [12], S_0 gives the functional dependence on the ratio of the top quark and $W^\pm$ masses (m_t and $m_{W^\pm}$) [13], η_B is a known quantum chromodynamics (QCD) correction factor of order one [14], $m_{B_q^0}$ is the average mass of the heavy and light mass eigenstates of the B_q^0 meson, $B_{B_q^0}$ is the 'bag' factor [15], $f_{B_q^0}$ is the decay constant for the B_q^0 meson [15] and V_{ij} are elements of the CKM matrix. The functional form of $S_0(x)$ can be approximated by the function $0.784x^{0.76}$ [16]. This then leads to SM values for Δm_q [17–19] of

$$\Delta m_d^{\text{SM}} = 0.543 \pm 0.091\,\text{ps}^{-1}\,,$$
$$\Delta m_s^{\text{SM}} = 17.3 \pm 2.6\,\text{ps}^{-1}\,.$$

If instead the ratio of the two mass differences is calculated, many of the terms cancel and one finds

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s^0}}{m_{B_d^0}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2}, \tag{1.7}$$

where $m_{B_d^0}$ and $m_{B_s^0}$ are the masses of the B_d^0 and B_s^0 respectively and $\xi = \left(\sqrt{B_{B_s^0}}\, f_{B_s^0}\right) \Big/ \left(\sqrt{B_{B_d^0}}\, f_{B_d^0}\right) = 1.268 \pm 0.063$ [21] is calculated from lattice QCD. Many of the theoretical uncertainties are removed in this ratio and so it allows the relatively precise calculation of the ratio of the two CKM elements $|V_{ts}|^2 \big/ |V_{td}|^2$. This ratio is proportional to one side of the unitarity triangle and is shown in Fig. 1.4, where the orange circle shows the limit from the measurement of the ratio of Δm_d and Δm_s. Currently the ratio of mass differences is one of the strongest constraints on the unitarity of this

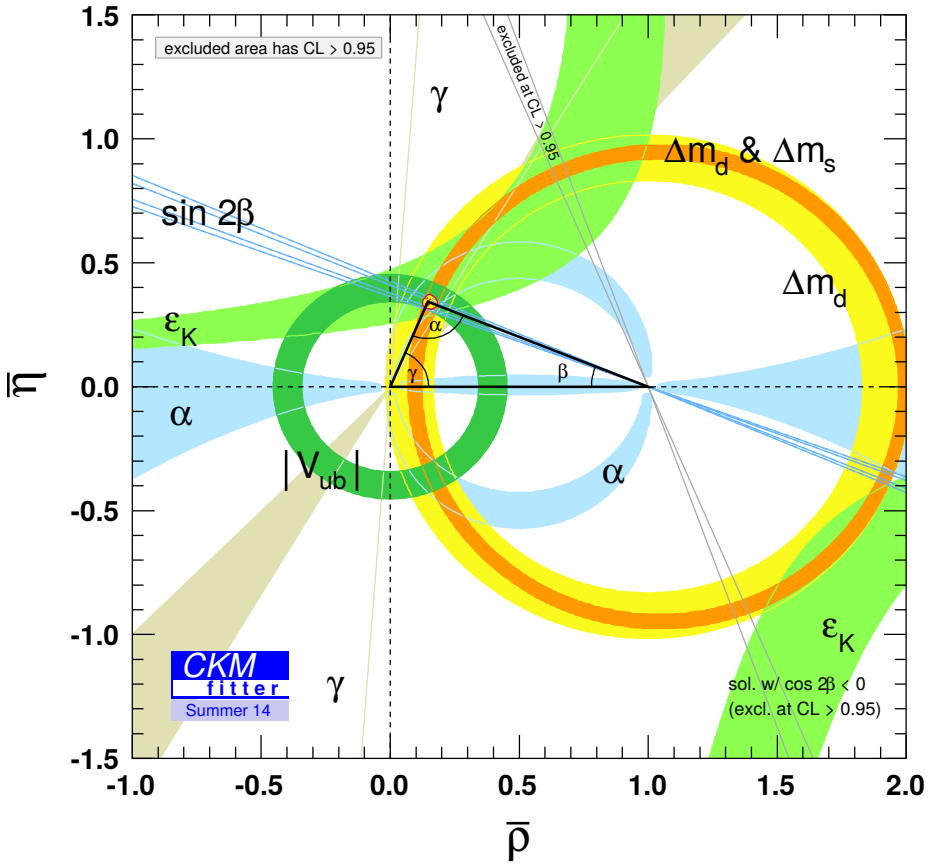

Figure 1.4: Diagram showing the constraints on the CKM matrix in the global CKM fit. The yellow circle shows the limits imposed by measurement of Δm_d, while the orange circle is derived from the ratio of Δm_d and Δm_s. The plot is taken from Ref. [20].

triangle and so the further reduction of the uncertainties on Δm_d and Δm_s would lead to improved constraints on the unitarity of the CKM matrix.

Many papers have been written on the implications that new physics would have on B^0_q mixing. As yet undiscovered particles can enter into the box diagrams as virtual particles and alter B^0_q mixing parameters. The effect of the minimal supersymmetric standard model (MSSM) on B^0_q mixing is studied in Ref. [22], while the effect of multiple Higgs bosons is studied in Ref. [23]. A more generic search for new physics is performed in Ref. [15]

and finds that the SM is disfavoured with p-values equivalent to $3.6\,\sigma$ and $3.3\,\sigma$ in the B_d^0 and B_s^0 systems respectively. To both constrain the CKM matrix and explore possible new physics contributions it is very interesting to measure B_q^0 mixing parameters such as Δm_q, the measurements of which are presented in Chapter 6.

1.2.1 Time dependence of mixing

In order to model the mixing and extract the mixing frequencies, the time evolution of the B_q^0 mesons must be understood. The time evolution of either a pure $|B_q^0(t)\rangle$ or $|\bar{B}_q^0(t)\rangle$ state can be written as

$$|B_q^0(t)\rangle = g_+(t)|B_q^0\rangle + \frac{q}{p}g_-(t)|\bar{B}_q^0\rangle$$
$$|\bar{B}_q^0(t)\rangle = g_+(t)|\bar{B}_q^0\rangle + \frac{q}{p}g_-(t)|B_q^0\rangle$$

where $g_\pm$ gives the amplitude that the initial particle is the same flavour (g_+) or has changed flavour (g_-). The probability of finding a certain flavour is given by the amplitude squared,

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma_q t}}{2}\left[\cosh\left(\frac{\Delta\Gamma_q t}{2}\right) \pm \cos\left(\Delta m_q t\right)\right], \tag{1.8}$$

where Γ_q is the average decay width $\frac{\Gamma_q^H + \Gamma_q^L}{2}$. The probability $|g_\pm(t)|^2$ is plotted in Fig. 1.5, where the oscillations can be seen and the decrease in the amplitude is caused by the exponential decay of the mesons. If the mixing asymmetry of Eq. (1.8) is studied instead, one finds

$$\frac{|g_+(t)|^2 - |g_-(t)|^2}{|g_+(t)|^2 + |g_-(t)|^2} = \frac{\cos\left(\Delta m_q t\right)}{\cosh\left(\frac{\Delta\Gamma_q t}{2}\right)}, \tag{1.9}$$

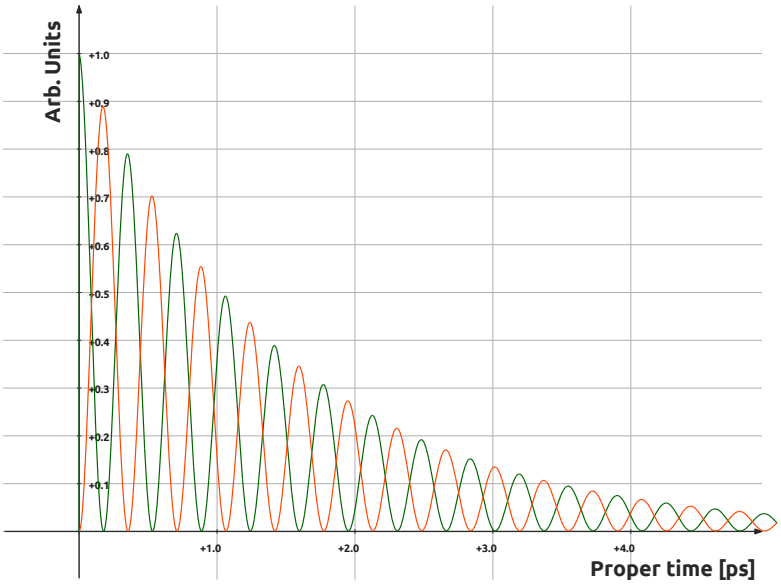

Figure 1.5: Plot showing Eq. (1.8), with realistic parameters of B_s^0 mixing. The red line shows $|g_+(t)|^2$, the probability of observing the initial B_s^0 flavour state, while the green line shows $|g_-(t)|^2$, the probability of observing the antiparticle of the initial state. The plot is taken from Ref. [25].

which removes the exponential decay component and more clearly shows the sinusoidal oscillations. The amplitude of this oscillation slowly decreases due to the hyperbolic cosine function. In the case of B_q^0 mixing the relative size of $\Delta\Gamma_q$ is small ($|\Delta\Gamma_d|/\Gamma_d < 0.18$ [24] at a 95 % confidence level (CL) and $\Delta\Gamma_s/\Gamma_s = 0.138 \pm 0.012$ [1]) and so the decrease in amplitude of the oscillation is often not detectable. However, the uncertainty of $\Delta\Gamma_q$ does contribute a small systematic uncertainty to the B_q^0 mixing analysis presented in Chapter 6.

1.3 Lepton flavour violation

In this section a general overview of lepton-flavour violation (LFV) is given and then specific theories which predict $D^0 \to e^{\pm}\mu^{\mp}$ are discussed. Noether's Theorem [26] states that there is a conservation law associated to every symmetry of a Lagrangian system. For example, time translation symmetry leads to the conservation of energy, space translation symmetry leads to the conservation of momentum and rotational symmetry leads to the conservation of angular momentum. In the SM there is no symmetry which leads to the conservation of lepton number or lepton flavour and so it is not a fundamental symmetry of the SM. Instead, conservation of these quantities is caused by an accidental symmetry, where no terms are included in the SM Lagrangian which could violate lepton number or lepton flavour. Since there is not a fundamental symmetry forbidding LFV, many extensions to the SM include terms that cause LFV.

Neutrino mixing in the νSM violates the conservation of lepton flavour and allows CLFV through neutrino oscillations in loops. Such loops are highly suppressed by a factor of $\left(\Delta m_{ij}/m_{W^{\pm}}\right)^4$, where $m_{W^{\pm}}$ is the $W^{\pm}$ mass and Δm_{ij} is the difference between the masses of the i-th and j-th neutrino mass eigenstates.

One of the most precisely probed CLFV decays is $\mu^{\pm} \to e^{\pm}\gamma$. In the SM this decay is forbidden, but in the νSM, $\mu^{\pm} \to e^{\pm}\gamma$ can occur at very low rates. Given the high suppression factor, the branching fraction of this decay is expected to be $\mathcal{B}\left(\mu^{\pm} \to e^{\pm}\gamma\right) \sim 10^{-54}$ [27]; this is well below experimental sensitivities, where the world's current best measurement by the Mu to E Gamma (MEG) experiment has determined that

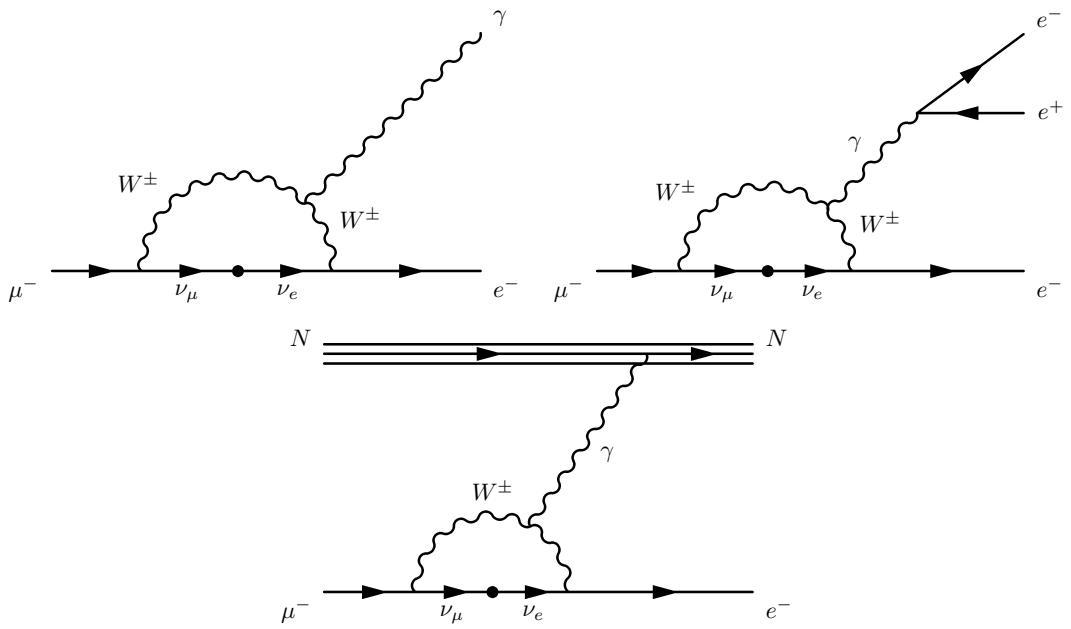

Figure 1.6: Three Feynman diagrams showing possible ways in which CLFV decays could occur. The three diagrams show the decays: $\mu^{\pm} \to e^{\pm}\gamma$ (top left), $\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp}$ (top right) and $\mu^{\pm} + N \to e^{\pm} + N$ (bottom). In all of these diagrams the ν_μ neutrino mixes to become a ν_e.

$\mathcal{B}\left(\mu^{\pm} \to e^{\pm}\gamma\right) < 5.7 \times 10^{-13}$ [28] at a 90 % CL. Other CLFV decays could occur at similar levels within νSM and so any detection of CLFV is a clear sign of new physics beyond the νSM. Many new physics models enhance the branching fractions of CLFV decays to the point that they reach the experimental sensitivity. In addition to $\mu^{\pm} \to e^{\pm}\gamma$ there are two other muonic CLFV decays that are often searched for and these are $\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp}$ and $\mu^{\pm} + N \to e^{\pm} + N$, where the last decay occurs while the muon is captured by a nucleus (N). These decays can occur through very similar diagrams in the νSM and these are shown in Fig. 1.6. The experimental limits placed on $\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp}$ are $\mathcal{B}\left(\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp}\right) < 1.0 \times 10^{-12}$ [29] at a 90 % CL and for $\mu^{\pm} + N \to e^{\pm} + N$ the limit is slightly more stringent with a value of $\mathcal{B}\left(\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp}\right) < 7.0 \times 10^{-13}$ [30] at a 90 % CL.

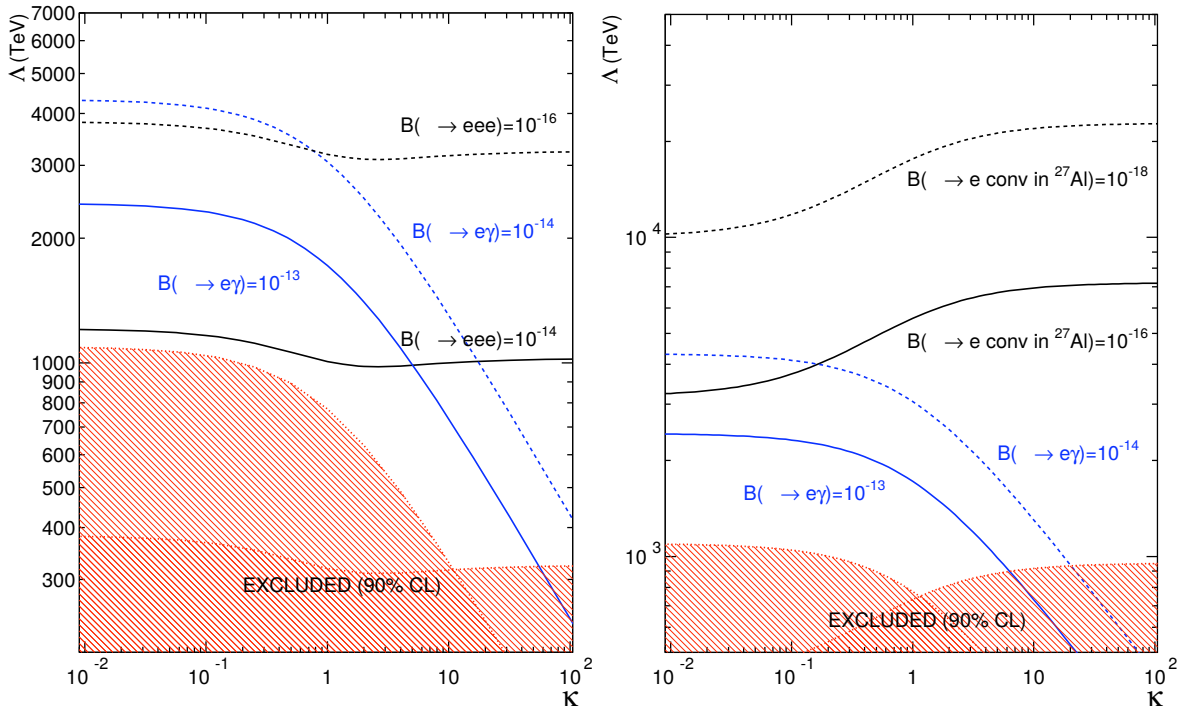

Figure 1.7: The plots show regions on the $\Lambda - \kappa$ plane that are excluded by searches for CLFV processes. The left plot compares $\mu^{\pm} \to e^{\pm}\gamma$ and $\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp}$, while the right plot compares $\mu^{\pm} \to e^{\pm}\gamma$ and $\mu^{\pm} + N \to e^{\pm} + N$. Measurements from 2012 are used to exclude the regions in this plot; this means that the previous measurement of $\mathcal{B}\left(\mu^{\pm} \to e^{\pm}\gamma\right) < 2.4 \times 10^{-12}$ [31] at a 90 % CL was used. The plots are taken from Ref. [32].

These three muonic CLFV decays can occur in many different physics models. The physics models can be categorised into two types: where the CLFV decay can occur at tree-level by adding a contact term into the Lagrangian and where the CLFV decay occurs at the loop-level by adding a dipole term into the Lagrangian. Generically new physics models involving CLFV decays can be modelled using two parameters: the mass scale of the new physics model (Λ) and the interaction type (κ). Large values of κ indicate that CLFV decays are primarily caused by the contact term in the Lagrangian, while small values of κ indicate that the CLFV decays occur

mainly through the dipole term. Figure 1.7 shows the regions on the $\Lambda - \kappa$ plane that are excluded by searches for CLFV processes. Since this plot was created the limit on $\mathcal{B}(\mu^{\pm} \to e^{\pm}\gamma)$ has become more stringent and with this updated limit all values of κ have been excluded up to $\Lambda \sim 1000\,\text{TeV}$. It is thought that with the current search methods in the MEG experiment, $\mathcal{B}(\mu^{\pm} \to e^{\pm}\gamma)$ has an ultimate experimental sensitivity of $\mathcal{O}(10^{-14})$, while the limit on $\mathcal{B}(\mu^{\pm} \to e^{\pm}e^{\pm}e^{\mp})$ could be pushed down further to $\mathcal{O}(10^{-16})$. There is no foreseen maximum experimental sensitivity in the $\mu^{\pm} + N \to e^{\pm} + N$ channel until $\mathcal{O}(10^{-18})$ and so an experiment to study $\mu^{\pm} + N \to e^{\pm} + N$ is planned to replace the MEG experiment. These various limitations in sensitivity are also shown in Fig. 1.7.

The search for $D^0 \to e^{\pm}\mu^{\mp}$ is very similar to the searches for $B_d^0 \to e^{\pm}\mu^{\mp}$, $B_s^0 \to e^{\pm}\mu^{\mp}$ and $K_L^0 \to e^{\pm}\mu^{\mp}$ and so the processes through which these could occur are also similar. The D^0 decay mode involves down-type quarks in the loop of the box diagram (as shown in Fig. 1.8), while the equivalent K^0 and

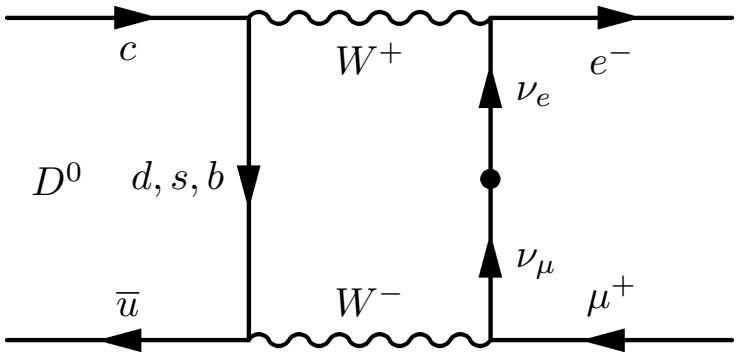

Figure 1.8: A Feynman diagram showing a possible way in which a $D^0 \to e^{\pm}\mu^{\mp}$ decay could occur. In this diagram the ν_μ neutrino mixes to a ν_e in the same way as the $\mu^{\pm} \to e^{\pm}\gamma$ decay; this highly suppresses the rate of this decay in the νSM.

B_q^0 decay modes involve up-type quarks. The large mass of the top quark leads to poor Glashow-Iliopoulos-Maiani (GIM) cancellations in box diagrams involving up-type quarks and so the K and B sectors have relatively large predicted values for mixing and rare decays such as $b \to s l^+ l^-$. The larger values of these quantities has led to more attention being focused on the K and B sectors, while similar processes in the D sector have been less exploited due to their smaller predicted values. However, extensions to the SM often spoil the effective GIM cancellations in the D sector and give significant enhancements over the SM values.

Searches for new physics in the D sector are also complementary to the searches in the K and B sectors as they are the only mixing processes which can probe the dynamics of up-type quarks. In some extensions to the SM, such as the R-parity violating MSSM (RPV-MSSM) (as discussed in Section 1.3.3), it is possible that the R-parity violation (RPV) effects are limited to only the up-type or only the down-type quarks [33]. This makes it vital to probe the dynamics of both up-type and down-type quarks in mixing type processes.

There is a very stringent limit on $\mathcal{B}\left(K_L^0 \to e^\pm \mu^\mp\right) < 4.7 \times 10^{-12}$ [34] at a 90 % CL from the E871 Collaboration. In the B sector, the LHCb collaboration has published the results of the search for $B_d^0 \to e^\pm \mu^\mp$ and $B_s^0 \to e^\pm \mu^\mp$, which found $\mathcal{B}\left(B_d^0 \to e^\pm \mu^\mp\right) < 1.1 \times 10^{-8}$ [35] at a 90 % CL and $\mathcal{B}\left(B_s^0 \to e^\pm \mu^\mp\right) < 2.8 \times 10^{-9}$ [35] at a 90 % CL. Before the measurement published in this thesis, the world's best experimental limit on $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ was produced by the Belle collaboration, who found $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) < 2.6 \times 10^{-7}$ [36] at a 90 % CL. A limit on $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ that was of the same order as the limits

on $B_d^0 \to e^\pm\mu^\mp$ and $B_s^0 \to e^\pm\mu^\mp$ would help to further constrain new physics models.

As with other CLFV decays, in the SM the decay $D^0 \to e^\pm\mu^\mp$ is forbidden; however, in the νSM the decay can occur through the diagram shown in Fig. 1.8 and so the branching fraction is non-zero, but still far beyond current experimental sensitivities. There are several extensions to the νSM which enhance the branching fraction of $D^0 \to e^\pm\mu^\mp$ to possibly detectable rates and these extensions are discussed in the following sections. Depending on the theory being studied, the predictions of the branching fraction vary by many orders of magnitude.

1.3.1 Unparticles and the Z'

The measured amount of D^0 mixing may be larger than the amount predicted by the SM [37]. If it is assumed that the D^0 mixing is caused primarily by new physics, the existing D^0 mixing measurements can be applied as constraints to the amount of new physics. The unparticle [38] can couple to different flavours, and so it allows flavour-changing neutral currents (FCNC) to occur at tree-level. The tree-level unparticle mediated $D^0 \to e^\pm\mu^\mp$ decay is shown in Fig. 1.9. This would alter the mixing of the D^0, and it would also create couplings through which a $D^0 \to e^\pm\mu^\mp$ decay could occur. Assuming that the D^0 mixing is primarily caused by the unparticle couplings, one finds that $\mathcal{B}\left(D^0 \to e^\pm\mu^\mp\right) < 1.0 \times 10^{-22}$ [37]. Another similar possibility could be for a non-universal gauge boson, Z', to create similar tree-level FCNC [39]. Assuming that the D^0 mixing is primarily caused by the new boson couplings, one finds that $\mathcal{B}\left(D^0 \to e^\pm\mu^\mp\right) < 7.9 \times 10^{-20}$ [37]. Both of these predictions

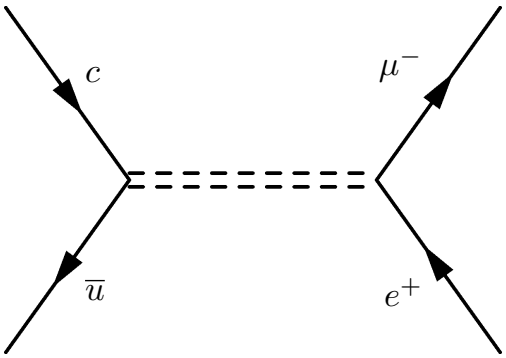

Figure 1.9: A Feynman diagram showing a $D^0 \to e^{\pm}\mu^{\mp}$ decay mediated by an unparticle. The double dashed line represents the unparticle.

are orders of magnitude above the νSM prediction, but many times smaller than the current experimental sensitivity.

1.3.2 Extra leptons

If either an additional lepton doublet or an extra neutral lepton singlet are added to the νSM, then once again the branching fraction of $D^0 \to e^{\pm}\mu^{\mp}$ could be enhanced. The Feynman diagram for this process is very similar to the νSM diagram, except the neutrino undergoes the transition $\nu_\mu \to \nu_L \to \nu_e$, where ν_L is the extra neutrino added to the νSM; this is shown in Fig. 1.10. The invisible decay width of the Z^0-boson is compatible with only three light neutrino flavours [40] and so if the additional neutrino was lighter than $\frac{1}{2}m_{Z^0}$ it couldn't couple to the Z^0. To avoid this extra complication the mass of the extra neutrino is assumed to be large and is taken to be $m_N \simeq 50\,\mathrm{GeV\,c^{-2}}$. Using the experimental constraint placed on $\mathcal{B}\left(\mu^{\pm} \to e^{\pm}\gamma\right)$ as an input, a limit can be placed on the amount of mixing that occurs between the new heavy neutrino and the first two generations of neutrinos (ν_e and ν_μ). Ref. [33] (published in 2002) has calculated that $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right) < 1.0 \times 10^{-14}$ using

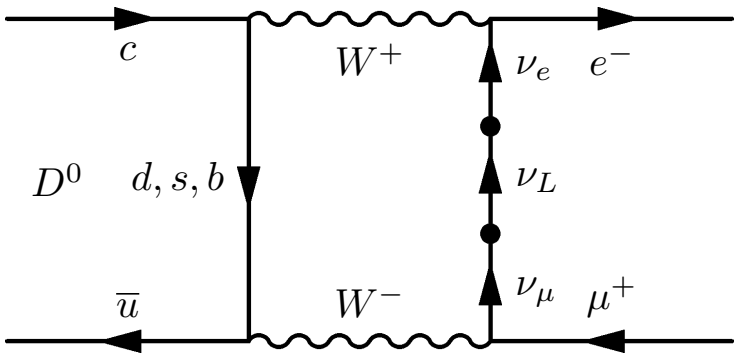

Figure 1.10: A Feynman diagram showing a possible $D^0 \to e^\pm\mu^\mp$ decay mode, where the neutrino under goes a $\nu_\mu \to \nu_L \to \nu_e$ transition, where ν_L is an extra neutrino that is added to the νSM.

$\mathcal{B}(\mu^\pm \to e^\pm\gamma) < 1.2 \times 10^{-11}$ from Ref. [41]. Since these publications the limit placed on $\mathcal{B}(\mu^\pm \to e^\pm\gamma)$ has become much more stringent and so it is likely that this would lead to a more stringent constraint on $\mathcal{B}(D^0 \to e^\pm\mu^\mp)$. However, even this relatively weak constraint is still beyond the expected experimental sensitivity to $\mathcal{B}(D^0 \to e^\pm\mu^\mp)$ using LHCb data.

1.3.3 *R*-parity violation in the MSSM

The most widely discussed [33, 42, 43] new physics model which predicts $D^0 \to e^\pm\mu^\mp$ is an R-parity violating MSSM. In the MSSM a new quantum number known as R-parity is conserved. R-parity is defined as $P_R = (-1)^R = (-1)^{3B+L+2s}$, where B and L are the baryon and lepton quantum numbers and s is the spin. All SM particles have an R-parity of +1, while the super-symmetric particles have an R-parity of -1. Masses of particles in the MSSM have been experimentally excluded up to large values and so extensions to the MSSM are now being studied including theories with RPV.

In the RPV-MSSM, the decay $D^0 \to e^\pm \mu^\mp$ is predicted to occur at relatively large rates (when compared to the other theoretical predictions) via the diagram shown in Fig. 1.11. There are several predictions of $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ in the RPV-MSSM that are calculated using different methods and these vary by two orders of magnitude. In Ref. [33] a limit is placed on the branching fraction $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) < 1.0 \times 10^{-6}$ using model-independent constraints on single RPV couplings. However, there is a product of two RPV couplings used to calculate the branching fraction of $D^0 \to e^\pm \mu^\mp$ in the RPV-MSSM. The combination of these two RPV couplings can also be constrained using experimental data, specifically the two limits from the BaBar collaboration of $\mathcal{B}\left(D^\pm \to \pi^\pm e^\pm \mu^\mp\right) < 2.9 \times 10^{-6}$ [44] at a 90 % CL and $\mathcal{B}\left(D_s^\pm \to K^\pm e^\pm \mu^\mp\right) < 1.4 \times 10^{-5}$ [44] at a 90 % CL. Reference [42] performed the study which constrained the product of the two RPV couplings using the limit on $\mathcal{B}\left(D^\pm \to \pi^\pm e^\pm \mu^\mp\right)$ to find $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) < 1.8 \times 10^{-7}$, while Ref. [43] found $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) < 2.6 \times 10^{-8}$ using the RPV couplings constrained with the limit on $\mathcal{B}\left(D^\pm \to \pi^\pm e^\pm \mu^\mp\right)$. The limit placed on $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ by Belle is in the middle of these predictions and so a further measurement of

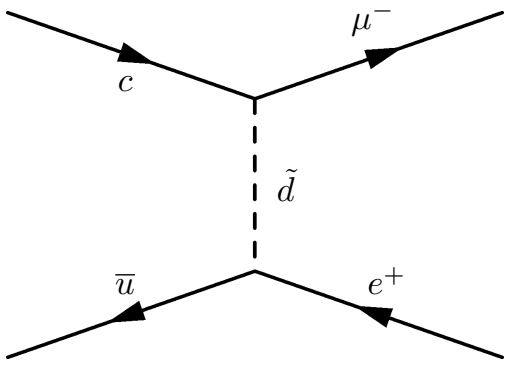

Figure 1.11: A possible Feynman diagram showing a way in which a $D^0 \to e^\pm \mu^\mp$ decay could proceed within the framework of an RPV-MSSM model.

$\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$ could place highly stringent limits on RPV couplings in the MSSM.

CHAPTER 2

The LHC and the LHCb experiment

The large hadron collider beauty (LHCb) experiment detects particles created from proton-proton collisions. The protons are collided by the Large Hadron Collider (LHC), which is the largest and most energetic particle collider in the world. The LHC is a circular synchrotron with a 27 km circumference and is built at CERN, the world's largest particle physics research laboratory. This chapter will first discuss the LHC and then describe the LHCb experiment.

2.1 The Large Hadron Collider

The LHC is the last accelerator in a chain of accelerators as shown in Fig. 2.1. The tunnel in which the LHC is situated crosses the French-Swiss border near Geneva and is nominally 100 m underground as pictured in Fig. 2.2. The same tunnel previously contained the Large Electron–Positron (LEP) collider;

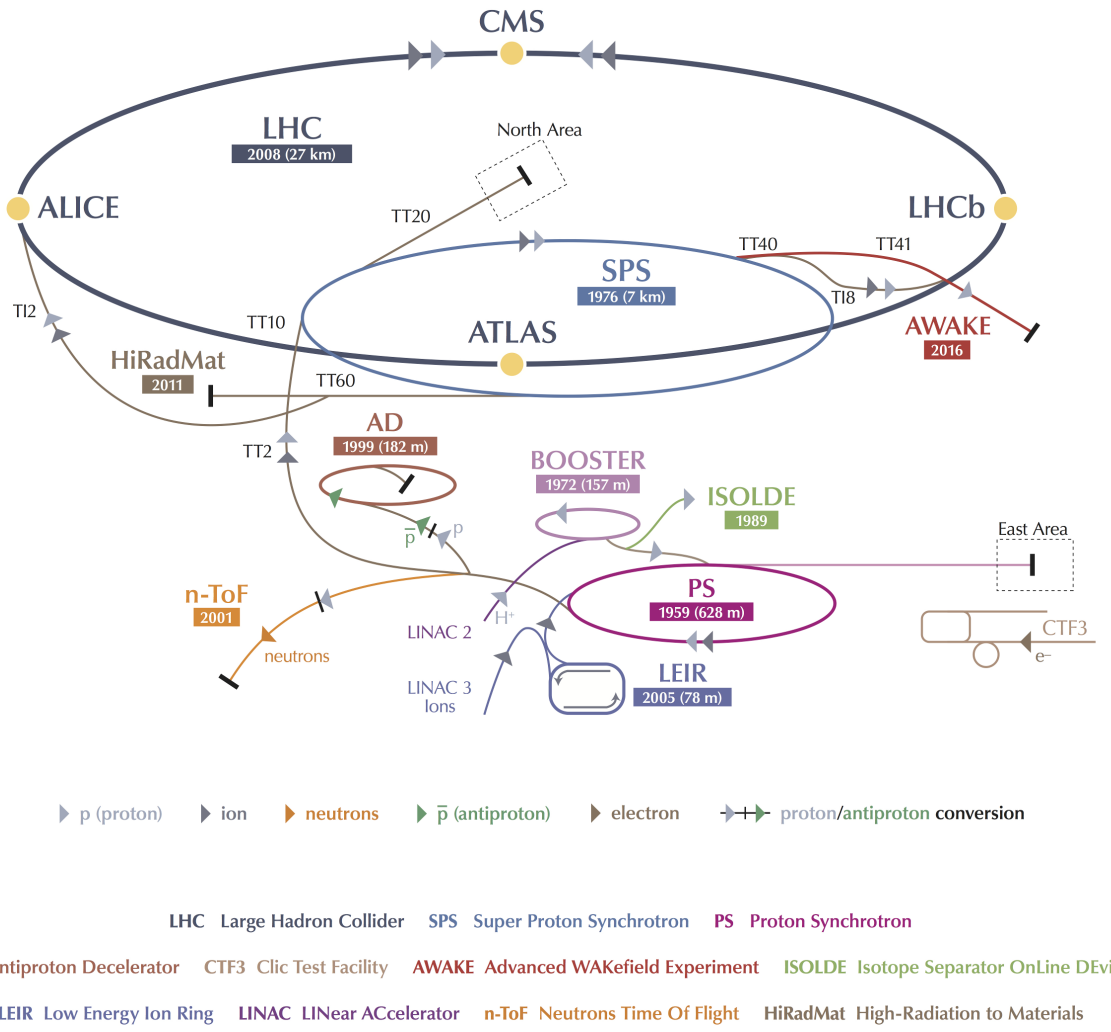

Figure 2.1: Diagram of the accelerator complex at CERN (not to scale). The figure is taken from Ref. [45].

this was an accelerator that provided electron–positron collisions to four experiments designed to test the standard model. The process of colliding protons starts by accelerating bunches of protons to 50 MeV using a linear accelerator called LINAC2. These protons are then transferred to a series of synchrotrons where the energy of the protons is increased to 1.4 GeV, 25 GeV and 450 GeV by the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), respectively. From the SPS the protons are then transferred to the LHC where they are

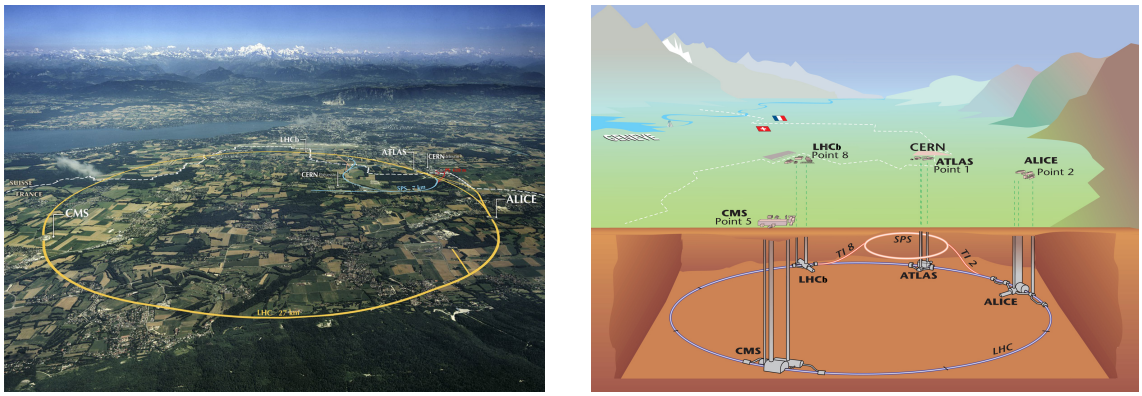

Figure 2.2: Aerial photo with the position of the LHC overlaid (left) and a diagram of the LHC ring (right). In both images the positions of the four major experiments are shown along with the airport, which can be seen next to the LHCb surface buildings. The images are taken from Refs. [46, 47].

accelerated to their maximum energy. They are injected into the LHC in two oppositely rotating rings, which will be collided after reaching the desired center-of-mass energy ($\sqrt{s}$). The design center-of-mass energy of the LHC is 14 TeV, but due to issues with the connections between the LHC dipole magnets, this energy has not been reached.

Nominally, the LHC runs for approximately 200 days per year. Each year there is usually a shut down of the LHC over the winter period, when electricity is in higher demand in Europe. The LHC takes data for several consecutive years (known as runs) and then has a long shutdown to allow for repairs and upgrades to both the detectors and the accelerators. A full list of all of the LHC run periods are shown in Table 2.1.

During Run 1 of the LHC, the center-of-mass energy of the beams was changed; in 2010 and 2011 the beams had a center-of-mass energy of 7 TeV, but this was increased to 8 TeV in 2012. During Long Shutdown 1 (LS1),

Table 2.1: The start and end year of each previous and planned period of the LHC run schedule. This is the current plan as of 2015 and is subject to change. Long shutdown is abbreviated to LS.

Period	Run 1	LS1	Run 2	LS2	Run 3	LS3	Run 4	LS4	Run 5
Start	2010	2013	2015	2018	2021	2023	2026	2030	2032
End	2013	2015	2018	2021	2023	2026	2030	2032	2035

many of the connections between the dipoles were repaired, which allows $\sqrt{s}$ to be increased to 13 TeV for Run 2 of the LHC.

After accelerating the proton beams to the desired center-of-mass energy, the beams are brought together at four collision points, where the experiments are situated. There are two general purpose detectors (GPDs), A Toroidal LHC Apparatus (ATLAS) [48] and Compact Muon Solenoid (CMS) [49]; it is these two experiments which announced the observation of a new boson in 2012 [50, 51]. The LHC was primarily designed to accelerate protons, but also accelerates lead ions to 2.76 TeV per nucleon. One experiment, A Large Ion Collider Experiment (ALICE) [52], is designed specifically to measure the ion collisions from the LHC, with an aim to measure the properties of the quark gluon plasma (QGP). Finally the last of the major experiments is the large hadron collider beauty (LHCb) experiment [53], which is a forward single-arm spectrometer designed primarily to measure CPV and rare decays of particles containing b and c quarks. Three other smaller experiments are placed around the interaction regions: the Monopole and Exotics Detector at the LHC (MOEDAL), the large hadron collider forward (LHCf) experiment and the total, elastic and diffractive cross-section measurement (TOTEM) experi-

ment, which respectively aim to detect monopoles [54], model high energy cosmic rays [55] and measure the total proton-proton cross-section [56].

An important measure of the performance of the LHC is the instantaneous luminosity ($\mathcal{L}$), which determines the rate of proton-proton interactions. The design value of the LHC is $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Since the proton beam is grouped into bunches, the instantaneous luminosity can be increased in several different ways: by increasing the number of colliding bunches, by increasing the number of particles in each bunch, by decreasing the physical size of the bunches (to increase the density of particles) or by increasing the revolution frequency. Nominally the LHC was designed to run with a bunch collision rate of 40 MHz which leads to a bunch spacing of 25 ns; however, in Run 1 the

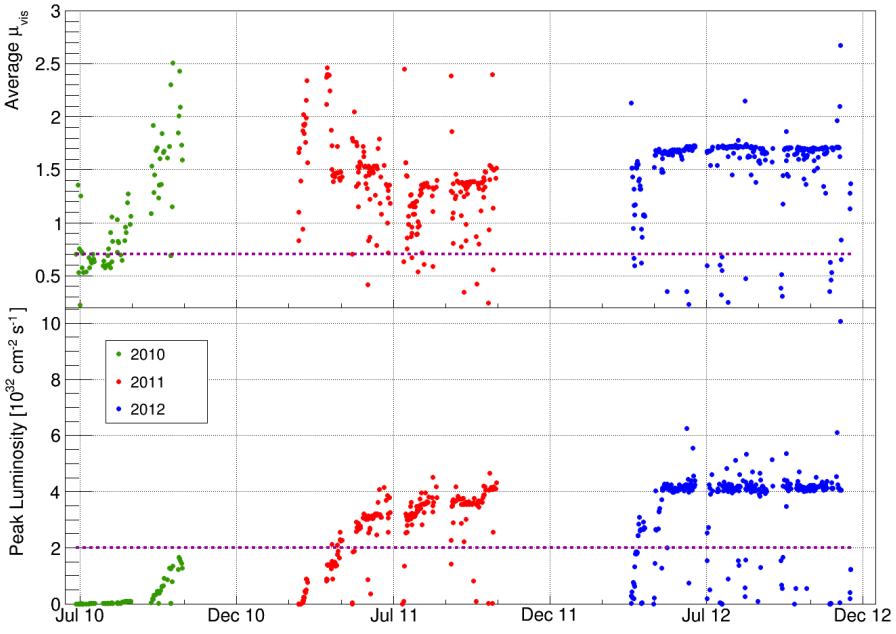

Figure 2.3: Average number of visible interactions per bunch crossing (top) and peak luminosity in each fill (bottom), both as a function of time. The design values of these parameters are highlighted by the dashed purple line. The figures are taken from Ref. [57].

LHC ran with a bunch spacing of 50 ns. To maximise luminosity, LHCb ran with a high number of collisions per bunch crossing. Events which contain multiple *pp*-collisions also contain more tracks and so are computationally more difficult to reconstruct. LHCb was designed to run with a mean number of visible *pp*-interactions per bunch crossing (μ) of about 0.7. Figure 2.3 shows the average μ over time, which during 2011 and 2012 peaked at 2.5; this is over four times the design value. When compared to the GPDs, the luminosity received by LHCb is significantly lower. The luminosity is lowered by laterally separating the beams. As the number of particles in the beams is reduced (due to the collisions), the beams are bought closer together to

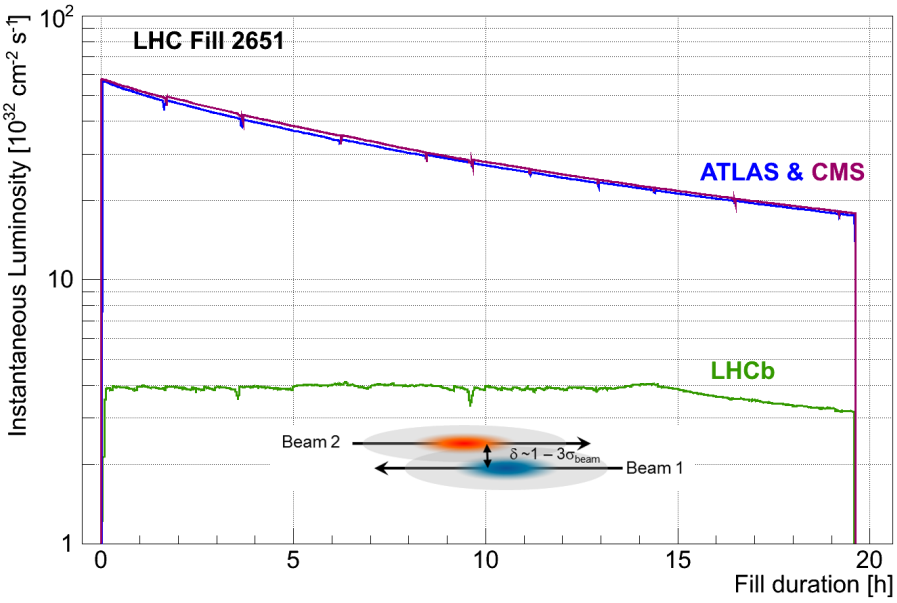

Figure 2.4: Instantaneous luminosity at the collision points for ATLAS, CMS and LHCb. At the bottom of this figure is a diagram indicating how the beams are separated at the LHCb collision point to reduce the instantaneous luminosity; usually the beams are separated by 1 to 3 standard deviations of the beam width (σ_{beam}). The figure is taken from Ref. [57].

keep the luminosity constant over time; this is known as luminosity levelling and is illustrated in Fig. 2.4. During 2011 the average luminosity at LHCb was steadily increased until it reached about $3.5 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, but in 2012 the average luminosity was approximately constant at a value of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

2.2 The LHCb experiment

The LHCb experiment aims to study heavy flavour physics, specifically CPV and rare decays of b and c quarks. This is particularly interesting as these types of measurements often have very precise predictions from the SM and so any measured deviations from the values predicted by the SM can indicate new physics. Due to the large $b\overline{b}$ and $c\overline{c}$ production cross-sections at the LHC, around 10^{12} b and c decays have been recorded [57]; this makes the LHC a great place to study this type of physics.

The LHCb experiment consists of a series of detectors placed to detect particles coming from the collision point at angles from the beam axis between 10 mrad and 300 mrad; this will be referred to as the angular acceptance of the detector. The experiment consists of several discrete detectors which are shown in Fig. 2.5. Starting from the collision point the detectors are: the VELO, which is a silicon-strip detector that surrounds the collision point; the first of two ring imaging cherenkov (RICH) detectors, used for identification of low-momentum particles; the tracker turicensis (TT), a silicon strip detector; the magnet, with a bending power of about 4 T m; the Inner Tracker (IT) and outer tracker (OT) of the three tracking stations (T1-T3), which use silicon-strip and straw tube technologies

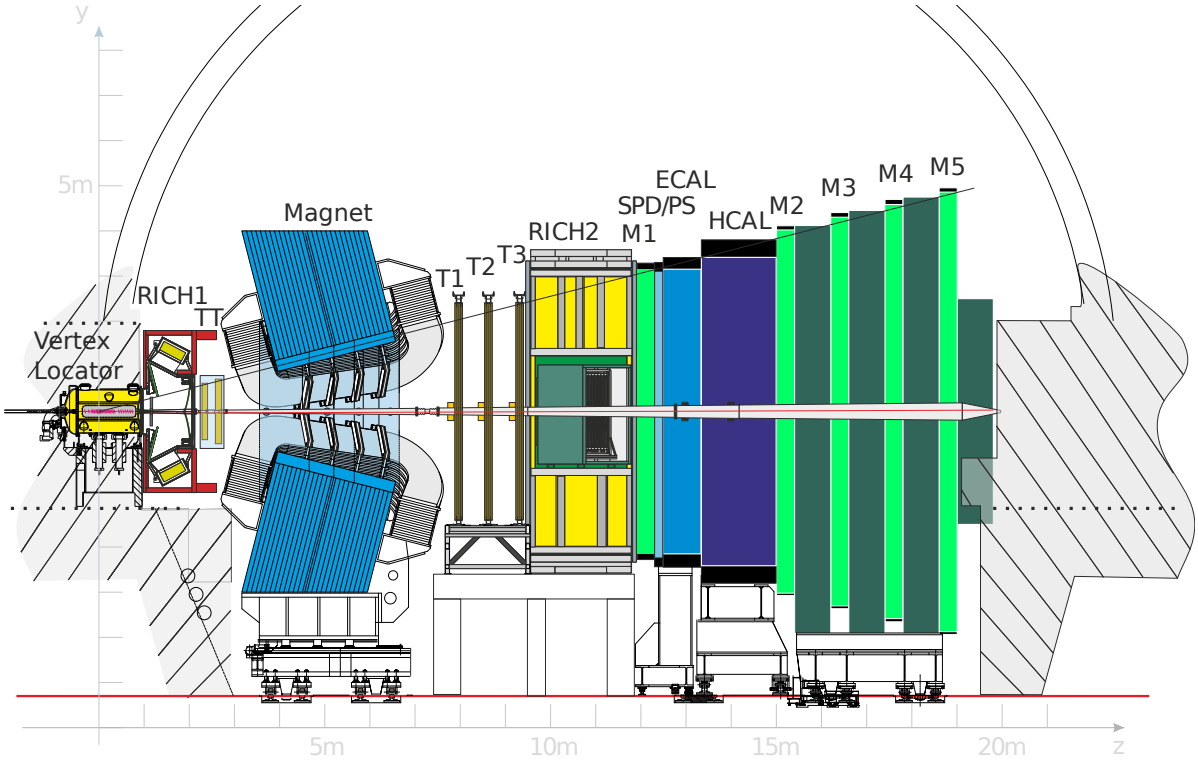

Figure 2.5: Layout of the LHCb experiment inside underground experimental area, with the various detectors labelled. The diagram is taken from Ref. [58].

respectively; the second RICH detector, used for identification of higher-momentum particles; the first of five muon stations (M1-M5), which rely on multi wire proportional chamber (MWPC) technology except for the innermost region of M1 where gas electron multiplier (GEM) detectors are used; the scintillating pad detector (SPD) and pre-shower (PS) detector, which use scintillating pads to reduce backgrounds in the next detector; an electromagnetic calorimeter (ECAL), which measures the energy of electrons and photons; the scintillating-tile iron-plate hadronic calorimeter (HCAL), which measures the energy of hadrons; and finally the remainder of the muon stations (M2-M5).

The majority of the detectors have two halves, which can be separated to ease assembly and maintenance. The two sides are referred to as the A and C-sides, where the C-side is closer to the centre of the LHC. A right-handed coordinate system is defined with z along the beam axis into the detector, y in the vertical direction and x in the horizontal direction, as shown in Fig. 2.5. Where appropriate, spherical coordinates (r,θ,ϕ) are also used, but with the polar angle (θ) component changed to pseudorapidity (η) via the transformation

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] . \tag{2.1}$$

The LHCb detectors have an angular coverage of 10 mrad to 300 mrad in θ; this corresponds to a range in η of $2 < \eta < 5$. It is this unique geometry

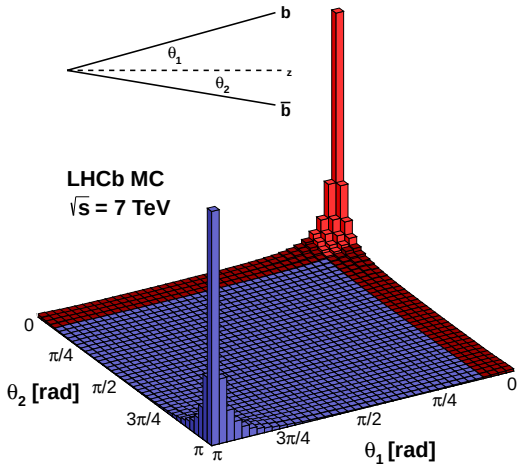

Figure 2.6: Relative $b\bar{b}$ production intensity as a function of the angle between each of the quarks and the beam axis as simulated with Pythia [59]. The $b\bar{b}$-pairs are mostly produced in a cone around the beampipe in the forwards or backwards directions. The figure is taken from Ref. [53].

which sets LHCb apart from the GPDs, whose tracking detectors only cover approximately $-2.5 < \eta < 2.5$. The geometry of LHCb only covers 4 % of the solid angle, but 40 % of the produced $b\bar{b}$-pairs are expected to fall in this region; this can be seen in Fig. 2.6.

2.3 Charged particle tracking

The tracking system consists of the VELO, the TT, the magnet and the three tracking stations (T1-T3). Each of the three tracking stations contains a region of high resolution sensors close to the beampipe known as the IT and a region of lower resolution tracking detectors further from the beampipe known as the OT. The IT only covers 1.3 % of the sensitive surface of each tracking station, but about 20 % of all charged particles from the interaction region will pass through the IT. It is because of the very large the track multiplicity (the number of particles produced in the collision) in this region, that a higher resolution detector is needed. The TT and IT were developed in tandem in a project known as the silicon tracker (ST) [60].

2.3.1 Vertex Locator

The VELO is a silicon strip-detector consisting of 42 modules (as shown in Fig. 2.7) and its primary purpose is to accurately measure the positions of all vertices in the collisions [58, 62]. This includes primary pp interaction vertices (PVs) (created from the collisions of protons) and secondary vertices (SVs) (created from decays of short-lived particles). This is particularly important as the particles that LHCb is most interested in have lifetimes from 10^{-13} s to 10^{-12} s and so can travel around 1 cm before decaying. Finding

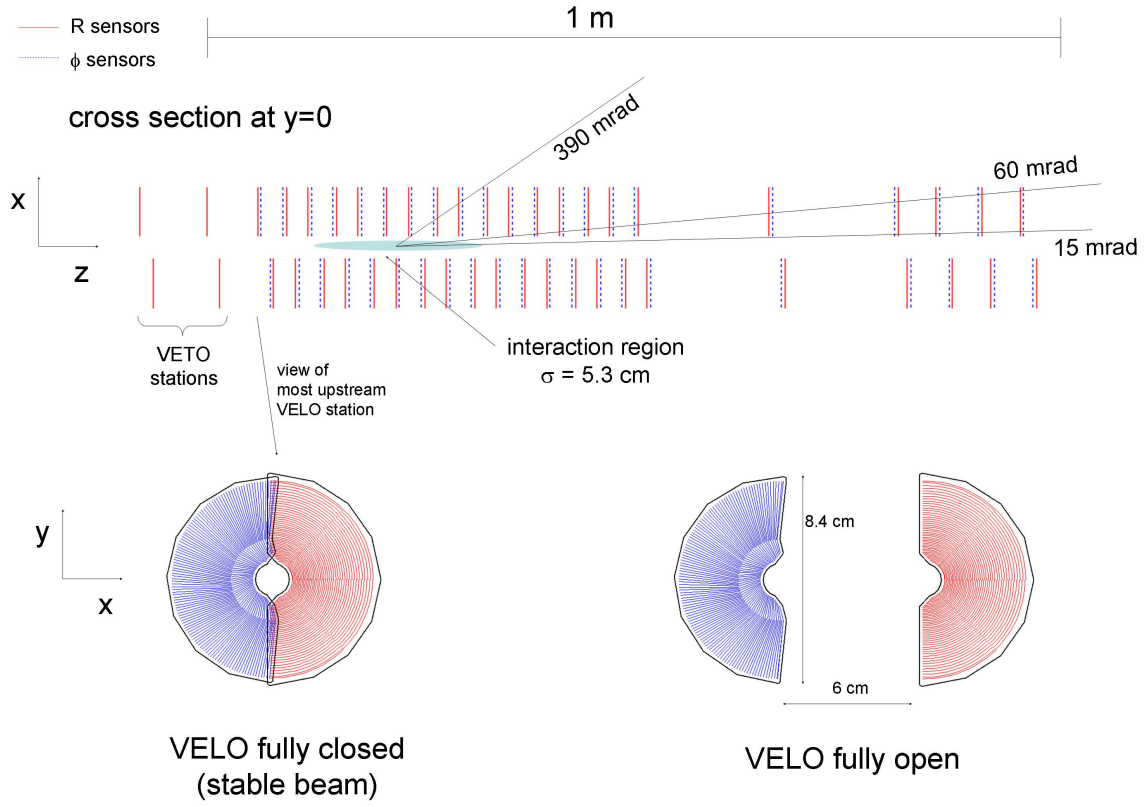

Figure 2.7: Layout of the sensors in the VELO (top) and diagram showing how the sensors overlap when closed and separate when open (bottom). The approximate luminous region is shown on the top diagram, along with the angular acceptance of the detector. The figure is taken from Ref. [61].

these secondary vertices is a simple way to determine which collisions produced particles containing *b* and *c* quarks. The SVs can be found quickly by finding reconstructed tracks with a large distance of closest approach to a PV; this distance is known as the impact parameter (IP). Because the IP is so important for finding SVs the IP resolution is a key performance measure of the VELO. A diagram of the VELO and a photo of a VELO module are shown in Fig. 2.8.

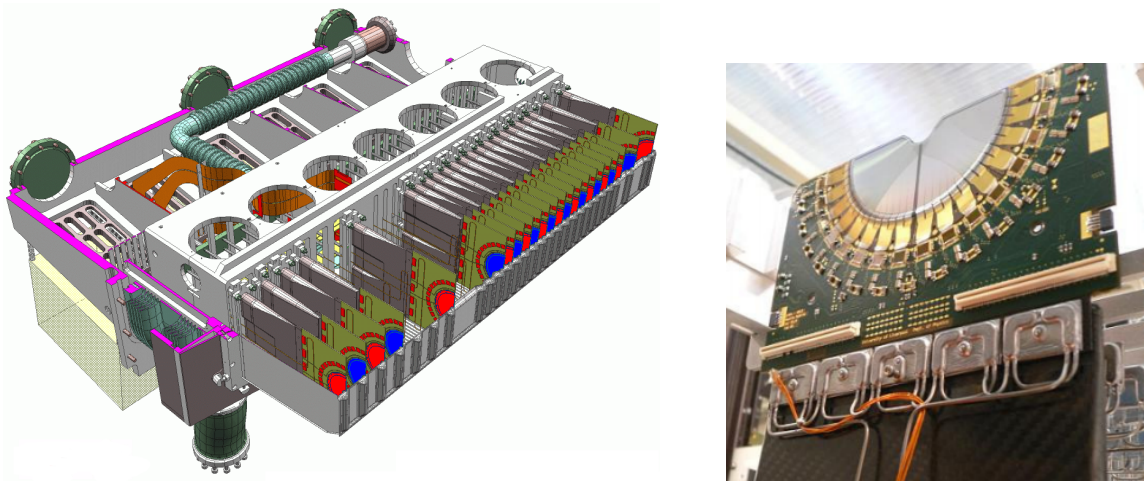

Figure 2.8: One half of the VELO cut open to show the inside of the detector (left) and a photo of a VELO module. The RF-foil and RF-box can be seen in the bottom right of the left hand diagram and then inside are the r (red) and ϕ (blue) semi-circular sensors. The figures are taken from Refs. [63, 64].

When a charged particle traverses a silicon sensor it deposits an electrical charge in the sensor. The charge deposited into the sensor can then be read out and a hit recorded. In strip detectors the sensor is segmented into strips such that two hits are needed, in perpendicular strips, to determine the position where the particle passed through the sensor. Each VELO module contains two silicon strip-based sensors, one with strips in the r-direction and the other in the ϕ-direction. The width of the strips varies from 38 µm at the inner edge to 102 µm at the outer edge of the sensor, which has a radius of 42 mm.

Nominally the VELO is around 6 cm from the beam, but when the beam stabilises the two sides of the detector (which are able to move independently) close tightly around the beam. This puts the first piece of active silicon just 8.2 mm from the beam. Between the beam and the sensors there is only a 300 µm thick corrugated foil which prevents radio frequency (RF)

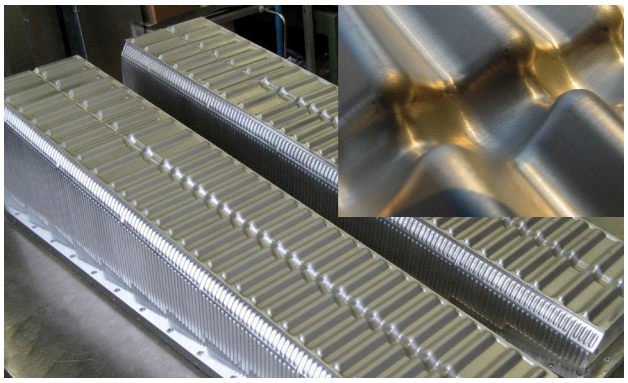

Figure 2.9: Images of the RF-foil [63]. The RF-foil and RF-box can be seen in the main image, while the inset image zooms into the central region where the corrugations can be more easily seen.

pickup in the modules and separates the LHC and the VELO vacuums; this foil is known as the RF-foil. The RF-foil has a complicated corrugated shape, which the modules are slotted into; it is pictured in Fig. 2.9. Because of its complicated shape, the RF-foil is hard to simulate, but it is the main contribution to the material traversed by particles in the VELO, and so it is very important to the simulated performance of the VELO. The modules are spread over a distance of about a meter and are cooled with evaporative carbon dioxide (CO_2) cooling.

2.3.2 Magnet

In LHCb the magnetic field is in the y-direction and so tracks of charged particles bend in the xz-plane. The radius of curvature of a track (r) in metres is given by

$$r = \frac{p\cos(\lambda)}{0.3zB},$$

where p is the particle's momentum in units of $\mathrm{GeV\,c^{-1}}$, λ is the angle between the particle's velocity vector and the y-axis, z is the charge of the particle in units of the elementary electric charge (e) and B is the magnetic field strength in tesla. The magnet used in LHCb is a warm dipole magnet and for particles that pass through the whole detector it has an integrated

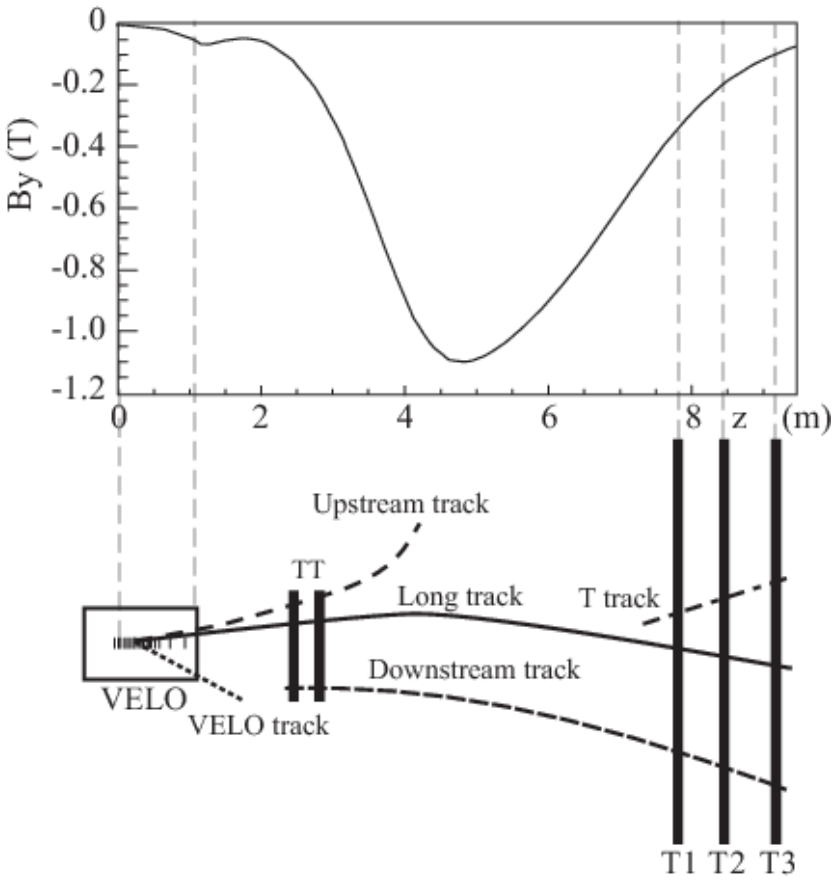

Figure 2.10: Magnetic field in the *y*-direction as a function of distance from the origin in the *z*-direction (top) compared to a top down illustration of the LHCb experiment (bottom). The figure is taken from Ref. [57].

magnet field of 4 T m [65]. Figure 2.10 shows the magnetic field strength in the *y*-direction as a function of distance from the origin along the *z*-direction, where it can be seen that there is minimal magnetic field in the VELO and after the tracking stations.

The polarity of the magnet can be reversed so that the field is parallel or antiparallel to the *y*-direction and this is known as magnet-up and magnet-down respectively. The magnet polarity is changed several times per year, such that a roughly equal amount of data is recorded in each configuration.

If one side of the detector was less efficient at tracking particles (for example due to a broken detector module) and the magnet polarity could not change, positive or negative particles would be preferentially detected. Changing the polarity of the magnet allows this effect to be negated and so is particularly useful for charge asymmetry measurements.

2.3.3 Silicon Tracker and Outer Tracker

The TT and the IT are silicon microstrip detectors [58, 60]. Simulations showed that a hit resolution of 50 µm in the TT and IT would lead to a momentum resolution which was dominated by multiple scattering. Following simulation studies a strip pitch of 200 µm was chosen to be used in these detectors. The hit resolution was measured in the 2011 data to be 52.6 µm and 50.3 µm for the TT and IT respectively [57]. Four layers of these strips are stacked on top of each other in a (x-u-v-x) layout; the first and last layers of strips are vertical with the second and third layers rotated from the vertical axis by small angles of $-5°$ and $5°$ respectively. The occupancy of a detector is the fraction of the sensitive region (in the case of the ST, it is the fraction of strips) which has a hit recorded in it. If the occupancy of strip detectors is too high, fake hits can be created because the strips can be grouped in different combinations to produce hits in many places. When reconstructing the position of the hit, this small stereo angle adds an additional constraint and so reduces the number of positions where there could be a genuine hit. It also allows the track position to be resolved very precisely in the x-direction, which results in a good momentum resolution.

The TT is placed in front of the magnet and covers the full angular acceptance of LHCb. It measures approximately 150 cm horizontally and

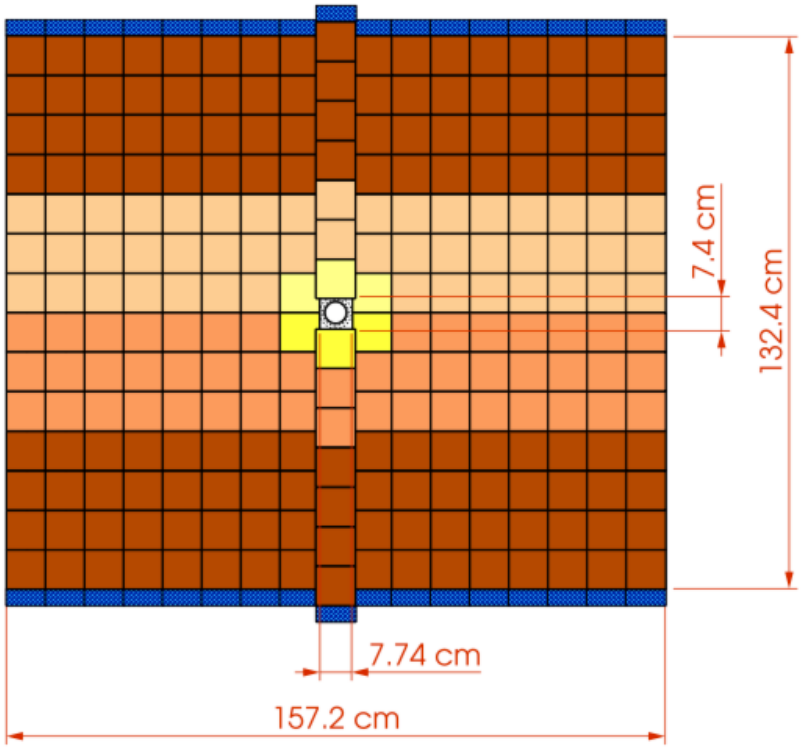

Figure 2.11: Position of TT sensors in the *x*-*y* plane. The circle in the middle shows the beampipe. The coloured regions show different readout sectors, while the readout hybrids are highlighted in blue. The diagram is taken from Ref. [61].

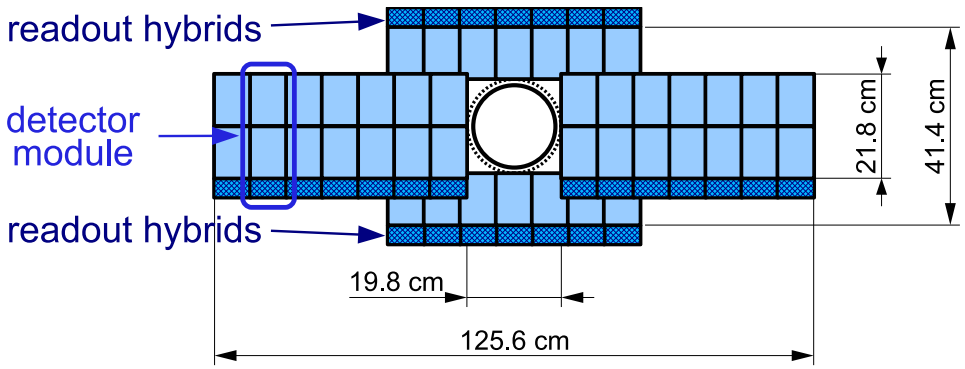

Figure 2.12: Position of IT sensors in the *x*-*y* plane. The circle in the middle shows the beampipe. Light blue shows the sensors and the dark blue shows the readout hybrids. The diagram is taken from Ref. [61].

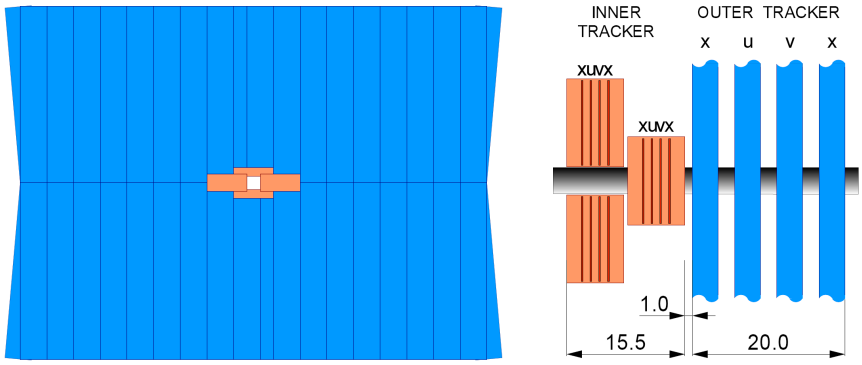

Figure 2.13: Relative positions of the IT (tangerine) and the OT (azure). The left diagram shows the IT and OT in the x-y plane, while the right diagram is in the y-z plane. Dimensions are given in cm. The diagrams are taken from Ref. [60].

130 cm vertically. The TT sensors are 9.64 cm by 9.44 cm and are placed as pictured in Fig. 2.11. There are 1.4×10^5 strips of up to 38 cm long. The four layers of the TT are separated into two pairs, the (x-u) pair and the (v-x) pair; these are placed 27 cm apart in the z-direction.

The IT is placed after the magnet and covers the region close to the beampipe, measuring approximately 120 cm by 40 cm in a cross-shape, as pictured in Fig. 2.12. The IT has 1.2×10^5 strips of 11 cm or 22 cm long. Each of the four (x-u-v-x) layers in the IT are placed next to one another and just in front of the OT, as illustrated in Fig. 2.13.

The OT is a gas straw-tube drift-time detector and is placed in the outer regions of the three tracking stations [66]. In gas straw-tube detectors charge is deposited in gas and a voltage is applied between the straw-tube and the central wire; the charge then drifts through the gas towards the wire (or straw-tube). The charged particles are then detected by measuring the current flowing through the straw-tube and central wire. Each station has

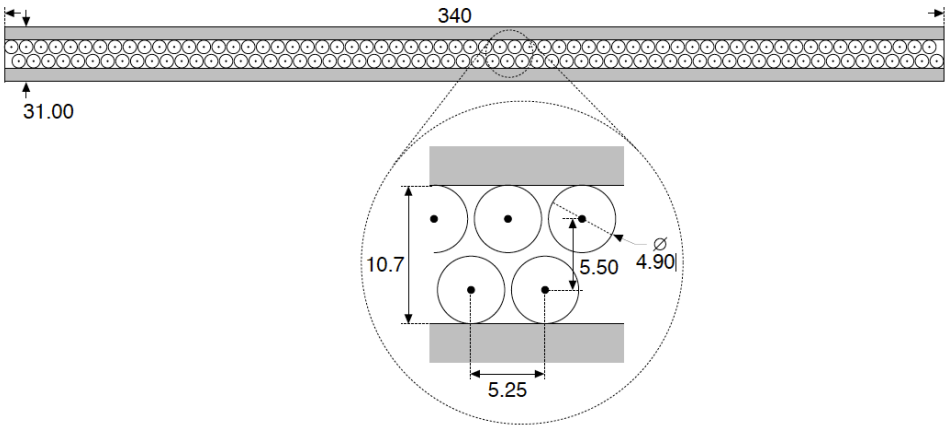

Figure 2.14: Cross-section of an OT module (top) and a section shown in greater detail (bottom). Dimensions are given in mm. The diagram is taken from Ref. [61].

four modules using the same (*x*-*u*-*v*-*x*) geometry as the ST. Each module consists of two staggered layers of drift-tubes with an inner diameter of 4.9 mm, as shown in Fig. 2.14. A mixture of Argon (70 %) and CO_2 (30 %) is used, which gives drift times of less than 50 ns. The active area of these stations is about 6.0 m by 4.9 m.

2.3.4 Tracking performance

The hits in the tracking detectors are combined to reconstruct tracks in the detector; an example event is shown in Fig. 2.15. This provides a relative momentum resolution, shown in Fig. 2.16, which varies from 0.5 % at low-momenta to 1.0 % at 200 GeV c^{-1}. This excellent momentum resolution allows LHCb to reconstruct invariant mass with a relative resolution of less than 0.5 % for masses smaller than 10 GeV c^{-2}.

Figure 2.17 shows the IP resolution as a function of inverse transverse momentum (p_T) and the PV position resolution in the *x*-direction for events

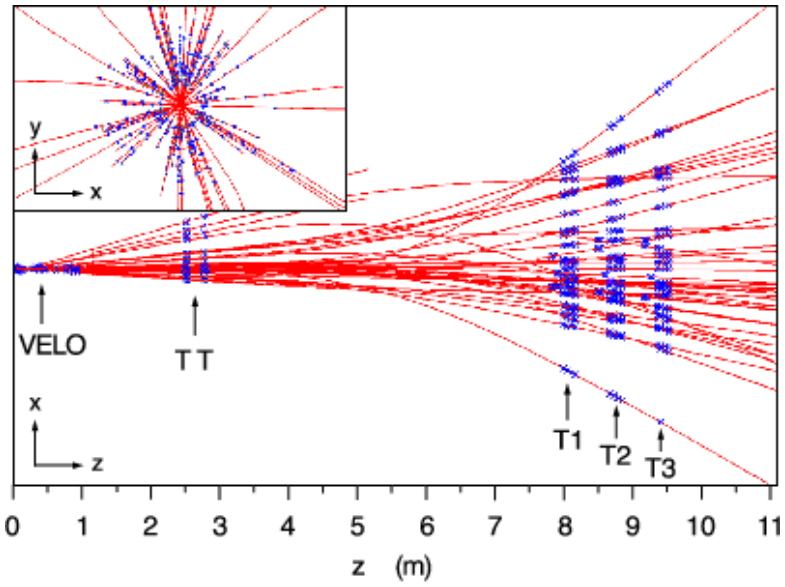

Figure 2.15: Example hits from the tracking detectors are reconstructed into tracks. The red lines show the reconstructed tracks, while the blue crosses show the reconstructed hits; the inset image shows the *x*-*y* plane and the main plot shows the *x*-*z* plane. The figure is taken from Ref. [61].

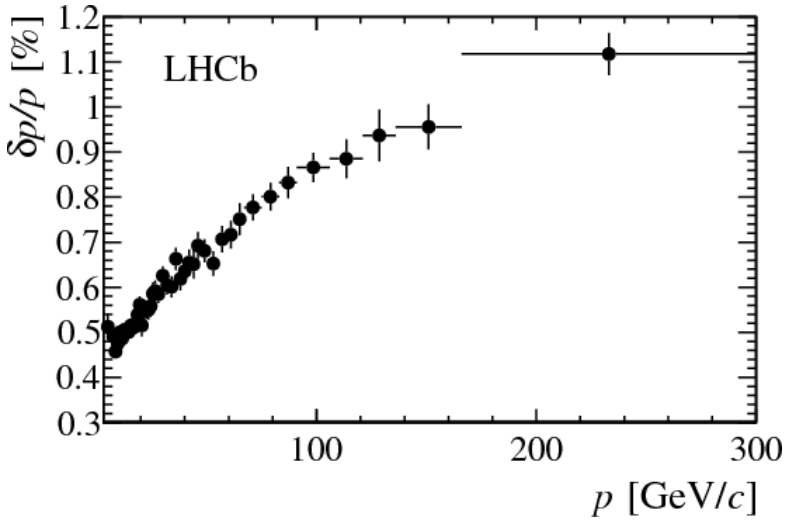

Figure 2.16: Relative momentum resolution as a function of the momentum. The figure is taken from Ref. [57].

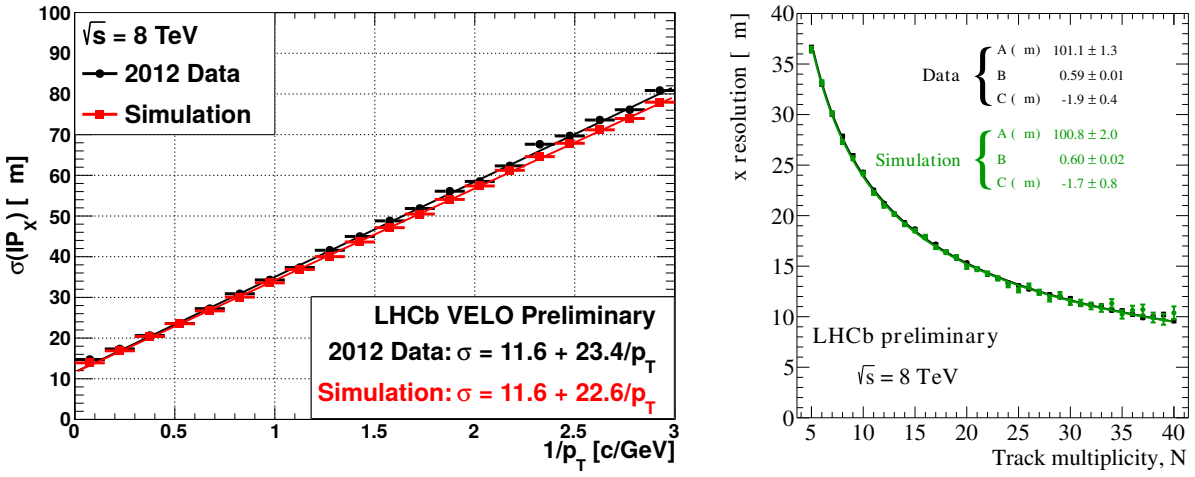

Figure 2.17: IP resolution of long tracks as a function of inverse p_T (left) and the PV resolution of events with one PV in as a function of the track multiplicity, N (right). Both plots are measured using 2012 data and compared to the simulation. The figures are taken from Ref. [57].

with one PV, as a function of the number of tracks in the event; these are measured with 2012 data and compared to the simulated data, which are in good agreement. It took several years of refining the simulation to create this high level of agreement, but there are still small discrepancies remaining between the simulation and data, particularly in IP resolution at high values of $1/p_T$.

2.4 Particle identification

The RICH detectors, the muon system and the calorimetry system all contribute to the identification of particles. The RICH detectors primarily separate charged hadrons, while the muon system identifies muons and the calorimeters separate the remaining particles.

2.4.1 RICH detectors

When a charged particle traverses a material faster than the speed of light in that material, it emits Cherenkov radiation (photons) in a cone around the direction of travel, similar to a sonic boom. The angle at which the cone is emitted is given by

$$\cos(\theta) = \frac{1}{vn}, \tag{2.2}$$

where v is the speed of the particle and n is the refractive index of the material. Figure 2.18 shows the angle of the Cherenkov cone reconstructed in the first RICH detector (RICH1) as a function of the momentum for the various particle types. This figure clearly shows that RICH1 is able to separate the different particles from one another until a certain momentum, at

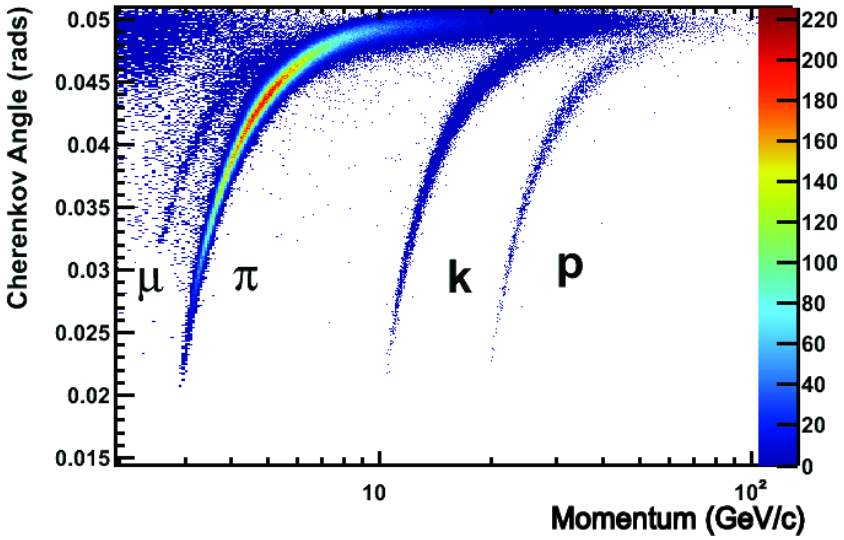

Figure 2.18: Reconstructed Cherenkov angle from the C_4F_{10} radiator as a function of momentum for a selection of particles. The figure is taken from Ref. [61].

Table 2.2: RICH momentum thresholds for the various radiators used in the LHCb RICH detectors. The values are calculated from Eq. (2.3).

	Radiators		
Particle	RICH1 – Aerogel	RICH1 – C_4F_{10}	RICH2 – CF_4
e [MeV c^{-1}]	2.1	9.7	16
μ [GeV c^{-1}]	0.4	2.0	3.3
π [GeV c^{-1}]	0.6	2.6	4.4
K [GeV c^{-1}]	2.0	9.3	15.6
p [GeV c^{-1}]	3.8	17.7	29.7

which point all particles have a similar Cherenkov cone angle; the second RICH detector (RICH2) helps separate these high momentum particles. The threshold above which particles emit Cherenkov radiation is given by

$$p_{\text{thresh}} = \frac{m}{\sqrt{2(1-n)}} \,. \tag{2.3}$$

The momentum thresholds of particles are tabulated in Table 2.2 for the various radiators that are used in the RICH detectors.

The LHCb experiment uses two RICH detectors to identify charged particles [58, 67]. In these detectors the particles pass through one or more radiators in which the Cherenkov radiation is generated; the photons are then reflected, by one or two mirrors, on to hybrid photon detectors (HPDs) which are placed outside of the angular acceptance of LHCb. To detect the photons the HPDs use a photocathode to absorb the photon and release an electron. This electron is then accelerated through a voltage in the range of 10 kV to 20 kV and detected on a silicon-pixel detector. The layouts of the two RICH detectors are shown in Fig. 2.19, where the mirrors, radiators and HPDs are labelled.

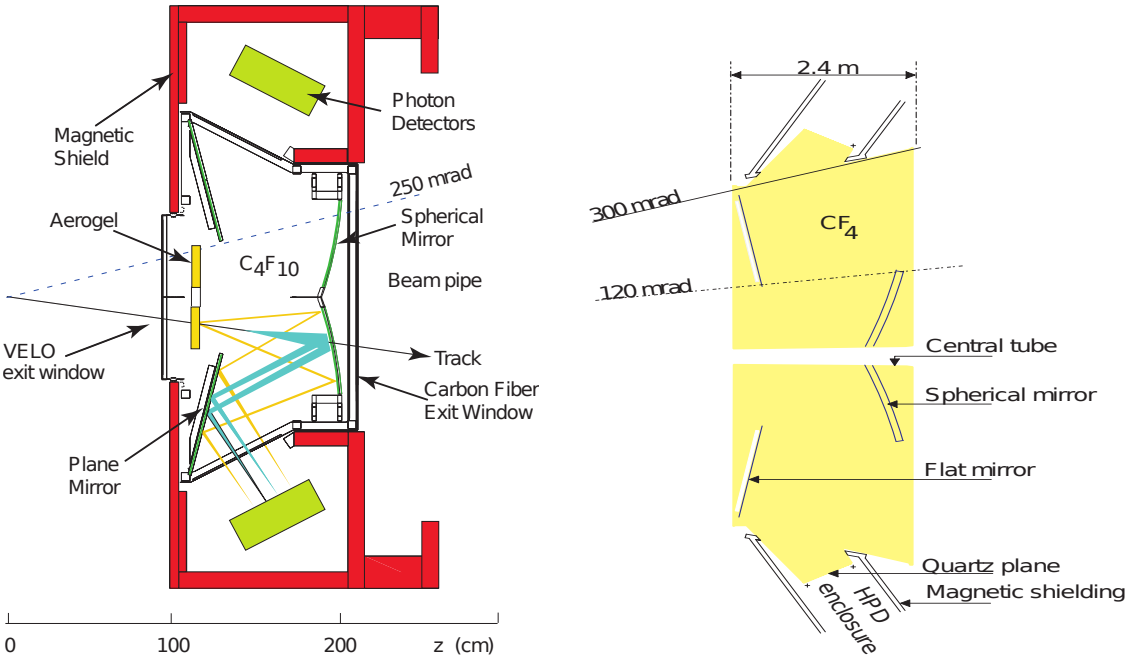

Figure 2.19: Schematics of RICH1 (left) and RICH2 (right). The diagrams are taken from [61].

The first of the two RICH detectors is designed to identify charged hadrons in the momentum range 2 GeV c^{-1} to 40 GeV c^{-1} and is placed between the VELO and the TT. This RICH detector contains two radiators, C_4F_{10} and Aerogel. The Aerogel has a higher refractive index to help identify particles below the momentum threshold of the C_4F_{10}, but it was removed during LS1 to allow particles to pass through more C_4F_{10}. Table 2.2 compares the various momentum thresholds for the radiators used in LHCb.

The second RICH detector is designed to identify charged hadrons in the momentum range 15 GeV c^{-1} to 100 GeV c^{-1} and is placed after the tracking stations. This RICH detector uses tetrafluoromethane (CF_4) as a radiator, with about 5 % CO_2 to prevent scintillation in the CF_4. RICH1 covers the full LHCb angular acceptance, but RICH2 only covers the region around the beampipe from 15 mrad to 120 mrad in the horizontal plane

and from 15 mrad to 100 mrad in the vertical plane. RICH2 has a limited angular acceptance as this is the region where simulations predicted that the majority of high momentum particles would pass through.

2.4.2 Calorimetry

The calorimeter system identifies hadrons, electrons and photons; further it measures their energies, positions and provides a trigger decision to the level-0 (L0) trigger within 4 µs of the *pp*-interaction [68].

Calorimeters work by converting the incident particle into showers of lower energy particles that can be absorbed by the calorimeter and then the energy deposited in the calorimeter is measured. The characteristic interaction length for an electromagnetic interaction is the radiation length (X_0), while the equivalent distance for strong interactions is the nuclear interaction length (λ_I); these distances characterise the length of electromagnetic and hadronic showers respectively. Both X_0 and λ_I are very important for a calorimeter, as a calorimeter which has insufficient material to absorb the energy of a particle will not be able to correctly measure the particle's energy. An ECAL should have a short X_0 and be many X_0 thick, while an HCAL should have a short λ_I and be many interaction lengths thick. Since electrons and photons are less penetrating than hadrons, the ECAL is placed in front of the HCAL and so the ECAL should have a long λ_I in order to minimise hadronic interactions in the ECAL.

The calorimeter system consists of the SPD, the PS, the ECAL and the HCAL. Each of the detectors contributes to the identification of the particles in different ways, as is illustrated in Fig. 2.20. Only charged particles will

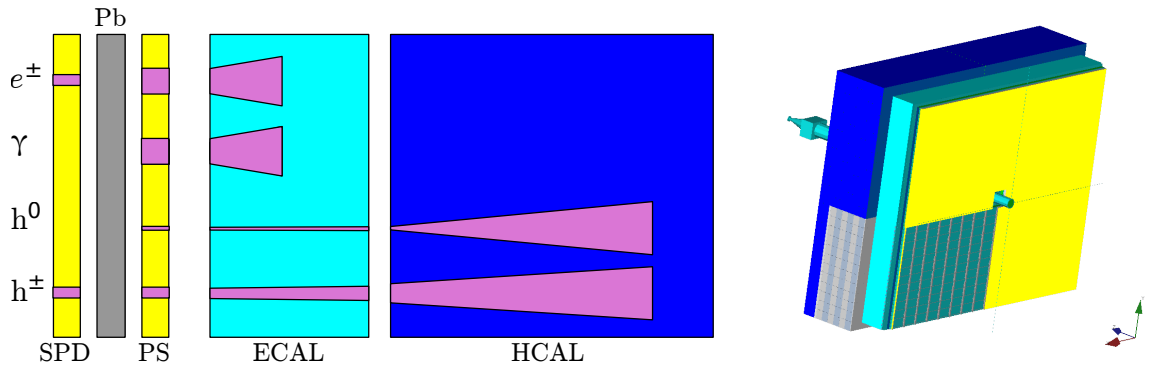

Figure 2.20: Illustration of how various particles interact with each of detectors in the calorimeter system [69] (left) and an image of the calorimeter system as it is described in the LHCb simulation [70] (right). The purple shapes in the left hand image indicate the amount of energy deposited in each of the detectors. The detectors are shown in various colours: the SPD and PS are shown in yellow, the ECAL in cyan and the HCAL in blue. In the right image the beampipe is also shown in the middle of the detectors.

produce hits in the SPD. After the SPD there is a layer of $2.5\,X_0$ lead (Pb), which produces electromagnetic showers; the showers then produce hits in the PS. The ECAL comes next and comprises $25\,X_0$ in order to absorb electrons and photons. Finally the $5.6\,\lambda_I$ thick HCAL is designed to absorb hadrons.

The SPD and PS are almost identical detectors, which help to improve the spatial resolution of energy deposits made in the calorimeters. They use blocks of scintillator to detect charged particles. Inside the blocks, loops of wavelength-shifting (WLS) fibres trap photons generated by the scintillator and direct them into multianode photomultiplier tubes (MAPMTs), which detect the photons. The SPD, PS and ECAL share a very similar geometry as shown in Fig. 2.21, but it is projective; this means transverse dimensions

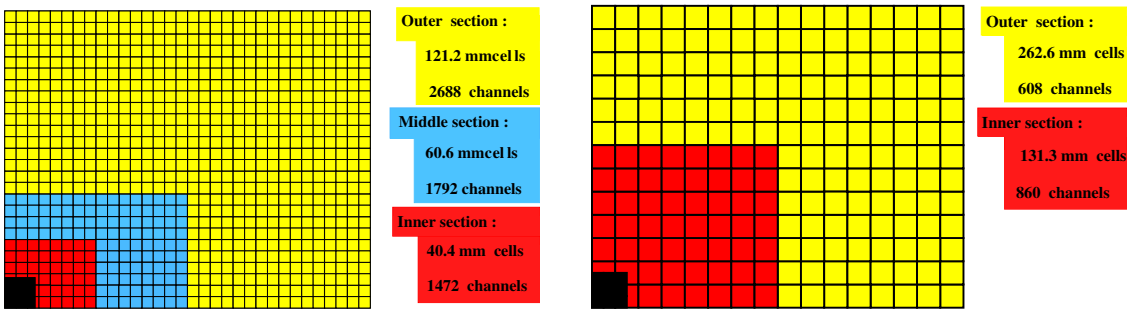

Figure 2.21: Granularity of one quarter the ECAL (left) and the HCAL (right). The SPD and PS have the same segmentation as the ECAL. The diagrams are taken from Ref. [68].

scale with distance from the collision point and so the size of the SPD and PS is scaled down by ∼1.5 % when compared to the ECAL.

The ECAL is a shashlik-type detector; this means it consists of alternating layers which detect and absorb the particles. In total 66 layers of 2 mm thick lead, reflecting TYVEK paper and 4 mm thick scintillator are stacked in the z-direction to form a 42 cm long module. The photons created in the scintillator are read out by WLS fibres which pass through all layers in the z-direction; these fibres are fed out of the back of the module into photomultiplier tubes (PMTs), which detect the photons. The modules are placed to cover an area 7.7 m by 6.3 m.

The HCAL follows a similar design to the ECAL, with alternating layers of 1 cm thick iron plates and 3 mm thick scintillating tiles; however, the difference is that the tiles are stacked laterally instead of in the z-direction. Scintillation photons are collected by WLS, which are placed along the edges of the scintillating tiles and deposit the photons in PMTs at the back of the HCAL cell. The HCAL is 8.4 m high, 6.8 m wide and 1.65 m in the z-direction. The weight of the HCAL is about 500 t.

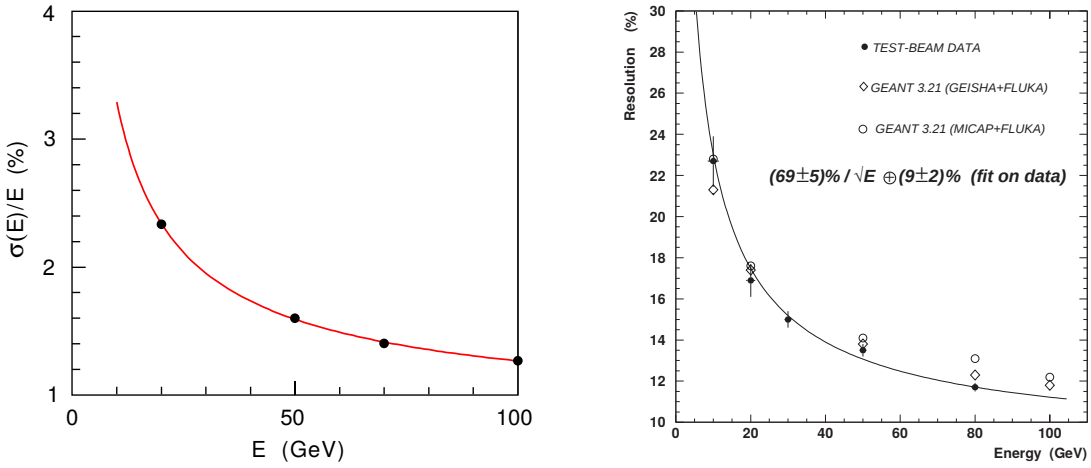

Figure 2.22: Relative energy resolutions as a function of energy determined for the ECAL (left) and HCAL (right). The hollow points show simulated events, while the solid points show the test-beam data which are fitted with the function $\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b$. The ECAL resolution was determined using a test-beam of electrons and found $a = 9.0 \pm 0.5\,\%$ and $b = \sim 0.8\,\%$. The HCAL resolution was determined using a test-beam of pions and found $a = 69 \pm 5\,\%$ and $b = 9 \pm 2\,\%$. The figures are taken from Ref. [61].

The relative energy resolution ($\frac{\sigma_E}{E}$) of calorimeters are often parametrised as using the following:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \,,$$

where E is the energy in units of GeV, a is the stochastic term, b systematic term and $\oplus$ represents addition in quadrature. The relative energy resolutions of both the ECAL and HCAL have been measured and are shown in Fig. 2.22. The a and b parameters of the relative energy resolution were found to be $a = 9.0 \pm 0.5\,\%$ and $b = \sim 0.8\,\%$ for the ECAL and $a = 69 \pm 5\,\%$ and $b = 9 \pm 2\,\%$ for the HCAL; these were measured using test-beams of electrons and pions for the ECAL and HCAL respectively. In the case of the

HCAL simulated datasets are also used to compare with the results from the test-beam and these show good agreement with one another.

2.4.3 Muon identification

The muon system is designed to identify muons, measure their p_T and provide high p_T muon candidates to the L0 trigger [71]. The muon system contains 5 stations, one before the calorimeter system (M1) and four after it (M2-M5). The detectors use a projective geometry, which is shown in Fig. 2.23. Lead of thickness 80 cm is placed between each of the last four muon stations; this means that a muon must have a momentum of at least $6\,\mathrm{GeV\,c^{-1}}$ to pass all five stations.

Each muon station has four regions (R1-R4), with R1 closest to the beampipe and R4 at the outer edge of the detector. The innermost region (R1) has the highest resolution and each subsequent region (as you move further away from the beampipe) has half the granularity of the previous region. The detector is also more finely grained in the x-direction than the y-direction; this improves the momentum resolution. The granularity of the various detector regions are shown in the top right of Fig. 2.23. The muon stations closer to the interaction point (M1-M3) have a higher spacial resolution than those at higher z (M4-M5). The muon system consists primarily of MWPC detectors, except for the innermost region of M1 where GEM detectors are used. The last four muon stations each have four layers of detector, organised into a logical OR, such that a muon will be detected if it leaves a hit in any of the four layers of a station. The first muon station only has two layers in order to reduce the material before the calorimeter.

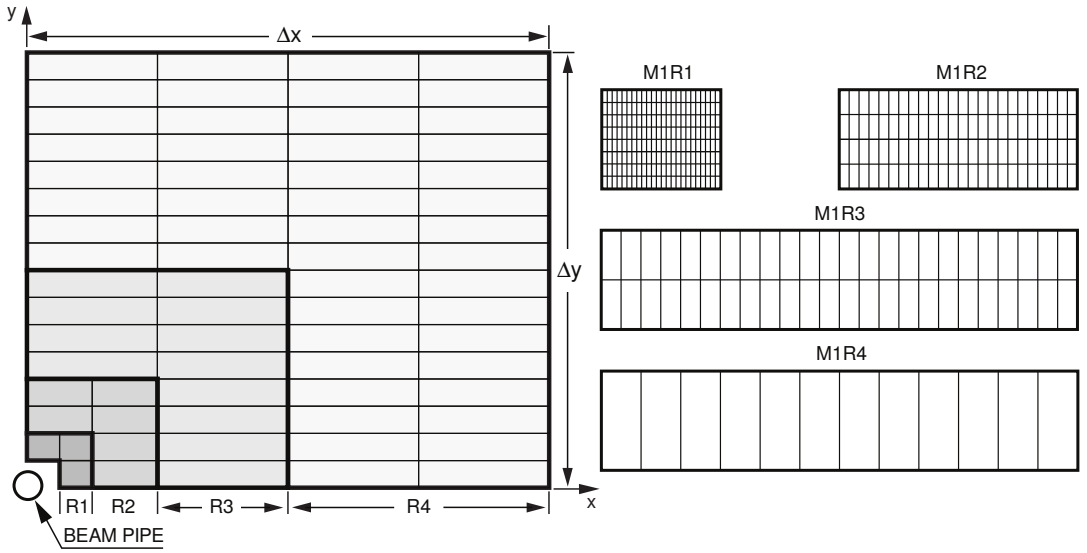

Figure 2.23: Top left image shows a quarter of a muon station, with each rectangle representing a muon chamber. The chambers are sub-divided into sensitive pads; this is shown on the top right. The muon system consists of five muon stations, one before the calorimeter system and four after the calorimeter system; this is shown on the left. A lead filter is placed between each of the muon stations M2-M5. The diagrams are taken from Ref. [61].

2.4.4 Particle identification performance

The information from the PID detectors is combined to determine the particle most likely to have caused each of the tracks reconstructed by the tracking

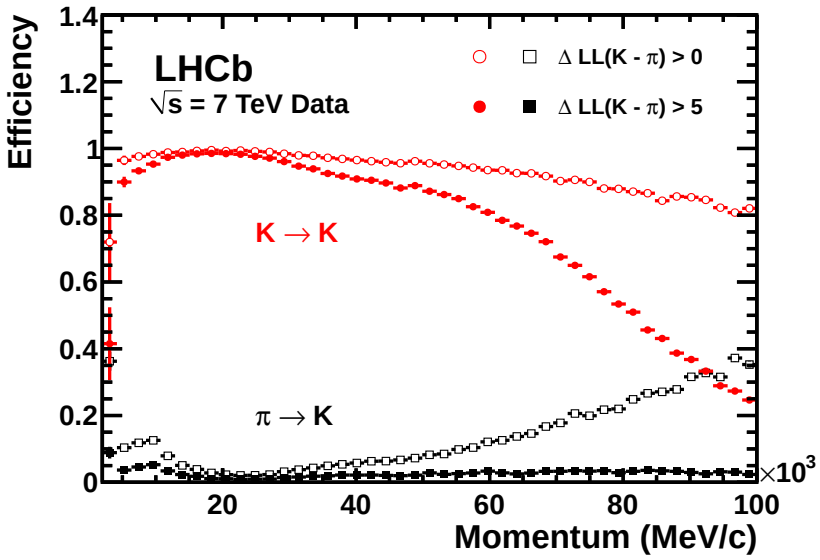

Figure 2.24: Fraction of correctly identified kaons and misidentified pions as kaons, as a function of momentum for two different requirements on the ΔLL variable calculated using 2011 data. The figure is taken from Ref. [61].

detectors. The difference in logarithms of likelihoods (ΔLL) of two PID hypotheses is used to separate one particle from another. The selection efficiencies of two different requirements on the difference between the kaon and pion log likelihoods, $\Delta\mathrm{LL}(K-\pi)$, are shown in Fig. 2.24. Using this technique charged hadrons can be identified in the momentum range from $2\,\mathrm{GeV\,c^{-1}}$ to $150\,\mathrm{GeV\,c^{-1}}$ [72].

In addition to the ΔLL method, LHCb also uses neural networks (NNs) for particle identification. The neural networks are given more variables than are input to the ΔLL calculation and so have greater PID separation power. The neural networks produce a probability that the track was caused by a given particle type. This ProbNN method generally has better performance than the standard ΔLL method; this is demonstrated for muons in Fig. 2.25, where it can be seen ProbNN rejects more background for a given signal

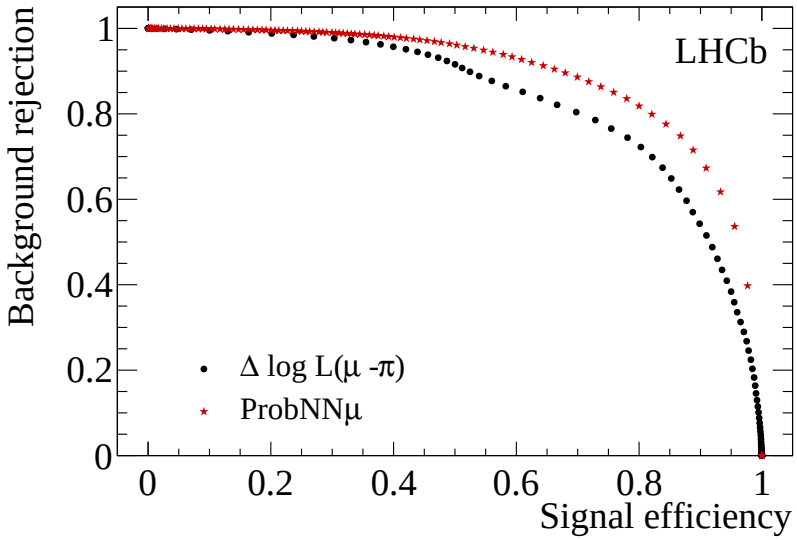

Figure 2.25: Muon background rejection as a function muon identification efficiency for ProbNNμ (red) and $\Delta\text{LL}(\mu - \pi)$ (black) in decays of the type $\Sigma^+ \to p^+\mu^+\mu^-$. Simulated events are used as the signal sample and data sidebands are used for the background sample. The figure is taken from [57].

efficiency. However, this new PID method was developed towards the end of the research presented in this thesis and so it was not available for all analyses presented here.

2.5 Trigger

The nominal bunch crossing frequency of the LHC is 40 MHz; however, LHCb can only save events at a rate of 2 kHz for analysis [74, 75]. This is due to various bandwidth limitations, storage limitations and also limitations in the readout speed of many components of the detectors. Thus most events must be discarded quickly, keeping only the most interesting collisions. This is done by the trigger in two stages (as shown in Fig. 2.26): first the

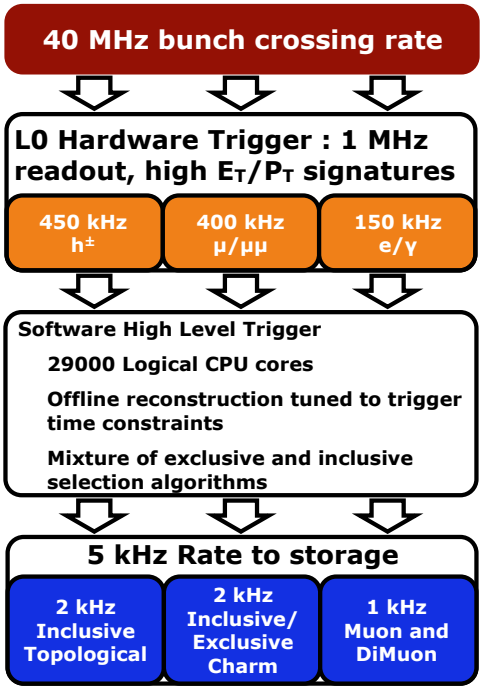

Figure 2.26: The diagram illustrates the data reduction performed by each level of the trigger. The figure is taken from Ref. [73].

hardware L0 trigger reduces the rate to 1 MHz and then the software high-level trigger (HLT) reduces the rate of events to 2 kHz.

The L0 is split into three separate units: the calorimeter trigger, the muon trigger and the pile-up trigger. The pile-up trigger is used for luminosity measurements and is not discussed here. The L0 calorimeter trigger uses information from the SPD, PS, ECAL and HCAL detectors to create candidates and then the computed transverse energy (E_T) must be larger than a certain value (listed in Table 2.3). An L0 hadron candidate contains the highest E_T HCAL cluster, plus anything in the matching ECAL cell. An L0 photon candidate is formed from the highest E_T ECAL cluster and requires 1 or 2 hits in the PS. The L0 electron candidates are formed from the L0 photon candidates, with the additional requirement that there is at least one hit in the SPD in front of the PS cells. The L0 muon trigger requires 5 aligned

Table 2.3: Typical L0 thresholds used in Run 1 [76]. Muonic triggers are activated when particles are reconstructed with p_T greater than the values shown below, whereas calorimeter triggers are activated when particles are reconstructed with E_T greater than the values shown below.

	minimum p_T or E_T		maximum N_{SPD}
	2011	2012	2011 and 2012
single muon	$1.48\,\text{GeV}\,\text{c}^{-1}$	$1.76\,\text{GeV}\,\text{c}^{-1}$	600
dimuon $p_{T1} \times p_{T2}$	$(1.30\,\text{GeV}\,\text{c}^{-1})^2$	$(1.60\,\text{GeV}\,\text{c}^{-1})^2$	900
hadron	3.50 GeV	3.70 GeV	600
electron	2.50 GeV	3.00 GeV	600
photon	2.50 GeV	3.00 GeV	600

hits in the muon stations (M1-M5) and the p_T to be greater than the amount listed in Table 2.3. For the L0 dimuon trigger the product of the highest muon p_T and second highest muon p_T in the event is used. The increase of the thresholds used between 2011 and 2012 was due to an increase in the luminosity between these years; the higher thresholds reduce the trigger rates to acceptable levels. To reduce processing time, events which have a particularly high multiplicity are vetoed by limiting the maximum number of hits in the SPD detector (N_{SPD}) to the values listed in Table 2.3.

After the L0 trigger selects events, the HLT further refines the event selection. When events pass the L0 trigger, they are read out of the detector into a computer cluster situated underground near LHCb, where approximately 29 000 instances of the HLT software (Moore) are run. The HLT is divided onto two parts: HLT1 and HLT2. HLT1 contains various trigger lines which aim to confirm the L0 decision. It only reconstructs the most interesting tracks in the event and looks for tracks with a large IP or high p_T; it also

ensures that the track is well fitted. HLT1 reduces the rate to 80 kHz, which is low-enough for the full event to be reconstructed with some simplifications in the algorithms used; this is performed in HLT2. Since all particles can be reconstructed, composite particles can be made by combining multiple tracks from a common vertex. This allows both inclusive and exclusive trigger lines to be produced. Events which pass HLT2 are saved to disk for full analysis later.

2.6 Simulation and reconstruction software

Simulations of LHCb are often required and the first step in creating these is to produce a pp-collision with Pythia [59, 77] using a specific LHCb configuration [78]. Hadronic decays are then modelled using EvtGen [79], and the final-state radiation is created using Photos [80]. The Geant4 toolkit [81] is used to describe the interaction of the generated particles with the detector, and its response [82]. The best values of branching fractions and form factors at the time of the simulation are input to EvtGen and are taken from Ref. [1]. This is all run within the Gauss software package which handles the generation of simulated events for LHCb. To simulate the response of the detector electronics, the Boole software package is used; this produces a very similar output to that from the detector so that the same reconstruction and analysis applications can be used for simulated and real events. The next steps for the reconstruction and analysis consist of: Moore, which implements the HLT; Brunel, which performs the offline reconstruction; and DaVinci where user analysis can be performed on the data [83]. This flow of data is visualised in Fig. 2.27. All selection of data before and including Moore is referred to as online, since it is performed in real time; all data

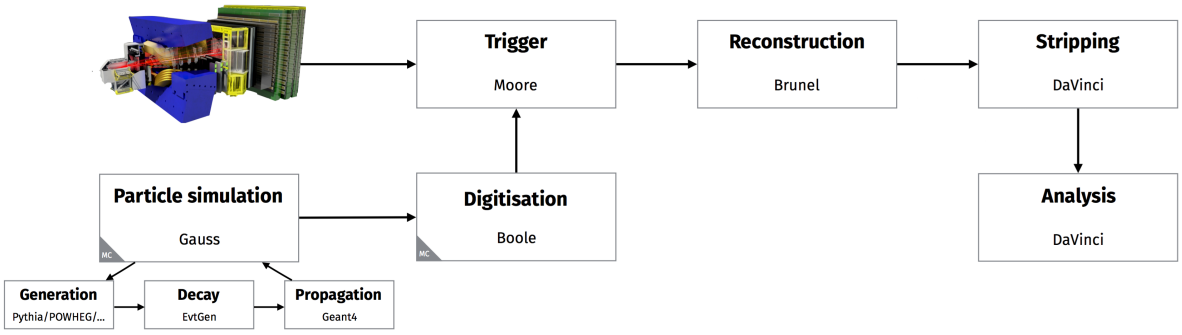

Figure 2.27: Illustration of the data flow through the LHCb software stack, for both simulated and real data. The figure was adapted from Ref. [84].

selection after Moore is referred to as offline. Due to the large amount of data produced by LHCb, interesting events are selected centrally for the collaboration in a process known as stripping. Each event selection algorithm is known as a stripping line and user analyses are run on the output from these stripping lines using DaVinci.

CHAPTER 3

The vertex locator upgrade

The LHC is capable of delivering a higher instantaneous luminosity to the LHCb experiment than it currently does, but instead the beams are separated to reduce the number of collisions (see Fig. 2.4). The GPDs are delivered a maximum instantaneous luminosity of $7 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, whereas LHCb only receives a maximum instantaneous luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$; this is a factor of almost 20 times lower. The main limitation that necessitates this low instantaneous luminosity is the triggering strategy of LHCb. The two main detectors responsible for triggering interesting events are the muon stations and the calorimeters and these require a particle to be reconstructed with high p_{T} or E_T. The various thresholds for triggering an event with large p_{T} or E_T in 2011 and 2012 are given in Table 2.3. In order to keep the data rate manageable, as the luminosity is increased these trigger thresholds must also be increased. In 2012 the hadronic E_T threshold was set to 3.70 GeV (about 70 % of the B_s^0 mass, $5.37\,\mathrm{GeV}\,\mathrm{c}^{-2}$), and so increasing this already very high threshold will rapidly

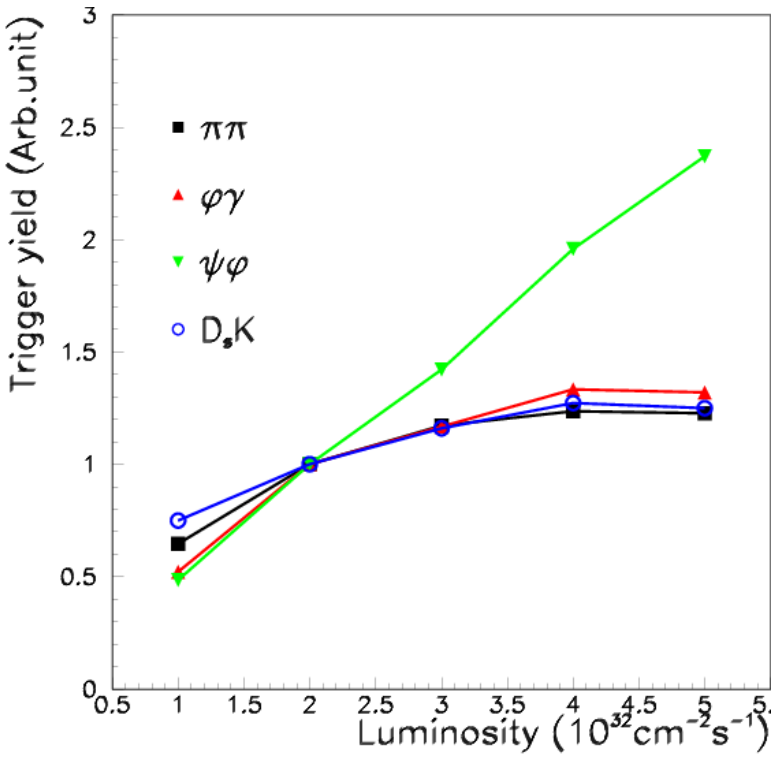

Figure 3.1: Trigger yield of various simulated B_s^0 decay modes shown as a function of the luminosity. Muonic triggers scale with luminosity, while calorimeter triggers saturate. The figure is taken from Ref. [85].

decrease the fraction of signal events which are triggered. Figure 3.1 shows the trigger yield as a function of the instantaneous luminosity for several different B_s^0 decay modes. The $B_s^0 \to J/\psi\,\phi$ decay is triggered using muons from the J/ψ decay and this scales linearly with the luminosity. For the other modes (which are triggered using the calorimeter), the increase in luminosity does not counteract the decrease in the fraction of signal which is triggered and the yield of triggered events saturates.

To overcome this limitation LHCb is planning a major upgrade during Long Shutdown 2 (LS2). This upgrade will allow the full detector to be read out into a software trigger, avoiding the need for a hardware trigger; this will

enable LHCb to efficiently run at a luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which is a factor of 5 increase in luminosity. With this higher instantaneous luminosity LHCb aims to collect an integrated luminosity of $50\,\mathrm{fb}^{-1}$ by the end of Run 4 in 2028. This rise in instantaneous luminosity increases the mean number of pp-interactions per bunch crossing (ν) from 2.5 to 7.6, which in turn increases the average number of particles created in each collision; these particles will cause additional radiation damage to the detectors, which will require some to be replaced in order to ensure that the detectors survive Run 3 and beyond. However, the requirement to read out the full detector at 40 MHz into the software trigger drives most changes in the upgrade, because it requires an upgrade of the front-end electronics of all detectors.

Full readout into a software trigger is foreseen to require a factor of 40 increase in computing resources, and so these resources will be increased in a staged manner. At the start of Run 3 a hardware low-level trigger (LLT) will reduce the rate of events to the software trigger. The LLT will be very similar to the current L0 trigger, but have more flexible thresholds so that as the computing resources are increased the thresholds are decreased.

The current LHCb detectors are very similar to the upgrade detectors, which are shown in Fig. 3.2. The muon stations (M2-M5) will remain mostly the same in the upgrade, with the exception of new front-end electronics for the stations and additional shielding in front of M2. However, the first muon station will be removed. It will no longer be required to precisely calculate the p_T in the trigger; instead information from the tracking detectors will be used. The ECAL and HCAL will also remain much the same, but with new front-end electronics. It is thought that the radiation damage to the calorimeters will start to become an issue at the end of Run 3, and so

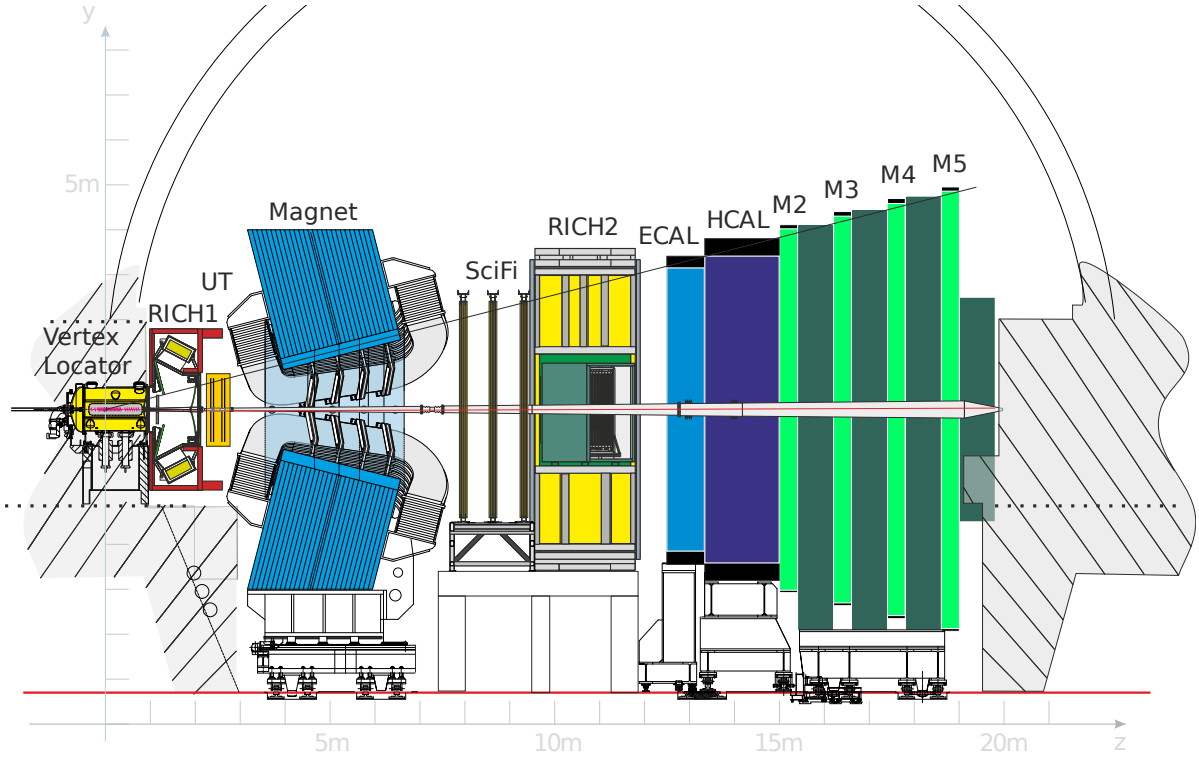

Figure 3.2: Layout of the planned upgrade to the LHCb experiment inside the underground experimental area, with the various detectors labelled. The diagram is taken from Ref. [86].

during Long Shutdown 3 (LS3) some calorimeter modules could be replaced. The SPD and PS will be removed as these were mainly used to suppress backgrounds when triggering on electron and photon candidates; the LLT has less stringent requirements on calorimeter background suppression, and the software trigger will be able to use information from the tracking detectors instead. The RICH detectors will have their PMTs replaced with MAPMTs in order to allow faster readout of the detectors. Further, the MAPMTs are smaller and so this allows the optics of the RICH detectors to be re-optimised. The IT detectors have their front-end electronics integrated into their detector modules and so the IT and the OT will both be removed and replaced with the scintillating fibre tracker (SciFi); this detector will use 5

to 6 layers of 250 µm diameter scintillating fibres to form modules. The SciFi modules will be placed in a $(x-u-v-x)$ layout and will be read out by silicon photomultipliers (SiPMs) placed above and below the modules, but outside of the angular acceptance of LHCb. The TT will also be replaced by a new high granularity silicon micro-strip detector called the upgrade tracker (UT). The TT is to be replaced for three reasons: the occupancy would be too high in the upgrade conditions, it would not survive the radiation damage expected in the upgrade, and the current front-end electronics read out the detector too slowly. The UT will have shorter strip lengths and thinner strips near the beampipe to cope with higher track multiplicities in this region. It will also move closer to the beampipe to improve its angular acceptance. The VELO will be replaced for the same reasons as the TT (high occupancies, too much radiation damage and slow readout); the following sections focus on the design of the upgrade VELO.

3.1 VELO design considerations

The most basic requirement of the VELO is that >99 % particles must leave four hits in the detector if they are within the geometric acceptance of LHCb; for the upgrade the geometric acceptance is defined to be in the pseudorapidity range $2 < \eta < 5$ and the particle must have originated from within the region $|z| < 2\sigma_{\text{lumi}}$, where the root mean squared deviation (RMS) of the luminous region ($\sigma_{\text{lumi}} = 63\,\text{mm}$) is derived from various LHC bunch parameters as expected in Run 3. For comparison, in Run 1 the RMS of the luminous region was $\sigma_{\text{lumi}} = 53\,\text{mm}$ and this is compared to the current VELO geometry in Fig. 2.7. Of the particles that are in the geometric acceptance of LHCb, the fraction of these which can leave four hits in the detector is known

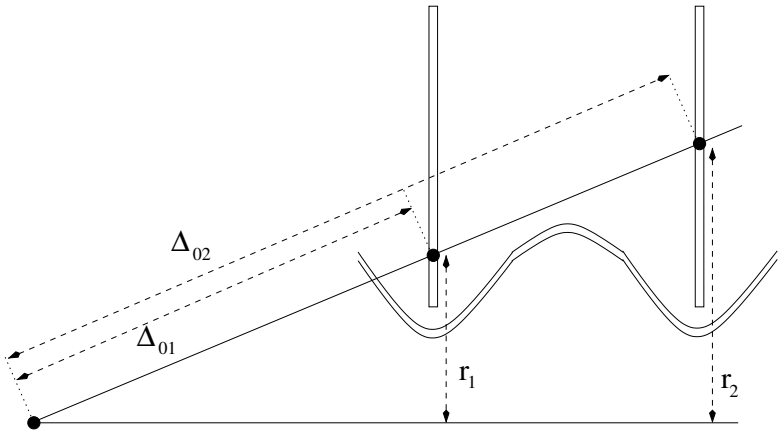

Figure 3.3: An explanation of the parameters used in Eq. (3.1). Two VELO modules and a section of the RF-foil are shown along with the beam-line (the bottom horizontal line) and a track (the line at an angle). The point at which the two solid lines meet is the origin of the track. The diagram is taken from Ref. [87].

as the geometric efficiency. We require four hits even though only three are necessary in order to reconstruct a track as this allows some redundancy. Once the geometric efficiency of the VELO is suitably high, the next most important consideration is the performance of the detector.

The primary measure of the performance of the VELO is the IP resolution (see Fig. 2.17 for the current IP resolution), because the IP is essential for quickly triggering events containing charm or beauty decays and the IP resolution also affects measurements of physics quantities (such as the decay-time resolution). Figure 3.3 illustrates the key parameters in the expression which approximately describes the IP resolution and this expression is given by

$$\begin{aligned} \sigma_{IP}^2 &= \frac{r_1^2}{{p_T}^2\sqrt{2}}\left[0.0136\sqrt{\frac{x}{X_0}}\left(1+0.038\,\ln\left(\frac{x}{X_0}\right)\right)\right]^2 + \frac{\Delta_{02}^2\sigma_1^2+\Delta_{01}^2\sigma_2^2}{\Delta_{12}^2} \\ &= \frac{1}{{p_T}^2}\,\sigma_{\text{MS}}^2 + \sigma_{\text{extrap}}^2\,, \end{aligned} \tag{3.1}$$

where σ_{MS} is the multiple scattering term, σ_{extrap} is the extrapolation term, r_1 is the radius of the first hit, p_T is the transverse momentum of the particle, $\frac{x}{X_0}$ is the fraction of the radiation length travelled before the second hit, σ_1 and σ_2 are the measurement errors on the first and second hits respectively and Δ_{ij} represents the distance between i and j, which can take the values: 0 (the track origin), 1 (the first hit) or 2 (the second hit). From Eq. (3.1), there are four ways that the VELO design can improve the IP resolution: reduce the material before the second measured point, increase the spacing between modules, reduce the measurement errors on the first or second points and finally reduce the distance from the first and second hits to the track origin.

3.2 The technology choice – strips or pixels

Initially two designs were proposed for the upgraded VELO detector: a silicon strip detector known as VL, and a silicon pixel detector known as the VP. The new VELO detector will be closer to the proton-beam than the current VELO; previously the inner edge of the active sensor was placed at a radius of 8.2 mm (after closing the VELO), while in the upgrade this will be reduced to 5.1 mm [88, 89]. The design of the VL is conceptually similar to the current detector, but with higher granularity, and it also has thinner sensors in order to reduce the amount of material traversed by particles. The widths of the strips range from 25 µm (at the edge closest to the beampipe) to about 110 µm (at the outermost edge of the sensor) [90]. The VP builds upon an ASIC called Timepix3 [91], which is widely used in medical detectors. It has square pixels, which are 55 µm × 55 µm in size. The active edges of the sensors of both the VP and VL are placed at 5.1 mm from the beam;

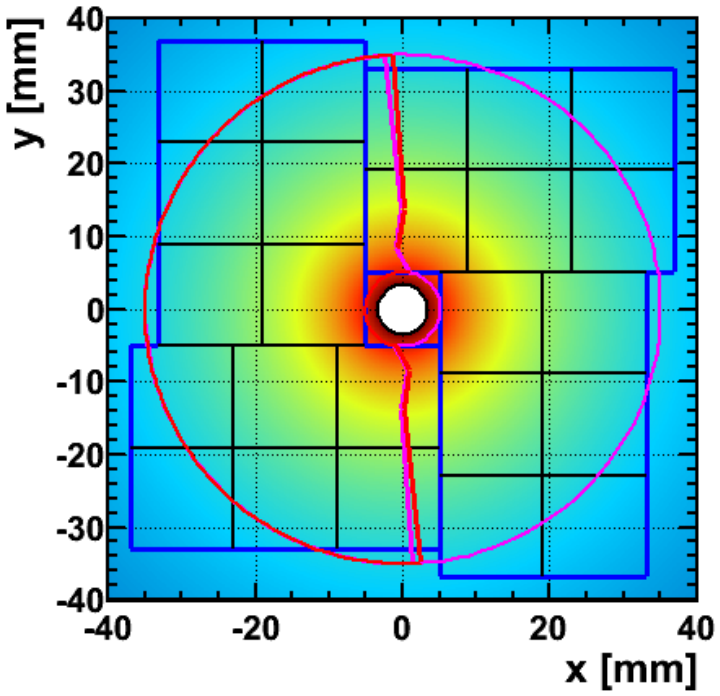

Figure 3.4: Approximate geometry of the sensitive areas of the detectors that could be used in the upgrade. The magenta lines show the VL sensors, and blue lines shown the VP ASICs. The coloured background indicates schematically the expected track multiplicity.

however, the VP modules are based on 12 square ASICs, which results in a square region for the beam to pass through (as shown in Fig. 3.4). This means the edge of the active sensor has a varying radius from the beam of 5.1 mm to 7.2 mm. Looking at these basic design properties of the two detectors and Eq. (3.1), the VL should outperform the VP. The resolution of the first and second hits will usually be better in the VL, and the VL is closer to the proton beam than the VP; both of these factors will contribute to improving the IP resolution of the VL when compared to the VP. Overall the hit resolutions of the two detectors are quite similar, as shown Fig. 3.5.

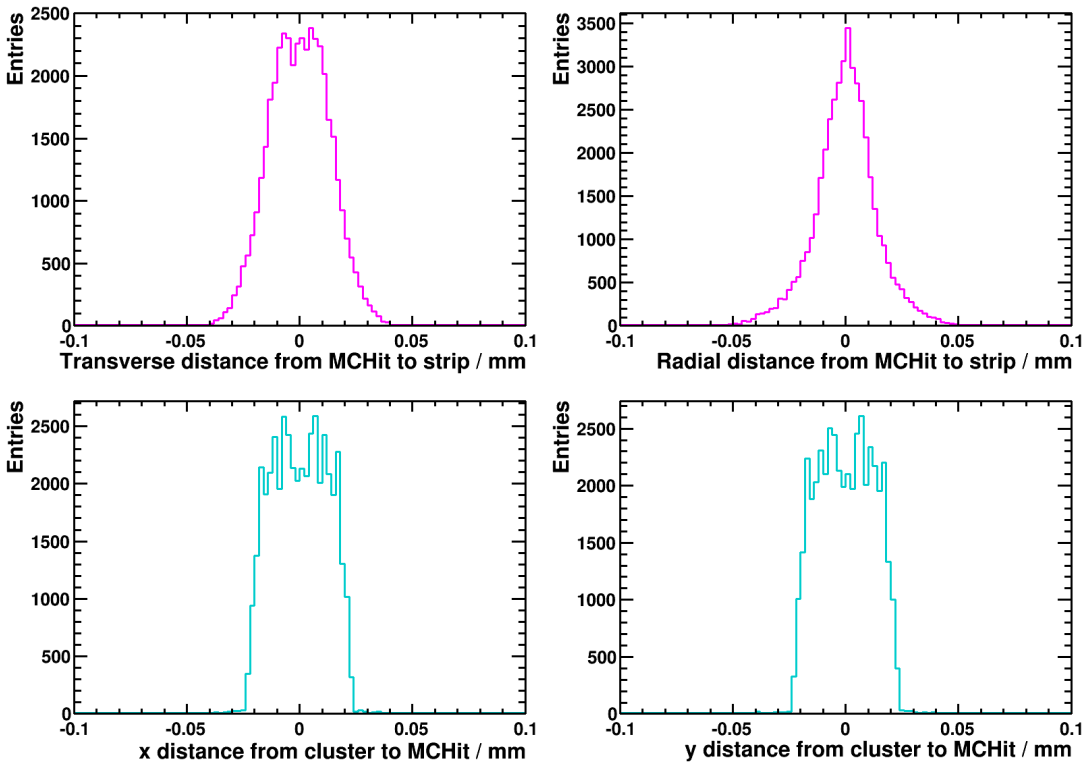

Figure 3.5: Hit residuals from the microchannel VL (top row, magenta) and VP (bottom row, turquoise) detectors. The top left, top right, bottom left and bottom right plots show residuals in the ϕ, r, x and y directions and the RMS of these distributions are 0.0134 mm, 0.0140 mm, 0.0123 mm and 0.0124 mm.

Both upgrade options would improve upon the IP resolution of the current VELO as both upgrade detectors will be closer to the proton-beam.

In addition to the choice of detector technology, the cooling method must also be chosen. The first choice was microchannel evaporative CO_2 cooling, where CO_2 evaporates in thin channels within the silicon substrate of the module. The second cooling method was a more conservative and conventional design, which was different for the VL and VP, but both add extra material into the design. For the VL a layer of TPG is added between

the R and ϕ sensors with cooling blocks placed around the sensors in which CO_2 would evaporate. The conservative cooling method for the VP involved adding a layer of graphite foam (pocofoam [92]) between the sensors with embedded pipes in which CO_2 would evaporate [93]. The microchannel cooling benefits from having significantly less material for particles to pass through; however, the microchannels only have 140 µm of silicon separating the VELO vacuum and the high pressure CO_2. If the microchannels were to burst, it would contaminate the LHC vacuum and stop the LHC from running.

Detailed simulations of the four detector options (the VL with two cooling options and the VP with two cooling options) were developed. Figure 3.6 shows the IP resolutions from the simulations of the four upgrade options and compares it to the current VELO. The two VL detectors and the VP with

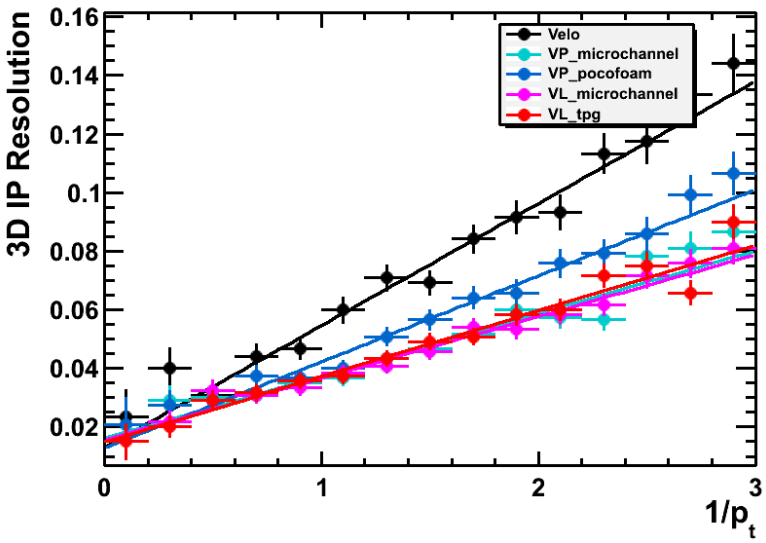

Figure 3.6: The average 3D IP resolution as a function of inverse p_T for tracks which traverse the whole of LHCb from simulated events. The lines are linear fits to the points. The five colours show: the current detector (black), the VP with microchannels (turquoise), the VP with pocofoam (blue), the VL with microchannels (magenta) and the VL with TPG (red).

microchannel cooling perform similarly. The more conservatively cooled VP detector performs slightly worse than the other three upgrade options as this design adds a large amount of extra material into the design of the modules; the amount of material traversed by particles is shown in Fig. 3.7 for the four designs.

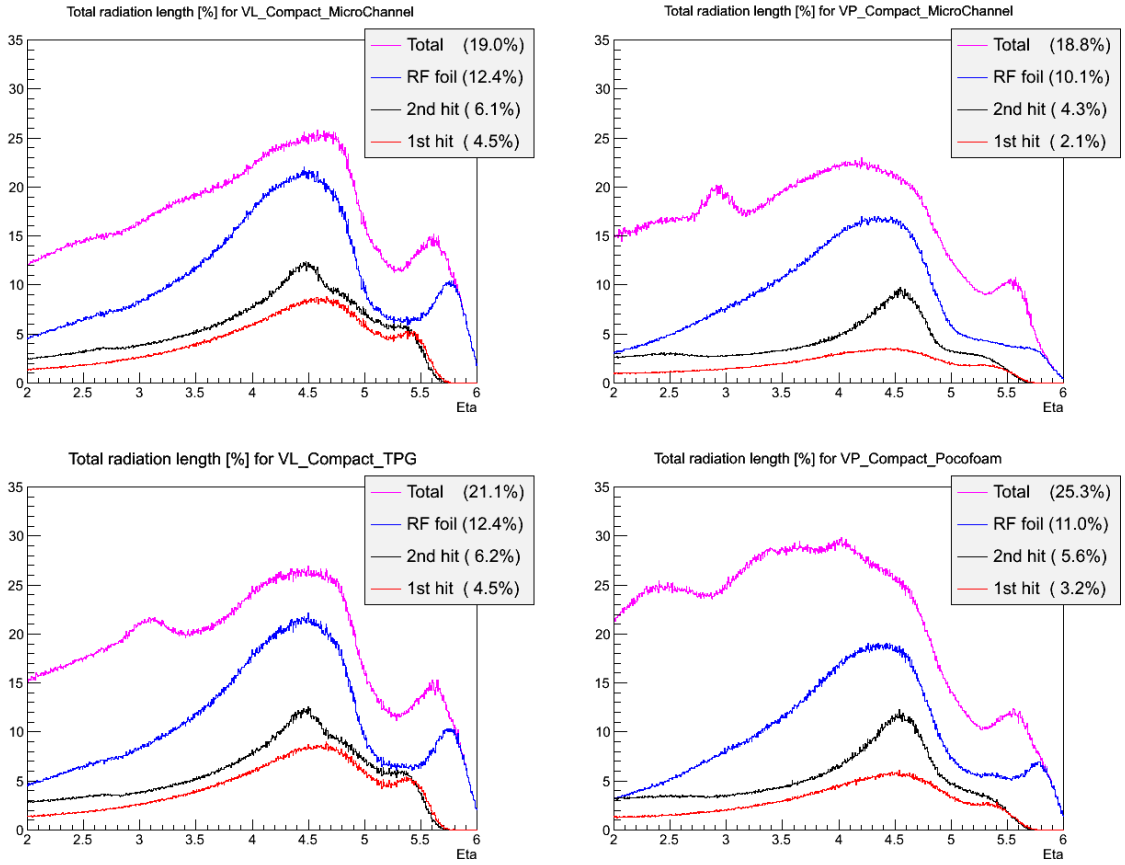

Figure 3.7: Material traversed by tracks as a function of η in units of percent of a radiation length. The four colours show the material: before the first hit (red), before the second hit (black), in the RF-foil (blue) and overall (magenta). The four plots show the four upgrade detector options: the VL with microchannels (top left), the VL with TPG (bottom left), the VP with microchannels (top right) and the VP with pocofoam (bottom right). Figure taken from Ref. [93].

Given the similar performances of the VL and VP, the VELO group decided to choose the pixel detector over the strip detector. The VP is more radiation hard than the VL; simulations show that after the VL has received the radiation dose equivalent to the planned integrated luminosity of $50\,\text{fb}^{-1}$ it will show a significantly larger degradation in IP resolution than the VP. The pixel-based detector also simplifies the tracking algorithms, as strip-based detectors require two strips to be matched in order to find the coordinate of a hit; avoiding strip matching will save processing time in the software trigger, and this leads to lower thresholds in the LLT and therefore higher selection efficiencies for interesting events. Further, because of the more complex tracking algorithms required for the strip detector, the fraction of ghost tracks (tracks reconstructed from randomly combined hits) is significantly higher for the VL; the ghost rate is about $\sim$13 % for the VL compared to 2 % to 3 % for the VP. The reduction in the number of ghost tracks also reduces the processing time needed in the software trigger.

In order to test the microchannels, several tests were performed; for example, the pressure was successfully cycled over a thousand times from 1 bar to 200 bar using a prototype, without any leakage [94]. Because these tests were so successful, the VELO group decided that the microchannels could safely contain the pressure as they would nominally only be operated at 20 bar to 30 bar [89]. In the VP, use of microchannels reduces the amount of material traversed by particles by about 26 % when compared to the pocofoam option.

3.3 The upgrade VELO detector

Each VP module will contain 4 planar silicon sensors, which will be used to collect charge deposited by ionising particles. The sensors are shown in Fig. 3.8, where the final design of the upgrade VELO modules is shown; initially the sensors were perpendicular to the x and y-axes (non-rotated), but due to issues discussed in Section 3.7 the sensors were rotated to the positions shown in this figure. The sensors will need a bias voltage of 1000 V after the radiation damage caused while collecting an integrated luminosity of $50\,\text{fb}^{-1}$ [89]. It is difficult to have this high voltage without sparking occurring between the sensor and ASIC, so a guard ring of 450 µm has been specifically designed to reduce the risk of sparking. Deposited charge in each of the sensors will be read out by 3 VeloPix ASICs, which measure

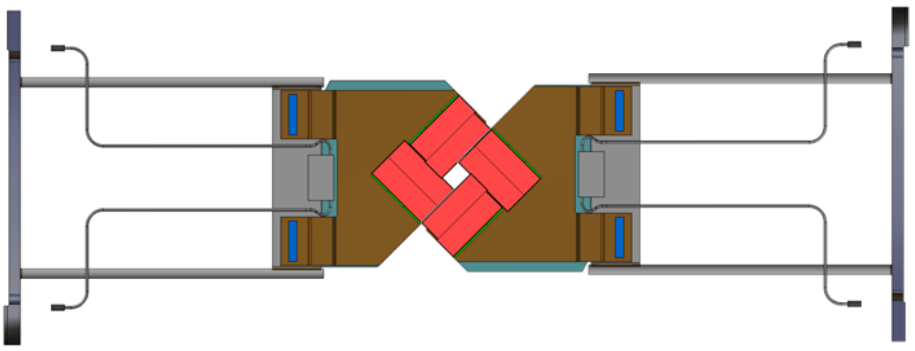

Figure 3.8: The upgraded VELO modules and one possible design of the supports which will hold the modules in place. The silicon substrate (turquoise) with microchannel cooling inside is shown, however it is mostly hidden beneath the hybrid (brown). The sensors (red) will be evenly distributed on opposite sides of the module and each sensor is read out by three ASICs.

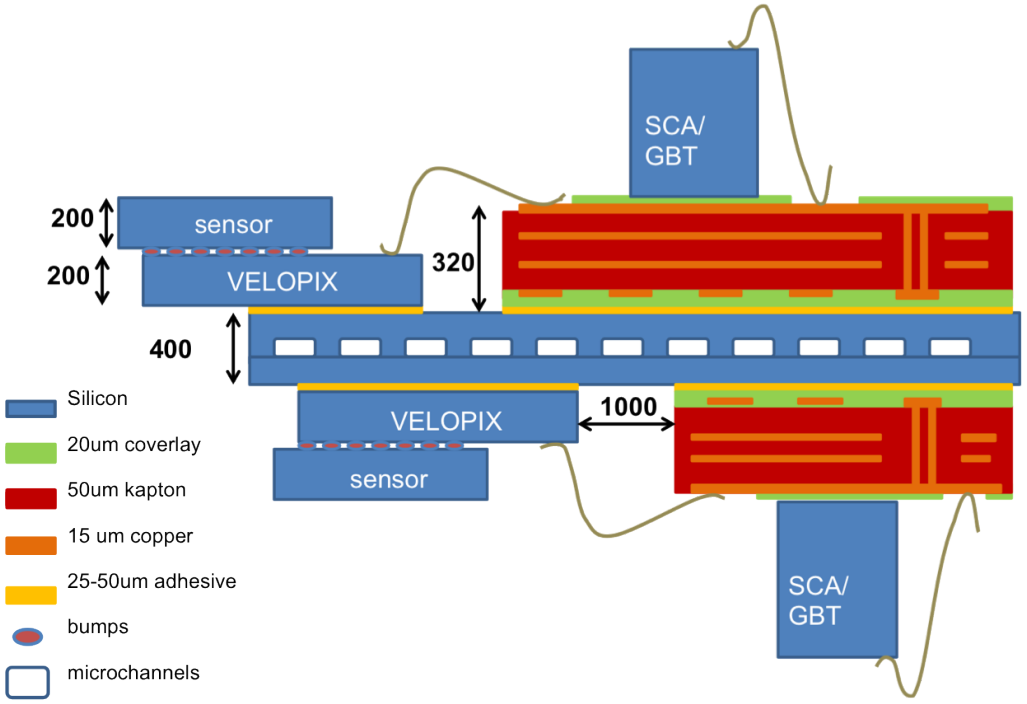

Figure 3.9: Cross-section of the pixel module (not to scale) with dimensions given in µm. It shows the sensors, VeloPix ASICs, GBT chips, microchannel silicon substrate, hybrids and bond wires. The diagram is taken from Ref. [89].

14.07 mm × 14.07 mm and contain 256 × 256 pixels. The VeloPix ASIC will use a binary data-driven readout of 4 × 2 super-pixels. The ASIC and sensors will be bump-bonded and together glued onto the silicon substrate. The bump-bonded sensors and ASICs are placed on alternating sides of the substrate to more evenly spread the heat generated by these components; this is shown in Fig. 3.9. The microchannels in the silicon substrate are made by etching 200 µm wide and 120 µm deep channels into 260 µm thick silicon, which is then overlaid with 140 µm of silicon to seal the channels. Liquid CO_2 is pumped through these channels where it evaporates and cools the module. Thermal simulations of the microchannel cooling have been performed; these were used to determine that the tip of the sensor could

extend 5 mm from the edge of the substrate, which further reduces the material in this design.

3.4 RF-Foil simulation

As shown in Fig. 3.7, the RF-foil is the largest single contribution to the material which is traversed by particles in the VELO; this makes its design critical. The RF-foil is shaped to fit tightly around the VELO modules and is corrugated to minimise the amount of material which particles pass through; this leads to a very complicated shape as pictured in Fig. 2.9. For initial simulations of the VP a relatively simple shape was used for the RF-foil.

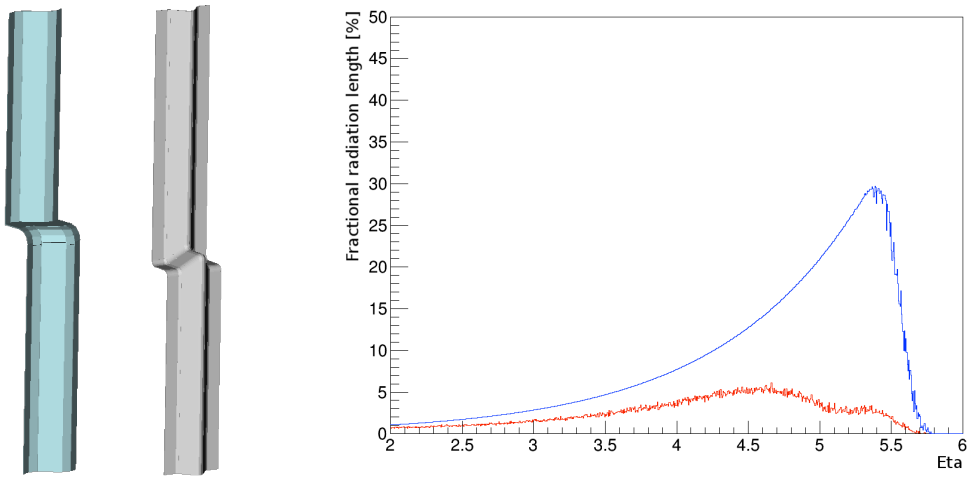

Figure 3.10: The two left images show single units of different RF-foils from the simulation. The left hand unit (light blue) is only partially corrugated (there is a flat region near the bend), while the right hand unit (grey) is fully corrugated. The right plot shows the fraction of a radiation length traversed by a particle in the VL before the first hit for a fully-corrugated RF-foil (red) and a tube-shaped non-corrugated RF-foil (blue). The right plot was taken from Ref. [93].

This simplified version the RF-foil was not corrugated along its full length. To demonstrate the importance of corrugations, a corrugated RF-foil is compared to a simple tubular RF-foil for the VL, and the results are shown in the right plot of Fig. 3.10. When using a tubular RF-foil, significantly more material is traversed by particles before the first hit when compared to a corrugated RF-foil. A more realistic RF-foil that was corrugated along its full length was implemented in the simulation of the VP . The simplistic, partially corrugated RF-foil and the fully corrugated RF-foil units are shown in Fig. 3.10; these units are repeated many times to form the full RF-foil.

3.5 VeloPix z-layout

Where the VP modules are placed along the z-direction (known as the z-layout) is very important as it strongly affects the performance of the detector. The number and position of the VP modules was optimised using a fast ray-tracing simulation and a full simulation of the detector. The primary concern of this optimisation is to track particles within the geometric acceptance of LHCb; once this is fulfilled, the IP resolution of the tracks is optimised to ensure that the best possible performance can be obtained. However, there are other important aspects to consider. Much of the existing infrastructure for the current VELO will be kept the same for the upgraded VELO, which leads to two design constraints for the z-layout. The cooling plant will not be changed; given the design of the VP modules, a maximum of 52 modules can be cooled by the existing cooling plant. Further, the structure which holds the VELO will not be changed and so the VP modules cannot exceed the limits along the z-axis of the current VELO. The first and last modules in the current VELO are at $z = -315\,\text{mm}$ and $z = 750\,\text{mm}$.

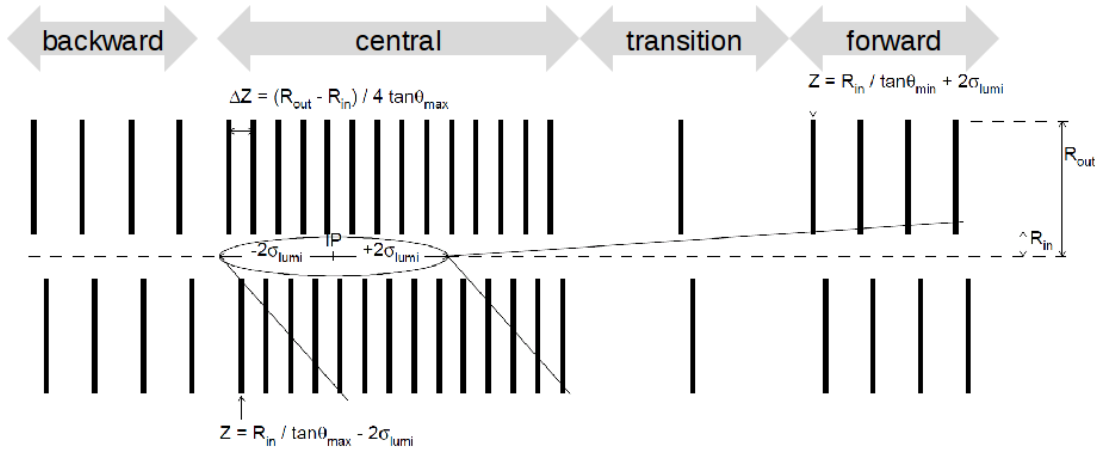

Figure 3.11: An example upgrade VELO layout. The figure introduces the nomenclature and the most important parameters which drive the choice of the z-layout. The diagram is taken from Ref. [93].

The lifetime of the K_S^0 ($\tau_{K_S^0} = 89.54\,\text{ps}$) means K_S^0-mesons are able to travel significant distances through the detector before decaying; the majority of K_S^0-mesons in LHCb decay after they have passed the end of the VELO. In order to maximise the yield of K_S^0 which can be detected in the VELO, the fourth from last VELO module should be as far as possible from the interaction region; this allows four hits to be made by charged particles produced from highly displaced vertices.

There are four groups of modules which will be referred to throughout this section, and these are shown in Fig. 3.11. The central modules ($-100\,\text{mm} < z < 300\,\text{mm}$) are tightly packed around the interaction region and are placed to detect tracks at low pseudorapidity. The backward modules ($z < -100\,\text{mm}$) are at negative z and these detect backward tracks. The forward modules ($z > 600\,\text{mm}$) are the four modules at largest z and are used to detect high pseudorapidity tracks. Finally the transition region modules ($300\,\text{mm} < z < 600\,\text{mm}$) are between the central and forward modules.

3.5.1 Geometric considerations

Initially trigonometry can be used to calculate various aspects of the z-layout. The distance between the modules in the central region is affected by two main factors: the minimum η of the angular acceptance and the requirement that particles leave at least four hits. The central module spacing (Δz) is shown in Fig. 3.11 and is defined as

$$\Delta z = \frac{R_{\text{out}} - R_{\text{in}}}{4 \tan(\theta_{\text{max}})},$$

where $\theta_{\text{max}} = 269\,\text{mrad}$ (corresponding to $\eta = 2$) and R_{in} and R_{out} are the inner and outer radii of the sensors.

Unlike the current VELO, the area covered by the active silicon on the pixel detector is approximately square (see Figs. 3.4 and 3.8), and so two inner and outer radii are considered as shown in Fig. 3.12. The first situation considered is when the tracks are parallel to the x or y-axis ($x-y$ tracks) and so the inner and outer radii are $R_{\text{in}} = 5.1\,\text{mm}$ and $R_{\text{out}} = 33.3\,\text{mm}$ respectively. The next situation considered is when the tracks are at 45° to the x or y-axis (45° tracks) and so the inner and outer radii are $R_{\text{in}} =$

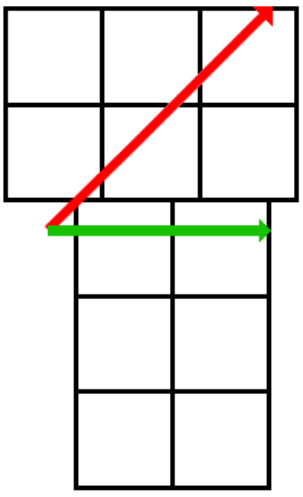

Figure 3.12: The outer radii distance of the two different track angles considered overlaid on a diagram indicating the position of the ASICs on the VP modules. The red arrow is at 45° to the x-axis and demonstrates the angle of 45° tracks, while the green arrow is parallel to the x-axis and demonstrates the angle of $x-y$ tracks.

7.2 mm and $R_{\text{out}} = 47.1$ mm respectively; using these values one obtains $\Delta z = 25.6$ mm for $x-y$ tracks and $\Delta z = 32.6$ mm for 45° tracks.

The central region should start and end at z_{start} and z_{end}; these are given by the equations

$$z_{\text{start}} = \frac{R_{\text{in}}}{\tan\theta_{max}} - 2\sigma_{\text{lumi}} \quad \text{and} \quad z_{\text{end}} = \frac{R_{\text{in}}}{\tan\theta_{max}} + 2\sigma_{\text{lumi}} + 4\Delta z\ .$$

Solving these numerically one finds that the central modules should start at $z_{\text{start}} = -108$ mm for $x-y$ tracks and $z_{\text{start}} = -100$ mm for 45° tracks, while the central modules should end at $z_{\text{end}} = 247$ mm for $x-y$ tracks and $z_{\text{end}} = 296$ mm for 45° tracks.

The forward region should begin at a z-position such that it can measure tracks with the smallest polar angle within the angular acceptance $\theta_{min} = 13.4$ mrad ($\eta = 5$) and which originate from $z = 2\sigma_{\text{lumi}}$. This position is determined using

$$z_{\text{forward}} = \frac{R_{\text{in}}}{\tan\theta_{\text{min}}} + 2\sigma_{\text{lumi}},$$

which is calculated to be $z_{\text{forward}} = 504$ mm for $x-y$ tracks and $z_{\text{forward}} = 660$ mm for 45° tracks.

Putting all these different values together one can create a basic layout as shown in green in Fig. 3.13; these positions will be referred to as Layout 1 and the exact module positions are listed in Appendix A.1. Layout 1 is further refined in the next section with a ray-tracing study. At this point backwards tracks have yet to be considered and so no backwards modules are shown in this layout.

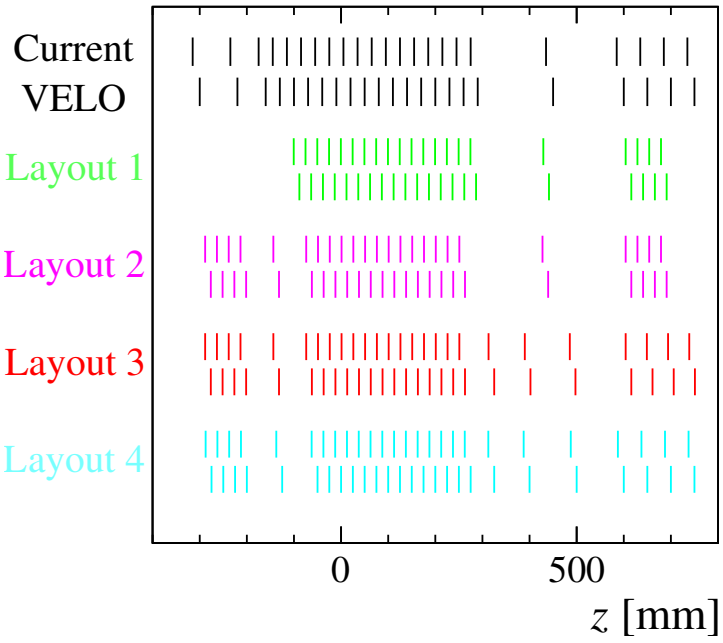

Figure 3.13: A comparison of various VELO z-layouts. All module positions are listed in Appendix A. The final z-layout used is Layout 4.

3.5.2 Ray-tracing simulations

A ray-tracing simulation is used to determine a minimal set of modules which satisfies the geometric acceptance requirements. In this ray-tracing study some approximations are made about the angular distribution of tracks and their origin; specifically it is assumed that tracks have a uniform distribution in η and ϕ and that the track origin z-position follows a Gaussian distribution with a mean at $z = 0\,\text{mm}$ and with a width of $\sigma_{\text{lumi}} = 63\,\text{mm}$.

The position of the central region, the position of the forward region and the central module spacing are varied until the layout with the smallest number of modules is found which covers the full geometric acceptance, with >99 % efficiency. A 48-module layout is chosen, where the modules in the central region are spaced 25 mm apart, spanning the region from $z = -74\,\text{mm}$ to $z = 263\,\text{mm}$, with the first forward module positioned at $z = 604\,\text{mm}$. This design will be referred to as Layout 2 and is shown in magenta on Fig. 3.13. The exact positions of Layout 2 are listed in

Appendix A.2. Compared to Layout 1, the central region of Layout 2 has four fewer modules and is shifted forward slightly; despite having fewer modules in the central region it maintains the same geometric acceptance. Layout 2 also adds 10 modules to measure backwards tracks, where the placement was chosen to maximise the number of detected tracks in this ray-tracing simulation. The ray-tracing simulation predicts an average of 7.9 hits per track and >99 % of tracks within the geometric acceptance passing through four sensors. This simulation also predicts the average first, second and third hit radii to be 7.2 mm, 10.0 mm and 12.5 mm. Having achieved the required geometric acceptance with the fewest possible modules, in the next section the IP resolution is optimised using the full LHCb simulation.

3.5.3 Full simulations

The use of the full simulation allows physics performance quantities to be produced for each layout which can then be used to optimise the layout further. In this simulation the centre-of-mass energy is $\sqrt{s} = 14\,\text{TeV}$ and the mean number of pp-interactions per bunch crossing (ν) is 7.6.

In Layout 2, the IP resolution of tracks with a first hit at $z > 300\,\text{mm}$ is substantially worse than tracks with a first hit in the central region. As discussed in Section 3.1, the IP resolution of tracks can be improved by reducing the distance from the first hit to the track origin; to this end four further modules were added in the transition region. This alone gives approximately a 13 % improvement in the IP resolution for tracks with a first hit at $z > 300\,\text{mm}$. Increasing the spacing between the last 4 module-pairs to 45 mm gives a further 14 % improvement in the IP resolution of these tracks. Layout 3 is created with these two improvements in mind; the

module positions of Layout 3 are listed in Appendix A.3 and shown in red in Fig. 3.13. The overall effect of the changes can be seen in Fig. 3.14, where it is shown that the IP resolution for tracks with a first hit at $z > 300\,\text{mm}$ is significantly improved with Layout 3. Despite adding four extra modules in the transition region, there is little change in the IP resolution of tracks with a first hit at $z < 300\,\text{mm}$. Overall, the IP resolution of Layout 3 is improved by approximately 7 % compared to Layout 2.

The average first hit radius in the full LHCb simulation is 7.9 mm, which can be compared to that calculated previously 7.5 mm from the ray-tracing simulation; these values are very similar given that the ray-tracing simulation simplifies many aspects of the simulation.

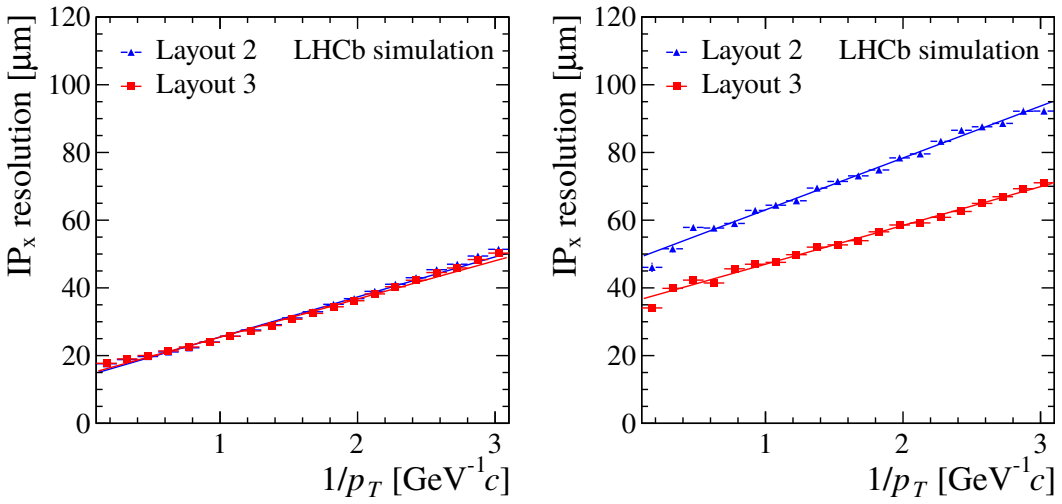

Figure 3.14: IP resolution in the x-direction for Layout 2 (blue triangles) and for Layout 3 (red circles). This is shown for tracks with first hit $z < 300\,\text{mm}$ (left) and tracks with first hit $z > 300\,\text{mm}$ (right). The lines show the line of best fit for a linear function. In the left plot the blue points and line are underneath the red points and line.

3.5.4 Regularised layout

The RF-foil would be simpler to produce if the z-layout required that all modules were at multiples of 12.5 mm from the origin. A new regularised layout, referred to as Layout 4, was created with modules positioned as listed in Appendix A.4 and this layout is shown in cyan in Fig. 3.13.

Using the ray-tracing simulation, it can be seen that the geometric efficiency changes when the layout is regularised. Layout 3 has an overall geometric efficiency of 99.9797 ± 0.0050 %, whereas the regularised layout in Appendix A.4 has an efficiency of 99.9486 ± 0.0050 %; this is about the same as reconstructing one fewer track in every 3200 tracks.

The geometric efficiency, IP resolution and PV resolution are studied in the full simulation with Layouts 3 and 4 to ensure that this change does not

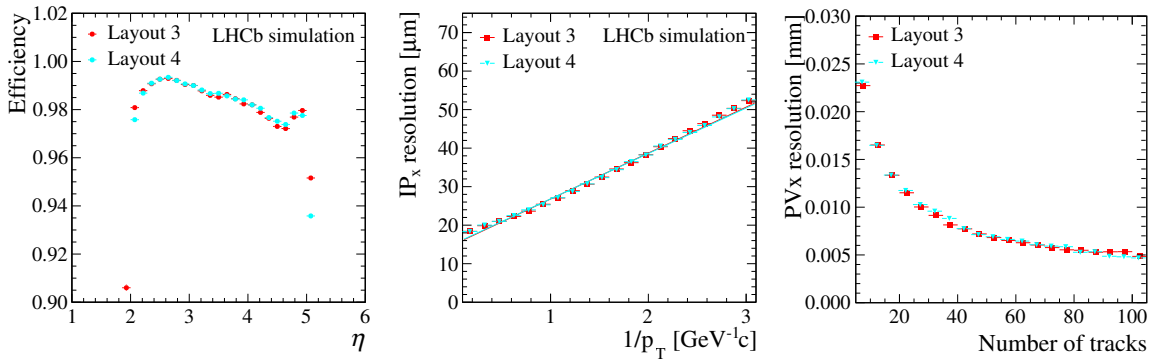

Figure 3.15: Comparison of the performance of Layout 3 (red) and Layout 4 (cyan) using the full simulation. The left plot shows the three-hit geometric efficiency from simulation as a function of pseudorapidity (η). The middle plot shows IP resolution as a function of inverse p_T. The right plot shows the PV resolution as a function of the number of tracks in the PV. In most places, the red points are almost directly underneath the cyan points.

affect the performance of the detector; the results can be seen in Fig. 3.15. The geometric efficiency from the full simulation shows the same effect from regularisation as was seen in the ray-tracing study; the geometric efficiency is slightly lower at the edges of the angular acceptance in η. The PV and IP resolution for Layouts 3 and 4 are similar.

Overall the ray-tracing and full simulations show only very small differences between the fully-optimised layout (Layout 3) and the regularised layout (Layout 4). Since it is desirable to use Layout 4 in order to make the design simpler and it has only a minor impact on the physics performance of the VP, the VELO group chose this layout as the final z-layout for the VP.

3.6 Sensor overlap

Given that the sensors are on alternating sides of the substrate and the thickness of the substrate plus the ASICs is about 800 µm, there is a gap through which particles could pass undetected, as demonstrated in Fig. 3.16. In the current design, where possible, the active areas of sensors have been designed to overlap by two pixels (which is equal to 110 µm). This number was an estimate and so the ray-tracing simulation was used to optimise the size of this overlap.

Figure 3.17 shows the geometric efficiency of tracks within the upgrade acceptance as a function of the overlap. The plot shows that the geometric efficiency linearly increases with overlap until around 220 µm at which point it levels off and then slowly decreases. From this an overlap of around 220 µm is found to be optimal; this is equal to the width of 4 pixels. Although 220 µm is the optimal overlap, changing the overlap from 110 µm to 220 µm

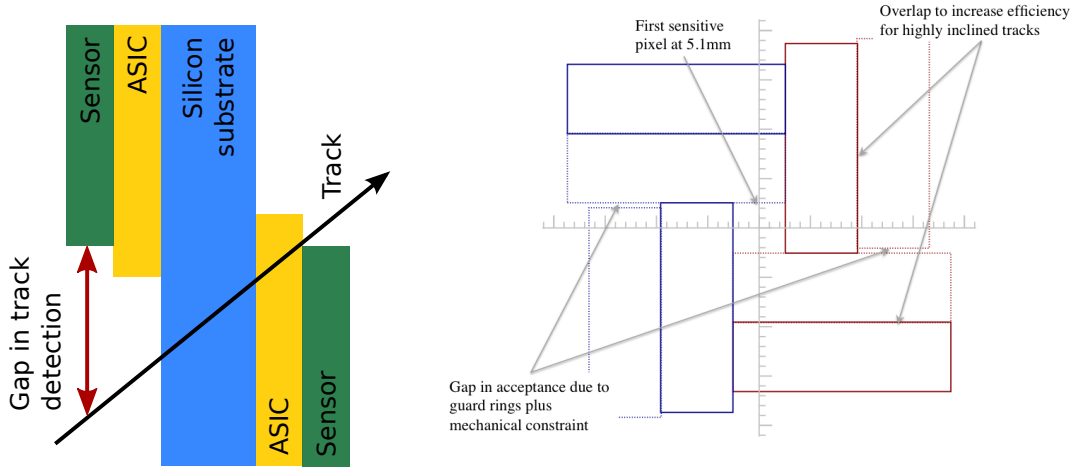

Figure 3.16: Diagrams of the pixel module design. The left diagram shows a cross-section through the module showing how tracks could escape detection. The right diagram shows the positions of the sensors in the $x-y$ plane and where these have been overlapped [89]; dashed rectangles show sensors on the far side of the module, while the solid rectangles represent sensors on the near side of the module.

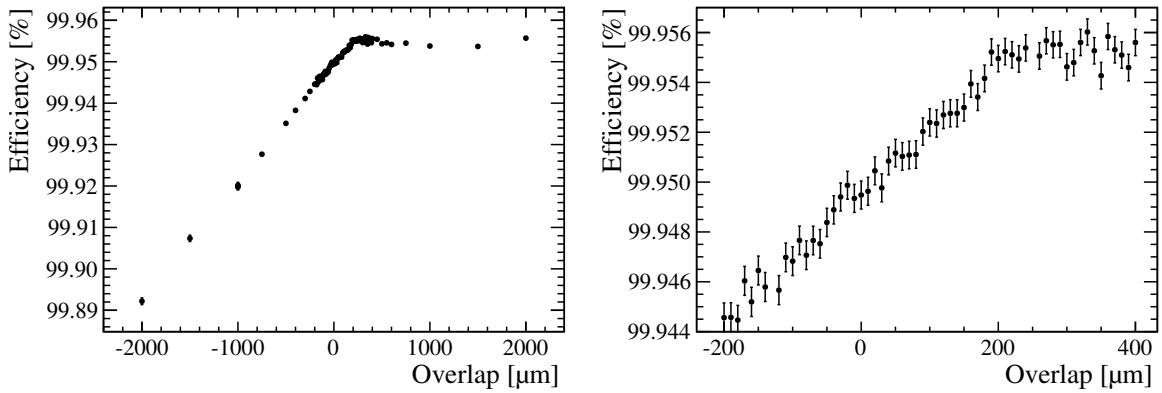

Figure 3.17: The fraction (in percent) of generated tracks within the geometric acceptance which leave 4 hits as a function of the overlap between the sensors. The right plot zooms into the region of interest around an overlap of 220 μm, which is the optimum.

only improves the overall geometric efficiency by 0.004 % (this equates to being able to detect one extra track in 25 000) and so it is a very small effect. Since this is such a small effect, the VELO group decided not to change the overlap to four pixels. It should be noted that this overlap could be useful to align sensors on the same module.

3.7 Simulations of a rotated VELO

During installation, the modules are slid into position along bearings attached to the inside of the RF-box. As the modules approach the end of the bearings they are positioned very precisely, but up until that point the position of the module is less certain. With the current geometry there is a length of several mm where the sensor would need to be 50 µm from the RF-foil without touching it; this is not possible given the tolerances of the installation system. One possible solution would be to rotate the VP sensors by 45° around the *z*-axis, as shown in Fig. 3.18; in this way the sensors would not be close to the RF-foil until they reached their final positions.

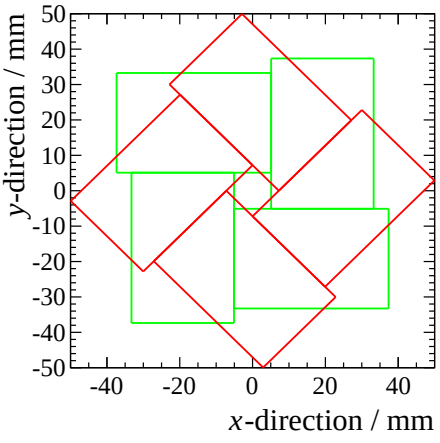

Figure 3.18: Approximate geometry of the non-rotated upgrade VELO sensors (green) compared to the rotated upgrade VELO sensors (red). The rotation is simply a rotation about the *z*-axis (the beamline) of 45°.

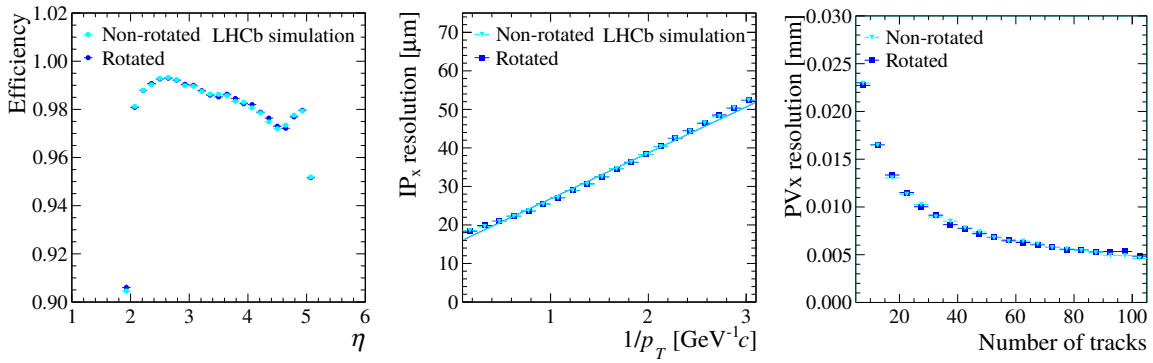

Figure 3.19: Comparison of the performance of the rotated (blue) and non-rotated (cyan) VP geometries. The left plot shows the three-hit geometric efficiency from simulation as a function of pseudorapidity (η). The middle plot shows IP resolution as a function of inverse p_T. The right plot shows the PV resolution as a function of the number of tracks in the PV. In most places, the blue points are almost directly underneath the cyan points.

The z-layout was designed such that the geometric efficiency was flat in ϕ and so rotating the sensor positions should have no effect on the performance. This is shown in Fig. 3.19, where the full simulation is used to demonstrate that the geometric efficiencies, IP resolutions and PV resolutions remain the same before (cyan) and after rotation (blue).

3.7.1 Partially open VELO performance

The current VELO does not close as far as it was designed to. The devices which prevent the two halves of the VELO colliding were placed such that the VELO closes by around 200 µm less than planned. If this were to happen in the VP there would be a reduction in the geometric efficiency; this can be simulated in the ray-tracing simulation by opening the VP in just the x-direction (to simulate a partially open non-rotated VP) or opening it along

Table 3.1: Comparison of geometric efficiencies from ray-tracing simulations when the VP is partially opened. It also compares the geometric efficiency when the VP is rotated or not rotated.

Distance open	Non-rotated modules	Rotated modules
0.0 mm	$99.949 \pm 0.052\,\%$	
0.5 mm	$99.550 \pm 0.052\,\%$	$99.651 \pm 0.052\,\%$
1.0 mm	$98.697 \pm 0.052\,\%$	$98.954 \pm 0.052\,\%$

the direction 45° to the x and y axes (to simulate a partially open rotated VP). The results of these ray-tracing studies are summarised in Table 3.1. The table shows that the rotated modules have a higher geometric efficiency than the non-rotated modules when the VP is partially open.

If the data rate during Run 3 is too high for the VP to cope with, the VP could be partially opened to reduce the occupancy and the data rate. However, opening the VP moves the modules further away from the beam and so will reduce the IP resolution. A full simulation study was performed to find the effect of a partially open VP on the IP resolution. Figure 3.20 shows the IP resolution when the VP is opened by various distances and for rotated and non-rotated modules. The direction in which the VP is opened is not reflected in the IP resolution; the IP resolution is the same in both the x and y-directions even if the VP is only opened along the x-direction. It is only the distance by which the VP is open that affects the degradation of the IP resolution.

Since the data rate is proportional to the number of hits in the detector, the number of hits can be used as a proxy to study the data rate. The

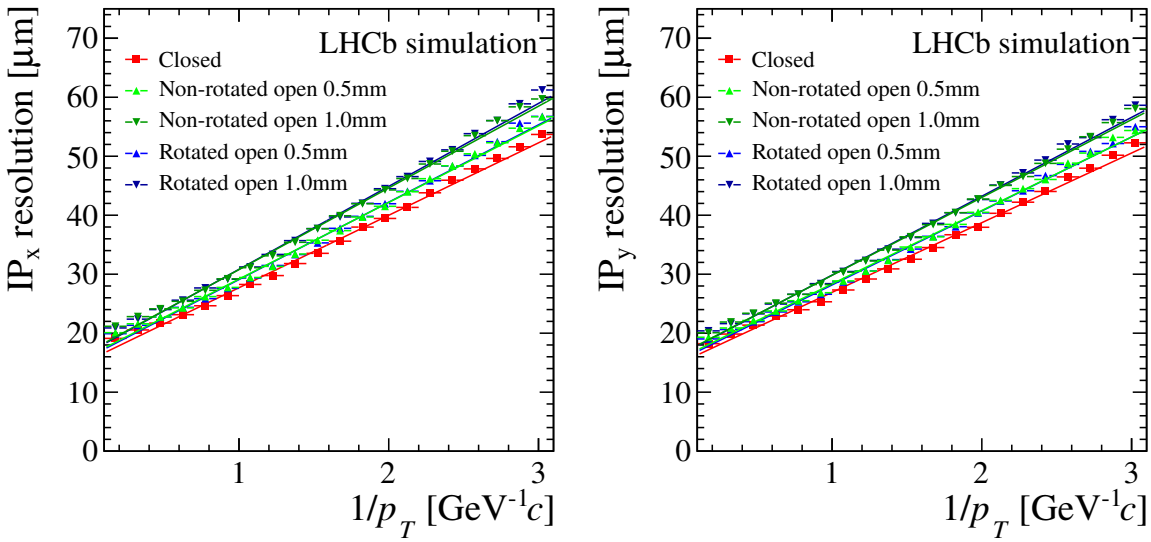

Figure 3.20: The IP resolution in the x-direction (left) and y-direction (right) as a function of inverse p_T. The red squares show the performance of a closed VP. Blue and green shapes show the performance of a rotated and non-rotated VP. Upward and downwards triangles show the performance when partially opened by 0.5 mm and 1.0 mm. The lines of best fit are also shown for the different detector configurations. The blue points are almost directly underneath the green points.

Table 3.2: Comparison of the relative number of hits when the VP is partially opened. The closed VP is defined to be 100% and the rotated and non-rotated VP options are compared to this.

Distance open	Non-rotated modules	Rotated modules
0.0 mm	100.00 ± 0.13 %	
0.5 mm	98.75 ± 0.13 %	96.90 ± 0.12 %
1.0 mm	95.10 ± 0.12 %	93.15 ± 0.12 %

reduction in the number of hits as a function of distance opened for the rotated and non-rotated modules can be seen in Table 3.2. This table shows that the rotated VP has a greater reduction in data rate for the distance opened. However, it is not the overall data rate of the VP which is important, but the data rate of the individual links from the VeloPix ASICs. There are

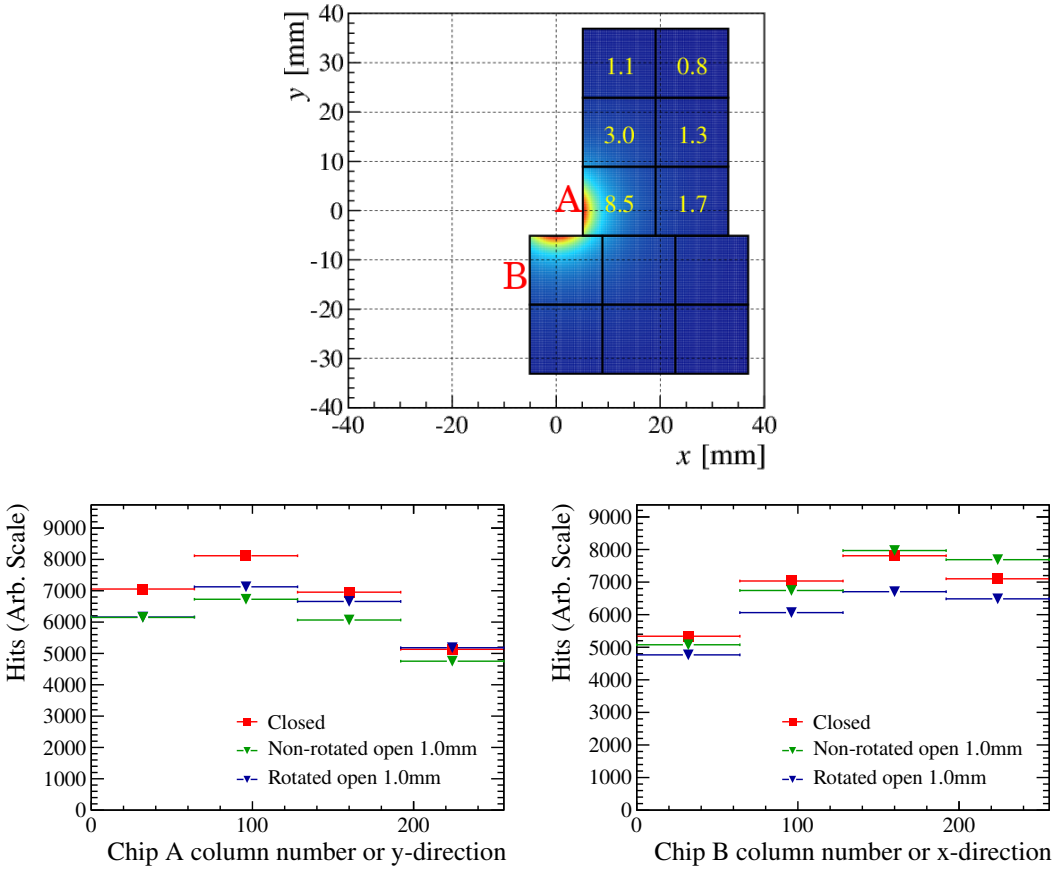

Figure 3.21: In the top plot chip A and B are labelled in red, while the average expected occupancy per event in the VeloPix chips is written in yellow, and the distribution of hits is shown in colour. The number of hits in chips A and B are shown in the bottom left and bottom right figures; the four bins represent the four data links in these chips. Red points show the closed VP, while the blue and green points show the rotated and non-rotated VP detectors open by 1.0 mm.

four data links in the two VeloPix ASICs closest to the beams; these two ASICs are labelled chips A and B in Fig. 3.21. Each data link reads out a quarter of the pixel columns; data link 0 reads out columns 0-63, data link 1 reads out columns 64-127, etc. Figure 3.21 shows what happens to the number of hits read out by each data link of chips A and B when the VP is partially opened. The nominal data rate of the closed VP is shown in red. When opening the non-rotated VP by 1.0 mm the data rate of chip A is reduced as expected, but chip B doesn't move further away from the beam and so the data rate in the two busiest data links actually increases. When the VP is rotated and opened by 1.0 mm, all of the busiest links show a reduction in data rate.

3.7.2 The decision to rotate

Rotating the VP causes no degradation in physics performance, but has several advantages if the VP needs to be run when not fully closed. The geometric efficiency of a 1.0 mm open VP would be higher if rotated, $(98.954 \pm 0.052)\%$, than compared to a non-rotated VP, $(98.697 \pm 0.052)\%$. The relative data rate when partially opened is less for a rotated VP, $(93.15 \pm 0.12)\%$, than compared to a non-rotated VP, $(95.10 \pm 0.12)\%$. Opening a non-rotated VP would make some data links above and below the beam busier; this does not happen for the rotated VP, where the data rate in all data links is reduced. Due to the issues with installing the modules in a non-rotated VP, the performance of the rotated VP when partially open and the fact that, when closed, rotating the VP has no effect on physics performance, the VELO group decided to rotate the modules in the VP. The final design of the rotated modules is shown in Fig. 3.8.

3.8 Upgrade and current VELO performance

The performance of the VP is shown in this section and is compared to the current VELO in upgrade conditions, where the centre-of-mass energy is $\sqrt{s} = 14\,\text{TeV}$ and the number of visible interactions per crossing is $\nu = 7.6$. This version of the VP simulation has: rotated modules, a rotated RF-foil which is fully corrugated along its length, the regularised layout (Layout 4) and realistically sized electronics for slow control of the VP (known as the GBT chip); this is the most recent and accurate simulation at the time of writing.

Figure 3.22 compares the three-hit geometric efficiency for the current VELO and the VP. The current VELO has a large variation in geometric efficiency as a function of ϕ. It is thought that this variation is caused by two

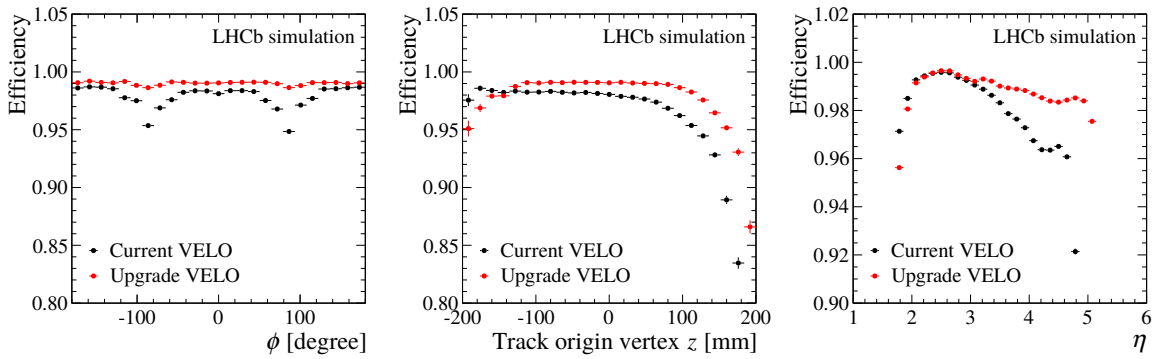

Figure 3.22: The efficiency of a long-lived charged particle leaving at least 3 MC hits (true hits) in the VELO as a function of various variables. The red and black points show the performance of the VP and the current VELO. The left, middle and right plots show the geometric efficiency as a function of the azimuthal angle (ϕ), track origin vertex z-position and pseudorapidity (η).

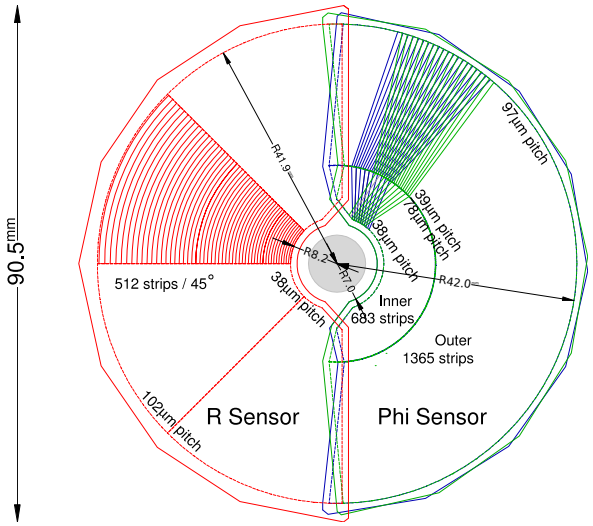

Figure 3.23: $r\phi$ geometry of the current VELO sensors [61]. For clarity, only a portion of the strips are illustrated. The strips of two ϕ-sensors on adjacent modules are shown in blue and green, while the r-sensor is shown in red.

effects. Firstly, as can be seen in Fig. 3.23, the inner edges of the sensors are not at a constant radius, as was originally planned when the z-layout of the current VELO was developed. In the region around $\phi = \pm 90°$, where the inner radius is greater, particles at high η are less likely to be detected and so contribute to the inefficiency at high η and $\phi = \pm 90°$. Secondly the current VELO is open by approximately 200 µm. In the nominal design shown in Fig. 3.23 there are small regions where the r and ϕ sensors don't quite overlap. These small areas where there is no overlap are made larger by the fact the VELO doesn't quite close; this also contributes to the inefficiency at $\phi = \pm 90°$. The variation as a function of ϕ is significantly smaller for the VP despite the square shape of the modules, since the z-layout was designed with this in mind. The VP geometric efficiency is also flatter in η and track origin vertex z-position than the current VELO.

Figure 3.24 compares the three-hit geometric efficiency for the current VELO and the VP with no requirement that the track originates from the region $|z| < 2\sigma_{\text{lumi}}$; this is plotted separately for particles which have decayed from bottom, charm and strange particles. There is a very good geometric

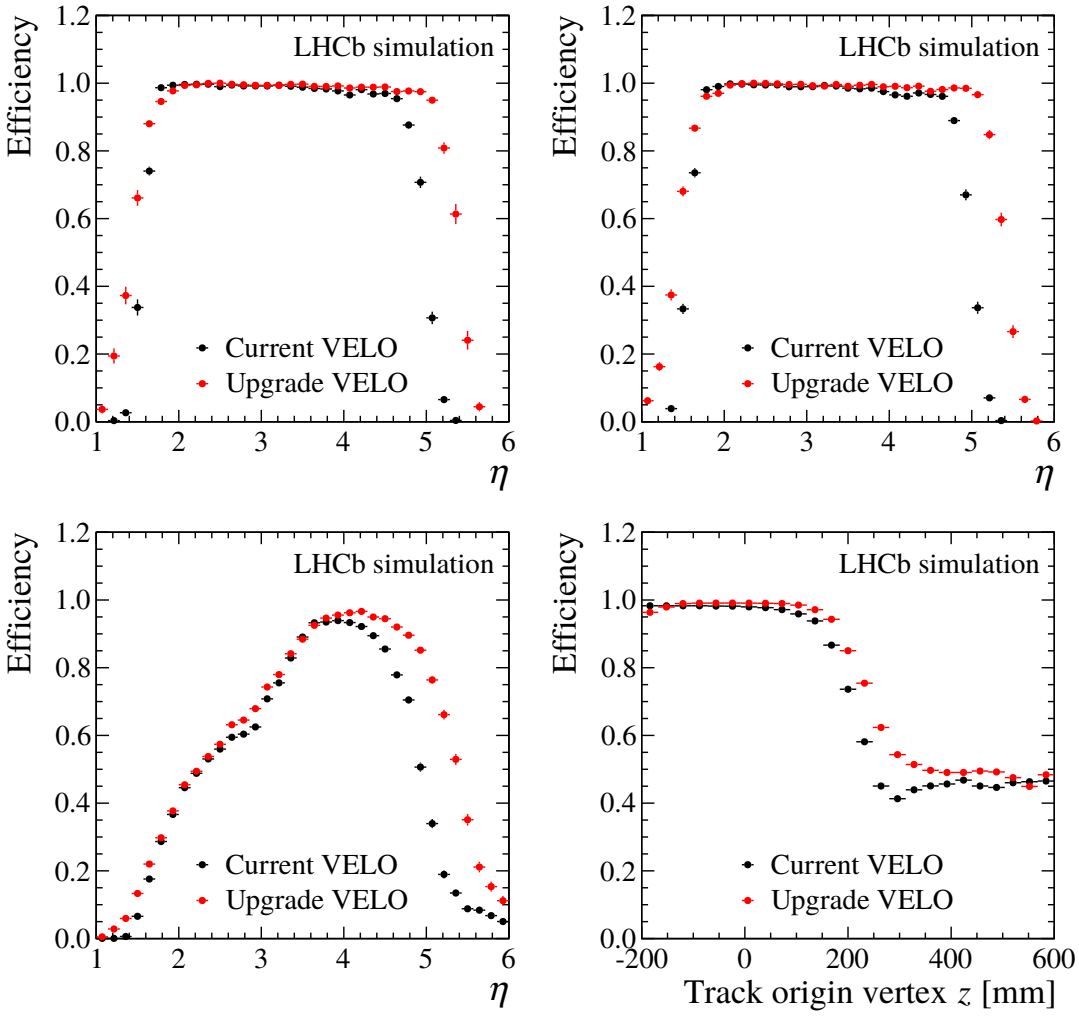

Figure 3.24: Efficiency for simulated long-lived charged particles to leave at least three true hits in the VELO with no requirement for the track to originate from within the region $|z| < 2\sigma_{\text{lumi}}$. The red and black points show the performance of the VP and the current VELO detectors. The top left, top right and bottom left plots show the geometric efficiency as a function of pseudorapidity (η) for particles which decay from bottom, charm and strange decays respectively. The performance shown in the bottom left plot appears poor, because strange particles often decay at particularly high z. The bottom right plot shows the geometric efficiency as a function of the track origin in the z-direction, where it shows that at very high z the VELO is unable to detect all tracks.

efficiency for particles from bottom and charm decays. The efficiency of strange decays appears poorer; this is due to the larger lifetime of strange particles, which means they travel significantly farther in the laboratory frame. The geometric efficiency for all particles as a function of the origin vertex z-position is also shown; for the VP this is > 95% in the region $-200\,\text{mm} < z < 200\,\text{mm}$ after which it quickly drops and flattens out to about 50% efficient in the region $300\,\text{mm} < z < 600\,\text{mm}$. The third from last VP module is at $z = 637.5\,\text{mm}$ and so this is the last point at which three hits could possibly be left. The VP can detect > 99% of particles up to a higher value of the track origin in the z-direction than the current VELO; this is important in case the interaction region is significantly displaced from the origin, as was the case during parts of 2011 and 2012. In almost all regions the VP has greater geometric efficiency than the current VELO.

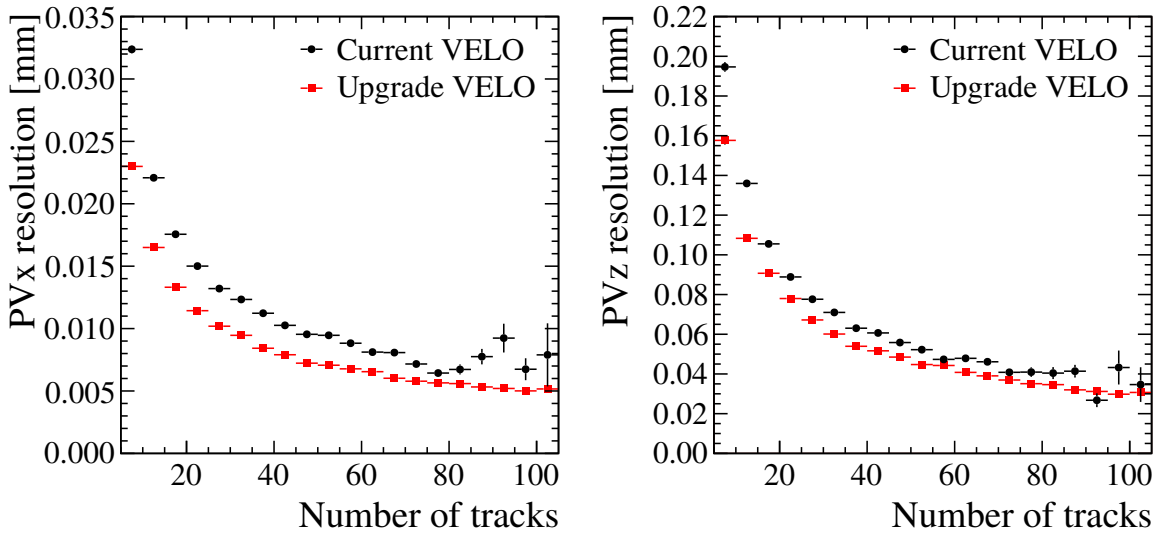

Figure 3.25: The PV resolution is shown as a function of the number of tracks associated to the PV for the upgrade VELO (red) and current VELO in upgrade conditions (black). The left plot shows the PV resolution in the x-direction, while the right plot shows the PV resolution in the z-direction.

Figure 3.25 shows the PV resolution as a function of the number of tracks associated to the PV for the VP and the current VELO. The VP PV resolution shows a significant improvement over the current VELO. The RMS of the PV resolution in the z-direction is $85.73 \pm 0.25\,\mu m$ and $109.03 \pm 0.36\,\mu m$ for the VP and the current VELO detectors respectively; this is an improvement of 21 %. The RMS of the PV resolution in the x-direction is $13.491 \pm 0.038\,\mu m$ and $19.608 \pm 0.065\,\mu m$ for the VP and the current VELO detectors respectively; this is an improvement of over 31 %.

Figure 3.26 shows the IP resolution as a function of inverse p_T, where the VP performs significantly better than the current VELO in the upgrade conditions. The multiple scattering term (gradient) of the fitted line is 11.708 ± 0.040 and 29.357 ± 0.089 for the VP and the current VELO detectors respectively; this is a reduction of 60 %. The IP resolution at infinite p_T is $15.173 \pm 0.056\,\mu m$ and $12.32 \pm 0.11\,\mu m$ for the VP and the current VELO detectors respectively.

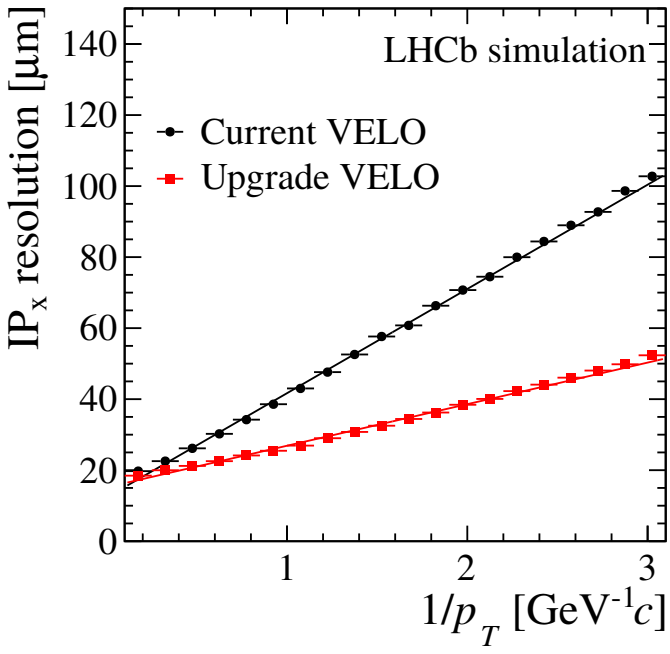

Figure 3.26: Comparison of IP resolutions for tracks which pass through the whole of LHCb as a function of inverse p_T for the upgrade VELO (red) and current VELO in upgrade conditions (black).

3.9 Conclusion

This chapter has described various aspects of the design, simulation, optimisation and performance of the upgrade VELO detector. First a decision was made to build a pixel-based microchannel-cooled detector, because it would be radiation hard, have the least amount of material for particles to traverse, require the least computationally expensive tracking algorithms and perform equally as well as the strip-based detector.

The process of optimising the placement of the sensors was started by using trigonometry to define the important parameters such as the inter-module spacing in the central region, the position of the first forward module and the extent of the central region. Using a ray-tracing simulation, a minimal set of modules was found which covered the geometric acceptance of LHCb and modules were added for reconstructing backwards tracks. A

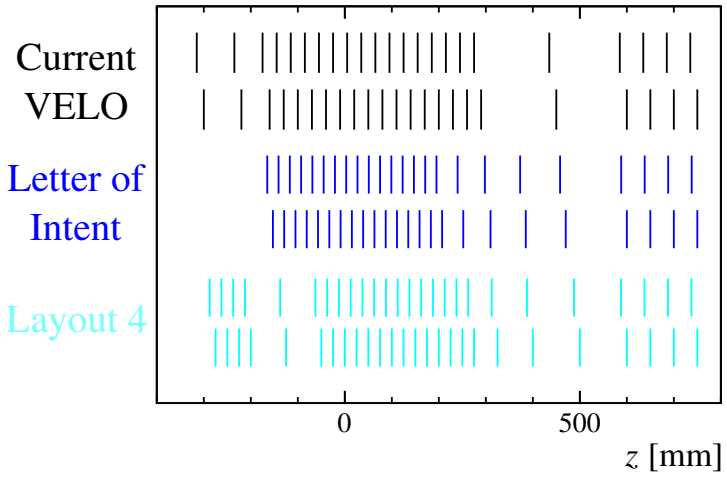

Figure 3.27: A comparison of VELO z-layouts. The first (top) layout is the current VELO (black). The second layout is the upgrade VELO layout proposed in the LHCb letter of intent [85] (blue). The third (bottom) layout is Layout 4 (cyan). All module positions are listed in Appendix A.

full simulation of the detector was used to optimise the layout and improve the IP resolution. To simplify the design, the modules were moved such that all modules are placed at multiples of 12.5 mm from the origin. This resulted in the module positions shown in Fig. 3.27, where they are compared to the current VELO layout and a proposed layout from the LHCb upgrade letter of intent [85]. This regularised layout (Layout 4) has been chosen by the VELO group as the final design of the z-layout.

The positions of the sensors in the xy-plane was also investigated to ensure that particles cannot pass between sensors. This is prevented by overlapping the edges of sensors where it is possible to do so. The study concluded that an optimal overlap would be 220 µm. Increasing the overlap from the nominal 110 µm to 220 µm would result in an efficiency increase of 0.004%; since this is such a small increase the VELO group decided to stay with the nominal overlap of 110 µm.

The sensors or RF-foil could be damaged if the two collided when sliding the modules into the RF-box during the installation process; rotating the sensors would solve this problem. Rotating the modules by 45° has benefits if the VP is required to run slightly open. The geometric efficiency is higher and the data rate is lower for an open rotated VP, than compared to the non-rotated VP; the lower data rate would be useful if there were particularly high occupancies and the data rate of the innermost data links became higher than their capacity. The VELO group decided to rotate the modules by 45° to solve the installation issue; this does not affect the performance of the VP.

The VP is compared to the current VELO in upgrade conditions, and the VP outperforms the current VELO in all metrics [95]. Geometric efficiencies

are increased and show less variation as a function of ϕ, η and origin of the track in the z-direction. The geometric efficiency of particles from bottom, charm and strange decays are higher for the upgrade when compared to the current VELO. Further the geometric efficiency as a function of track origin in the z-direction has been improved, which is important if the interaction region is displaced from the origin, as was the case during parts of 2011 and 2012. The average PV resolution in the z-direction has been improved by 21%. A typical particle with transverse momentum of $1\,\mathrm{GeV\,c^{-1}}$ will be reconstructed with a 35 % lower IP resolution in the upgrade than compared to the current VELO; this is caused by the 60 % reduction in the multiple scattering term of the IP resolution. The VP will provide excellent vertex reconstruction for LHCb during the upgrade.

CHAPTER 4

Electron identification and trigger efficiencies

Several results have shown that lepton universality may be violated [96, 97] and it is clear that in the neutrino sector lepton flavour is violated. To investigate lepton flavour it is essential that the full range of charged leptons are efficiently identified using the LHCb detectors. Muons are the easiest charged lepton to select since they are so penetrating, while electrons are much less penetrating. Electrons also radiate bremsstrahlung photons, which makes correctly reconstructing their momentum difficult. This chapter focuses on the measurement of the identification efficiency of electrons, along with various other improvements to the particle identification calibration (PIDCalib) software package, which measures PID efficiencies within LHCb.

4.1 PIDCalib methodology

PIDCalib determines the efficiency of PID requirements, but in order to do this, calibration datasets are required in which known particles have been selected without any PID requirements. To collect a dataset of well-identified particles without using PID, known decay modes of short lived particles are selected and their invariant mass distribution is fitted. The fit result is used to weight the selected candidates and candidates that are in mass regions where the signal probability density function (PDF) dominates are given high positive weights, while candidates in mass regions where there are more background events are given negative weights. The weighting is designed so that when a distribution that is not correlated to mass is studied, it is approximately the signal only distribution; this is known as ${}_s$weighting [98]. Using the ${}_s$weighted calibration sample, a histogram is created before and after the PID requirements. The histogram generally has three dimensions: momentum, pseudorapidity and track multiplicity. The ratio of the before and after histograms creates an efficiency map as a function of these variables.

The calibration datasets make use of a tag and probe methodology, where no PID is applied to the particle under study (the probe particle) and stringent PID requirements are made on the remaining particles (the tag particles). It is also important that the probe particle is not biased by the trigger, since the trigger will bias particles to higher momenta, and so we require that the event was triggered independent of the probe electron. To measure $\pi^{\pm}$ efficiencies, $D^{*+} \to D^0\left(K^-\pi^+\right)\pi^+$ decays are selected with a stringent PID requirement on the $K^{\pm}$ from the D^0; similarly $D^{*+} \to D^0\left(K^-\pi^+\right)\pi^+$ decays

are used to measure $K^{\pm}$ efficiencies, but with no PID requirements applied to the $K^{\pm}$, while the $\pi^{\pm}$ has stringent PID requirements applied to it. Decays of $J/\psi \to \mu^{-}\mu^{+}$ are used to measure $\mu^{\pm}$ efficiencies and $\Lambda \to p^{+}\pi^{-}$ decays are used to measure $p^{\pm}$ efficiencies. Before the work presented in this thesis, there was no sample in PIDCalib to measure electron efficiencies.

After the efficiency map has been created using the calibration sample, a reference sample of signal is used to determine the efficiency of the PID requirements on that signal sample; simulated events are often used as the reference sample, since they contain only signal. The PID efficiency is then calculated by averaging over the reference sample.

4.2 Electron calibration sample

Several previous LHCb studies [35, 97] have used the tag and probe method to measure the identification efficiency of electrons. A selection for $B^{+} \to J/\psi \left(e^{+}e^{-}\right) K^{+}$ decays is created, which is similar to those used previously. However, the requirements on the invariant masses of the J/ψ and B^{+} are altered in order to suppress partially-reconstructed backgrounds and B^{+} decays involving K^{*} mesons and excited J/ψ states. This selection also prevents the truncation of the low-mass bremsstrahlung tail of the J/ψ mass peak. The full requirements of this selection are listed in Table 4.1.

4.2.1 Electron specific variables

As discussed in Section 2.4.2, electron candidates are mainly identified with the calorimeter system. Tracks are extrapolated towards the ECAL and

Table 4.1: Requirements made to select $B^+ \to J/\psi \left(e^+e^-\right) K^+$ candidates for the calibration dataset. The B^+ J/ψ mass difference requirement removes K^* backgrounds, while the J/ψ mass requirement removes excited J/ψ states.

Particles	Requirement
Both electrons	$\chi^2_{\text{Track}}/N_{\text{d.o.f.}} < 5$ IP χ^2 from PV > 25
Tag electron	$p > 6\,\text{GeV}\,\text{c}^{-1}$ $p_T > 1.5\,\text{GeV}\,\text{c}^{-1}$ $\Delta\text{LL}(e-\pi) > 5$
Probe electron	$p > 3\,\text{GeV}\,\text{c}^{-1}$ $p_T > 0.5\,\text{GeV}\,\text{c}^{-1}$ Event triggered independent of the probe electron
J/ψ	Mass in range $2.25\,\text{GeV}\,\text{c}^{-2}$ to $3.6\,\text{GeV}\,\text{c}^{-2}$ Vertex $\chi^2 < 10$ Decay length significance > 5
$K^\pm$	$\Delta\text{LL}(K-\pi) > 0$ $\chi^2_{\text{Track}}/N_{\text{d.o.f.}} < 4$ $p_T > 1.0\,\text{GeV}\,\text{c}^{-1}$ IP χ^2 from PV > 9
$B^\pm$	IP χ^2 from PV < 9 Vertex $\chi^2 < 9$ Mass $4.2\,\text{GeV}\,\text{c}^{-2}$ to $6.0\,\text{GeV}\,\text{c}^{-2}$ B^+ J/ψ mass difference $2.1\,\text{GeV}\,\text{c}^{-2}$ to $2.3\,\text{GeV}\,\text{c}^{-2}$

if hits have been found in the PS, SPD and not in the HCAL, an electron candidate is created. Energy deposits in the ECAL are then searched for that could have originated from bremsstrahlung photons emitted by the

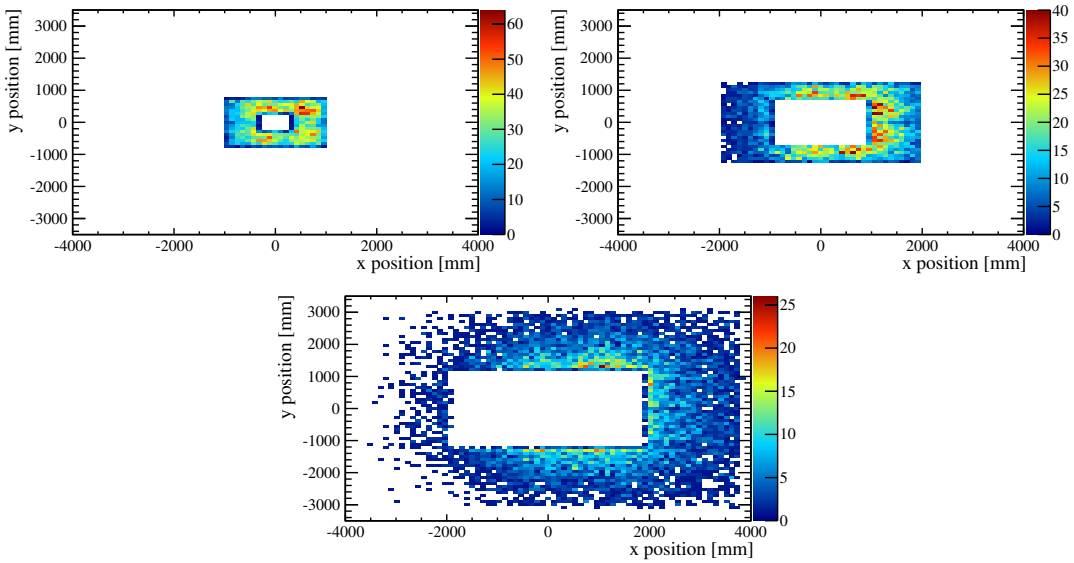

Figure 4.1: The number of e^- candidates as a function of their reconstructed position in the ECAL from magnet-down data. The three plots show only those candidates in the inner (top left), middle (top right) and outer (bottom middle) ECAL regions.

electron candidate; these energy deposits are added to the energy of the reconstructed electron.

The identification efficiency of electrons varies as a function of several electron specific variables. For example, the ECAL is divided into three distinct regions with different granularities (as discussed in Section 2.4.2); the three regions are shown in Fig. 4.1. Due to the higher granularity of the ECAL closer to the beampipe, there is a greater chance of correctly reconstructing an electron or its bremsstrahlung photon near the beampipe. The particle multiplicity is approximately flat in η and so there will also be more particles passing through the detector closer to the beampipe; this reduces the likelihood of reconstructing a bremsstrahlung photon closer to the beampipe. The region of the ECAL that the electron is reconstructed

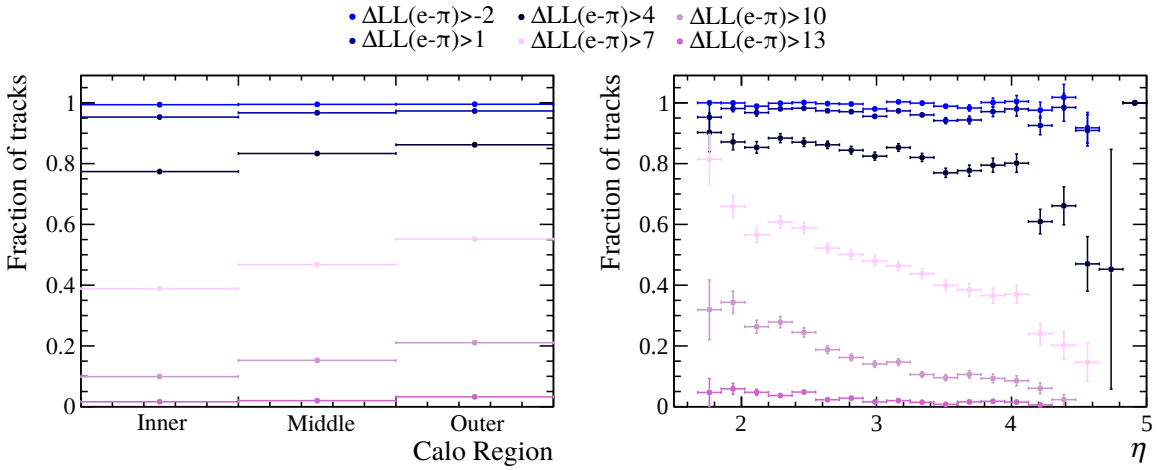

Figure 4.2: Fraction of electron tracks with a $\Delta LL(e-\pi)$ value greater than: -2 (blue), 1 (dark blue), 4 (black), 7 (light pink), 10 (dark pink) and 13 (pink). The left plot is shown as a function of the three ECAL regions: the left bin shows the innermost region and the right bin is the outermost region. The right plot is shown as a function of η; note that a high η track will most likely be reconstructed in the innermost region (first bin on the left plot), while a low η track will most likely be found in the outermost region (last bin of the left plot). The plots show a similar behaviour, except for a reflection along the x-axis.

in is highly correlated to the η of the track and so the dependence of the efficiency on the ECAL region can be accounted for by studying the efficiency as a function of the standard three variables (momentum, pseudorapidity and track multiplicity). The efficiencies of various electron PID requirements are shown in Fig. 4.2 as a function of the ECAL region and η. Both plots show very similar efficiencies, which decrease closer towards the beampipe.

Another electron specific variable that the PID performance depends on is whether a bremsstrahlung photon was added to the electron candidate. If a bremsstrahlung photon was added, it is more likely to pass an electron PID

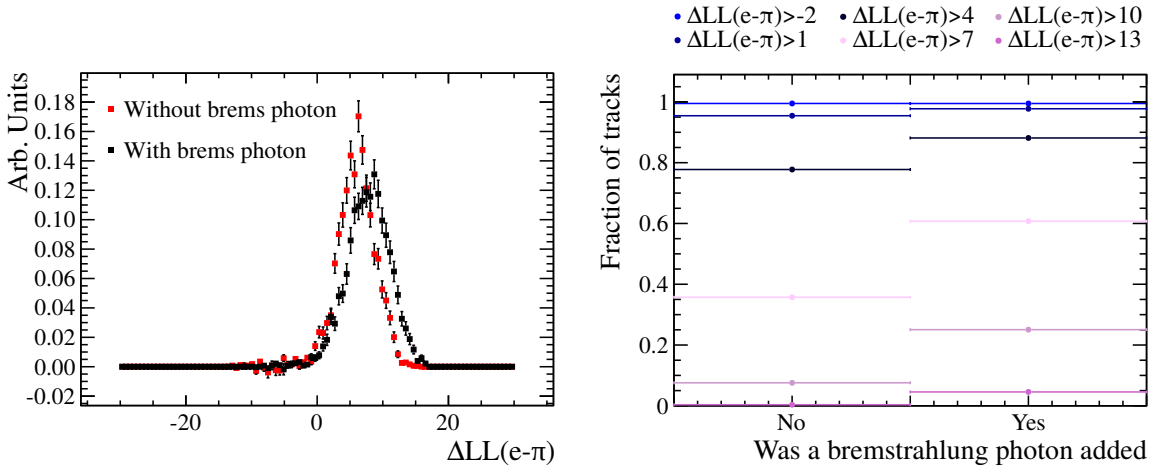

Figure 4.3: Left: $_s$weighted $\Delta\text{LL}(e-\pi)$ distributions for e^+ candidates reconstructed with (black) and without (red) a bremsstrahlung photon added, measured using the magnet-up 7 TeV dataset. Right: Fraction of electron tracks with a $\Delta\text{LL}(e-\pi)$ value greater than: -2 (blue), 1 (dark blue), 4 (black), 7 (light pink), 10 (dark pink) and 13 (pink); this is shown for simulated events with (left bin) and without (right bin) any added bremsstrahlung photons.

requirement; this is demonstrated in Fig. 4.3, where events reconstructed with a bremsstrahlung photon generally have a larger $\Delta\text{LL}(e-\pi)$ value. If the track multiplicity is higher in an event, then there is a greater chance of missing a bremsstrahlung photon that should have been assigned to an electron candidate or incorrectly assigning one to an electron candidate. Because of this, the event multiplicity and η are both highly correlated to the fraction of events with bremsstrahlung photons added and so once again calculating the efficiency as a function of the three standard variables will take into account the dependence of the efficiency on whether a bremsstrahlung photon was added to the candidate or not.

Ideally the number of hits in the SPD or RICH detectors would be used to study the multiplicity of the event, but this was not available in all of the

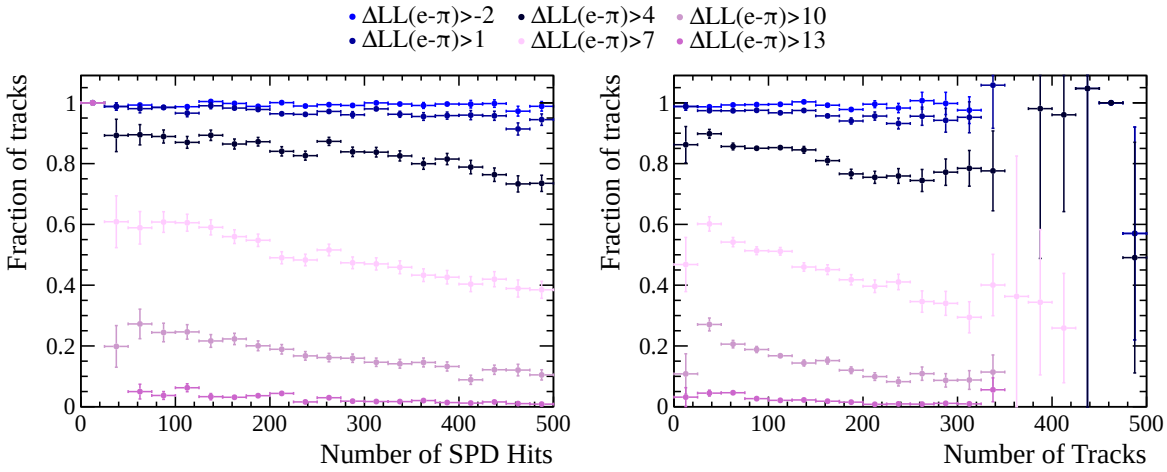

Figure 4.4: Fraction of electron tracks with a $\Delta LL(e-\pi)$ value greater than: -2 (blue), 1 (dark blue), 4 (black), 7 (light pink), 10 (dark pink) and 13 (pink). This is shown as a function of the number of hits in the SPD detector (left) and the number of reconstructed tracks (right).

calibration datasets at the time of these studies. Instead, the number of reconstructed tracks is used, but this is influenced by effects such as those from tracking, which could also depend on the track multiplicity in the event. However, it is assumed that this is a small effect for these studies. Figure 4.4 shows the fraction of candidates passing various PID requirements as a function of the number of reconstructed tracks and the number of hits in the SPD detector; both plots show a similar behaviour.

4.2.2 Calibration phase space overlap

In order to correctly measure the efficiency of a PID requirement on a certain reference decay mode, the momentum and pseudorapidity distributions of the calibration sample must fully cover that of the reference sample; if this is not the case, there could be regions of reference sample where there are

no entries in the efficiency map from the calibration sample. The $D^0 \to e^{\pm}\mu^{\mp}$ analysis (as discussed in Chapter 5) is taken as an example analysis in this section. The $p-\eta$ distributions of simulated $D^0 \to e^{\pm}\mu^{\mp}$ events and the electron calibration sample are plotted and compared in Fig. 4.5. In most analyses particles are required to have a momentum greater than a certain value and when this is combined with the angle of the track and the fact that there is a magnetic field in LHCb, kinematic limits are created in the $p-\eta$ distribution that run from the bottom left to the top of the plots in Fig. 4.5. The kinematic limit is less stringent in the calibration sample than in the reference sample; this ensures that, given a large enough calibration sample, efficiency measurements will be possible throughout the full phase space of the reference sample.

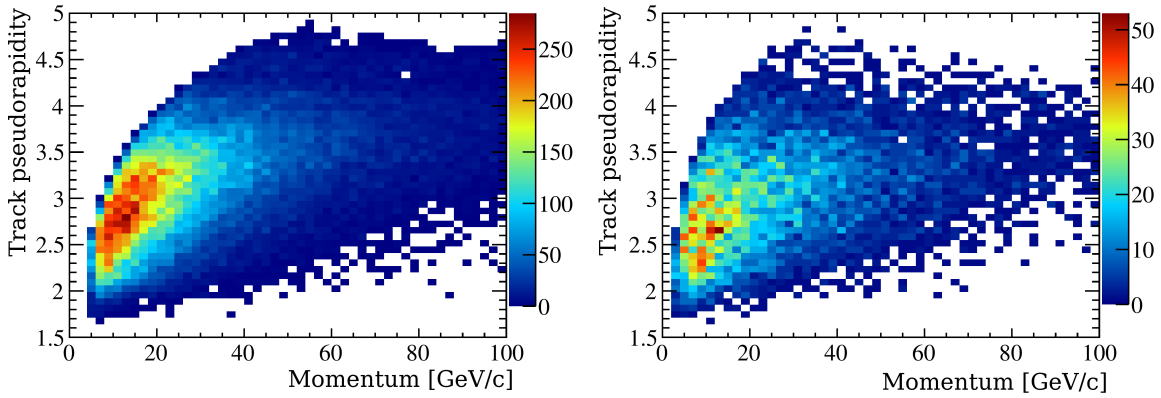

Figure 4.5: Distributions of the electron momentum and pseudorapidity from simulated $D^0 \to e^{\pm}\mu^{\mp}$ decays (left) and $B^+ \to J/\psi\left(e^+e^-\right)K^+$ candidates selected from data (right).

4.2.3 Fitting the calibration sample

It is necessary to fit the calibration sample in order to determine the $_s$weights. The data are fitted using a log-likelihood maximization program created using the Root package and the embedded RooFit framework [99, 100]. The calibration sample is split into two subsamples: electron candidates with and without an added bremsstrahlung photon; two different signal PDFs are used to fit the data in these two subsamples. When an electron is reconstructed with a bremsstrahlung photon, the J/ψ candidate mass spectrum is fitted using two CB functions [101] with a shared mean; one has low-mass tail to fit the energy loss through bremsstrahlung, while

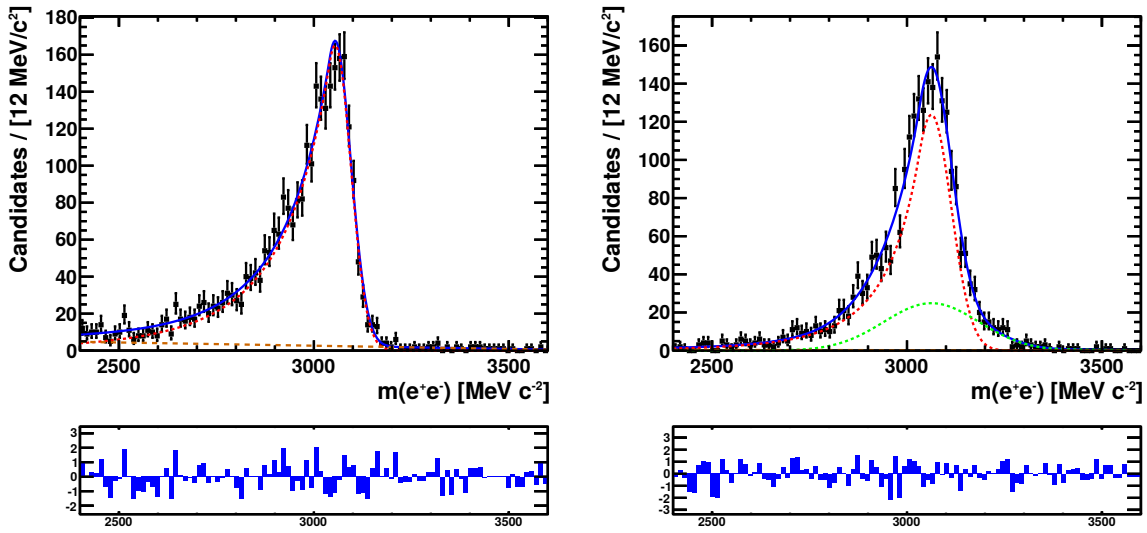

Figure 4.6: Fitted J/ψ candidate invariant mass distributions from magnet-up 7 TeV dataset. Candidates in the left figure have no reconstructed bremsstrahlung photons, while the candidates in the right plot have one or more. The dotted red and green lines show the signal CB functions, while the dashed orange lines show the linear Chebychev polynomial used to model the combinatorial background.

the other has a high-mass tail which fits events with an incorrectly assigned bremsstrahlung photon. If there is no bremsstrahlung photon assigned to the electron, only one CB function is used and it has a low-mass tail. Figure 4.6 shows an example fit result to the magnet-up 7 TeV dataset. A linear Chebychev polynomial is used to model the combinatorial background.

Figure 4.7 shows the unweighted and $_s$weighted $\Delta\text{LL}(e-\pi)$ distributions. The peak caused by background pions around the point $\Delta\text{LL}(e-\pi)=-7$ is removed by the $_s$weighting procedure; this shows the fit works well. Due to the very small amount of background for events with a bremsstrahlung photon (as shown in the right plot of Fig. 4.6), there is no background peak at $\Delta\text{LL}(e-\pi)=-7$ in the right plot of Fig. 4.7.

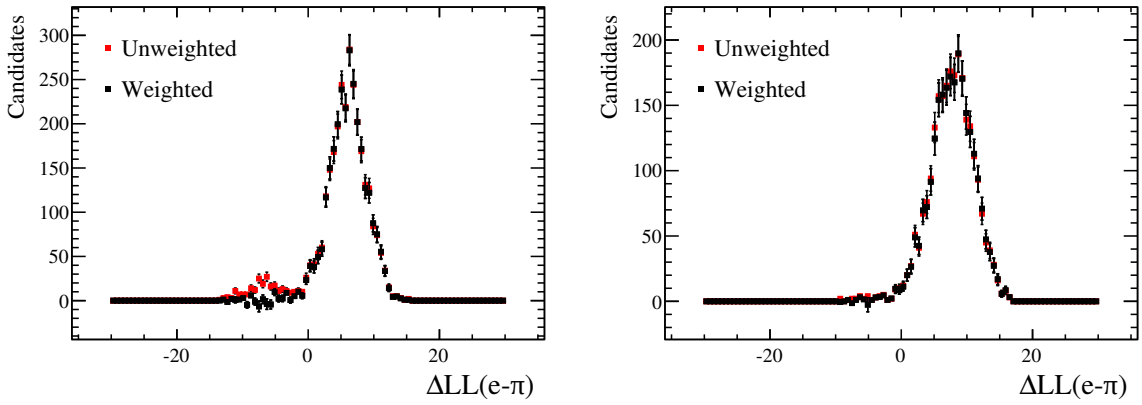

Figure 4.7: Distributions of $\Delta\text{LL}(e-\pi)$ for candidates with (left) and without (right) added bremsstrahlung photons. The red points are unweighted, while the black points are $_s$weighted.

4.3 Binning schemes

In order to accurately determine the efficiency of the PID requirement, a binning scheme is needed that can be used to create the efficiency map. If some bins are dominated by background (due to a low number of events in these bins), unphysical efficiencies can be found and so it is essential to ensure each bin in the efficiency map has sufficient events. It is also important to consider the RICH momentum thresholds (as listed in Table 2.2) as the identification efficiencies could sharply change at these points. One solution to a lack of events in certain bins is to create an equal-statistics binning scheme. In order to do this, the one-dimensional distributions of momentum, pseudorapidity and event multiplicity are individually divided such that there is an approximately equal number of events in each bin. Figure 4.8 demonstrates the binning scheme used for electrons in the $D^0 \to e^{\pm}\mu^{\mp}$ analysis. Due to the limited number of candidates available in the electron calibration sample, only 7 bins are used in each dimension. The plots show that the distributions of the calibration and reference samples are reasonably similar. The plots also show that generally the efficiency doesn't change too rapidly across the bins; an exception to this is the momentum distribution, where the identification efficiency quickly rises at lower momenta and because of this the binning scheme doesn't quite model the efficiency correctly. This introduces a systematic uncertainty into the $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$ measurement. In order to account for this, the number of bins is altered by ± 1 and the maximum difference from the nominal efficiency is taken as a systematic uncertainty. Scripts have been added to the PIDCalib package to simplify the creation of equal-statistics binning schemes.

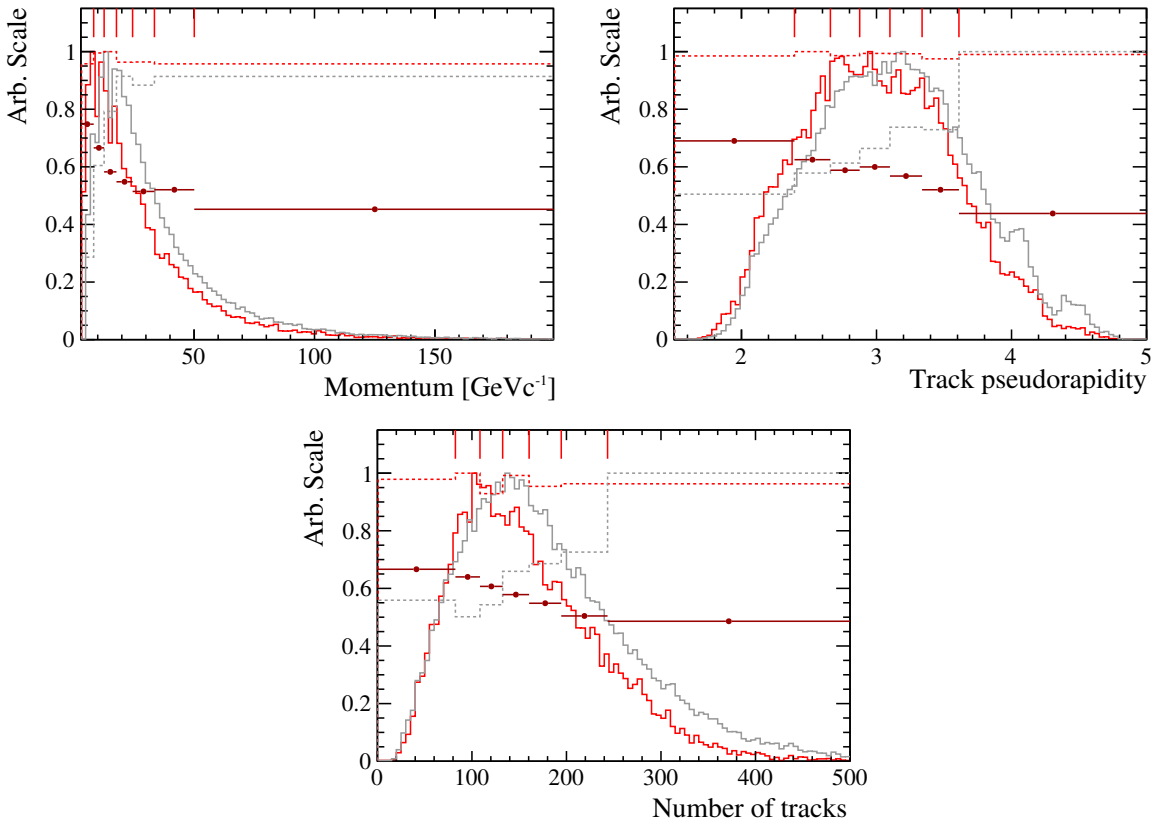

Figure 4.8: Example binning scheme for the electron calibration sample with 7 bins. The red solid histograms show the calibration sample and the grey solid histograms show the reference sample (signal $D^0 \to e^{\pm}\mu^{\mp}$ from simulation). The dashed red and grey histograms show the same data as their solid counterparts, but binned differently; they use the same binning scheme as used for the efficiency. The dark red points show the efficiency of $\Delta\text{LL}(e-\pi) > 6$, measured from the calibration samples. The red lines at the top of the figure divide the calibration samples into the 7 bins. The data are shown as a function of momentum (top left), pseudorapidity (top right) and the number of tracks in the event (bottom).

4.4 Single-track trigger efficiencies

There are several trigger categories into which reconstructed particles can fall, but the main two are: trigger independent of signal (TIS) and triggered

on signal (TOS). If a reconstructed particle is TIS, the event would have been triggered even if this particle were removed from the event. If a reconstructed particle is TOS, it would have triggered the event on its own. The TISTOS method can be used to calculate the efficiency of single-track triggers using data. First, a calibration sample is selected and $_s$weighted and then the selected tracks are required to be TIS; this ensures that all tracks considered were not biased by the trigger. A histogram is created before and after applying the TOS requirement and the ratio of the two histograms creates an efficiency map that measures the trigger efficiency. This is identical to the process in the PIDCalib software package and so the PIDCalib datasets were extended to include the trigger decisions; this allows PIDCalib to also measure the efficiency of single-track triggers. This method using PIDCalib cannot measure the efficiency of the more sophisticated HLT triggers that search for complex decay chains, as this would require a calibration sample with the same decay topology.

This method works well for the muon triggers which don't have any unusual variables that the efficiency depends on, but the calorimeter triggers are more complicated. The L0 hadron trigger is used to select $D^0 \to K^-\pi^+$ decays, which are used in the search for $D^0 \to e^\pm\mu^\mp$ as documented in Chapter 5. The calorimeter (CALO) group provide lookup tables [102] to determine the L0 hadron trigger efficiency as a function of both the HCAL region and the measured E_T in the HCAL. The HCAL region is highly correlated to η, while the E_T of a particle is highly correlated to both p and η; this means, once more, the three standard PIDCalib variables should correctly account for the dependence on the HCAL region and E_T. To ensure this is the case a simulation-based study was carried out and some results are shown in Fig. 4.9. In these plots it can be seen that this method correctly

measures the efficiency; the true value from the simulation closely matches that determined by PIDCalib on the simulation.

Since the L0 hadron trigger is activated by having a large amount of energy deposited in an HCAL cell, the more particles there are in an event, the more likely it is that an additional particle could also deposit energy in the same HCAL cell as the triggered particle. This means that as the track

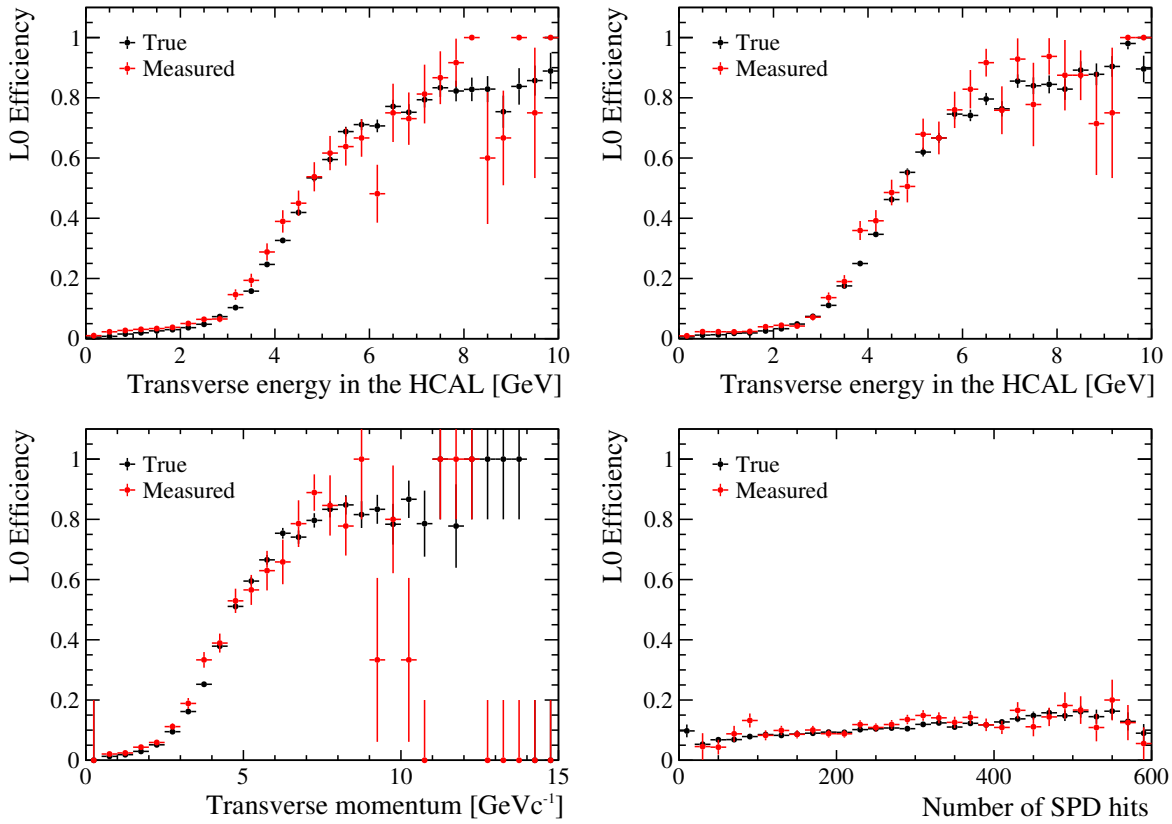

Figure 4.9: L0 hadron trigger efficiencies determined using simulated events. The black points show the true value and red points are determined using the PIDCalib method. The top row shows the efficiency as a function of the measured E_T in the HCAL for $K^{\pm}$ (top left) and $\pi^{\pm}$ (top right). The bottom row shows the $K^{\pm}$ L0 hadron trigger efficiency as a function of p_T (bottom left) and the number of hits in the SPD detector (bottom right).

multiplicity increases, the L0 hadron trigger efficiency also increases. This is not taken into account by the lookup tables provided by the CALO group, but is accounted for when using the PIDCalib method.

4.5 Systematic uncertainties

No matter the calibration dataset or the efficiency being measured (either PID or trigger efficiency), all measurements with PIDCalib share three common sources of systematic uncertainty. Firstly the calibration sample is limited in size and so its statistical uncertainty contributes to the uncertainty on the measured efficiency. The PDF used to model the mass distributions and determine the $_s$weights could be incorrect; this would result in incorrect $_s$weights and an incorrect efficiency map from the calibration sample. The effect of this was studied by the PID group and a 1 % systematic is applied. The binning could be too coarse and so incorrectly model the change in efficiency as a function of the three standard variables. As discussed in Section 4.3, when using equal statistics binning the number of bins can be altered by ±1 to determine the effect that this has on the overall efficiency; the largest variation is then taken as the systematic uncertainty.

4.6 PIDCalib modifications

It has been shown that the selected $B^+ \to J/\psi \left(e^+e^-\right)K^+$ sample can be used to measure the electron identification efficiency and that the standard PIDCalib binning takes into account the electron specific variables that affect the identification efficiency. The electron sample is now available for

collaboration wide use in PIDCalib. The groundwork has also been laid for TrigCalib, a new software package that measures trigger efficiencies.

In addition to the physics improvements of PIDCalib, there were several technical improvements which resulted in a factor of three speed increase and a reduction in memory usage by 50 %. The RooDataSet class of RooFit has two ways to store data: in standard C++ vectors (the default) or in a ROOT TTree object. By using incorrect constructors of the RooDataSet class, the dataset was unnecessarily converted back and forth between these data storage types; this was CPU intensive and resulted in unnecessary memory usage.

Many new user facing scripts have been developed for the PIDCalib package, the most interesting being the new equal-binning script. In addition there are new scripts to plot distributions in both the calibration and reference samples and to make efficiency plots.

CHAPTER 5

Lepton flavour violation in D^0 decays

In this chapter, the search for the LFV decay $D^0 \to e^\pm \mu^\mp$ is presented. Several other experiments have attempted to search for this decay, of which there has been no observation and so the previous experiments have placed limits on $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$. The Mark II experiment was the first to constrain $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ [103], but more recently constraints have been placed on $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ by E791 [104] and BaBar [105]. The most stringent limit has been published by Belle [36], which found $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) < 2.6 \times 10^{-7}$ at a 90 % CL. As previously discussed in Section 1.3, if a new limit were placed on $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right)$ below 10^{-7}, it would allow tighter constraints to be placed on the products of couplings in the RPV-MSSM [33, 42, 43].

This search for $D^0 \to e^\pm \mu^\mp$ uses the pp collision data collected by LHCb during 2011 and 2012, which includes $1\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$ and $2.0\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$. All D^0 decays in this analysis are required to come from D^* decays of the type $D^{*+} \to D^0 \pi^+$, as this significantly reduces the combinatorial

backgrounds of the selection [106]. The measurement of $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$ is normalised using the well-measured decay channel $D^0 \to K^-\pi^+$. To avoid any possible bias in the measurement, this analysis is performed blinded and so candidates with D^0 masses in the range $1815\,\text{MeV}\,\text{c}^{-2} < m(e\mu) < 1900\,\text{MeV}\,\text{c}^{-2}$ are not studied until the analysis procedure is finalised. The selection efficiency of $D^0 \to K^-\pi^+$ and $D^0 \to e^{\pm}\mu^{\mp}$ candidates is determined using a mixture of simulation and data-driven techniques. A boosted decision tree (BDT) is used to discriminate signal $D^0 \to e^{\pm}\mu^{\mp}$ decays from the background. The efficiencies and number of $D^0 \to K^-\pi^+$ candidates are input to a fitting program which fits the mass distributions of the $D^0 \to e^{\pm}\mu^{\mp}$ candidates in three categories based on the BDT output and this is used to determine $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$.

5.1 Data selection and reconstruction

Candidate $D^0 \to e^{\pm}\mu^{\mp}$ events are triggered using the L0 muon trigger and the muon track trigger in HLT1. In the second stage of the HLT, the triggered muon is combined with an electron candidate under a $D^0 \to \pi^-\pi^+$ mass hypothesis. The D^0 candidates that pass the trigger are required to have a vertex formed from two good quality tracks with opposite charge, that is well separated from any PV, with the summed momentum vector of the two tracks pointing towards a PV. During the offline event selection, energy deposits in the ECAL are searched for that could be from a bremsstrahlung photon; if they are likely to have originated from the electron candidate, the energy deposit is used to correct the measured electron momentum [107]. The two D^0 decay tracks are also required to have momentum $p > 4\,\text{GeV}\,\text{c}^{-1}$

Table 5.1: Requirements made to select $D^0 \to e^{\pm}\mu^{\mp}$ and $D^0 \to K^-\pi^+$ candidates from a $D^{*+} \to D^0\pi^+$ decay.

Particles	Requirement
D^0 daughters	$P > 4\,\text{GeV}\,\text{c}^{-1}$
	$p_T > 0.75\,\text{GeV}\,\text{c}^{-1}$
	Track fit χ^2 per degree of freedom < 5
	Minimum IP χ^2 from any PV > 3
	Ghost probability from ProbNN < 0.5
e	$\Delta\text{LL}(e-\pi) > 6$
μ	Positively identified by muon stations
	No shared hits with other reconstructed particles
	Muon probability from ProbNN > 0.4
K	$\Delta\text{LL}(K-\pi) > 6$
π	$\Delta\text{LL}(K-\pi) < 0$
D^0	$p_T > 1800\,\text{MeV}\,\text{c}^{-1}$
	DOCA between daughters $<0.1\,\text{mm}$
	$m(D^0) <0.3\,\text{GeV}\,\text{c}^{-2}$ from nominal
	Maximum child $p_T >1.1\,\text{GeV}$
	Cosine of momentum w.r.t. flight distance > 0.9997
	Minimum IP χ^2 from any PV of the daughters > 8
	χ^2 distance from related PV > 20
	Minimum IP χ^2 from any PV < 15
	Vertex fit χ^2 per degree of freedom < 10
π from D^*	$p_T >0.11\,\text{GeV}\,\text{c}^{-1}$
	Minimum IP χ^2 from any PV < 10
	Track fit χ^2 per degree of freedom < 7
	Ghost probability from ProbNN < 0.05
D^*	Mass difference from nominal D^0 $<0.3\,\text{GeV}\,\text{c}^{-2}$
	$m(D^*) - m(D^0)$ in the range $135.42\,\text{MeV}\,\text{c}^{-2}$ to $155.42\,\text{MeV}\,\text{c}^{-2}$
	Vertex fit χ^2 per degree of freedom < 10
	Decay tree fitter $\chi^2 < 100$

and transverse momentum $p_T > 0.75\,\text{GeV}\,\text{c}^{-1}$. The D^0 candidate is combined with a π that has transverse momentum $p_T > 110\,\text{MeV}\,\text{c}^{-1}$ to create a $D^{*+} \to D^0\pi^+$ candidate that originates from a PV. A geometric fit of the decay tree is performed [108], which constrains the two D^0 decay tracks to originate from a common secondary vertex, while the π and D^0 candidates are also constrained to come from their common associated PV. The decay tree fit (DTF) improves the resolution on the mass difference between the D^* and D^0 candidates, which is then required to be in a $20\,\text{MeV}\,\text{c}^{-2}$ window covering the range $135\,\text{MeV}\,\text{c}^{-2}$ to $155\,\text{MeV}\,\text{c}^{-2}$. In a very small number of events, there are multiple $D^{*+} \to D^0\pi^+$ candidates sharing reconstructed tracks and in these events all but one candidate are randomly discarded. The full list of the requirements made to select the dataset is given in Table 5.1.

A similar selection process is followed for $D^0 \to K^-\pi^+$ candidates, except that either the π or K are allowed to activate the trigger and that both the L0 hadron trigger and the track trigger in HLT1 must be activated by the same particle. The PID requirements applied to the $D^0 \to e^\pm\mu^\mp$ candidates were optimised using pseudo-experiments to maximise the sensitivity to $\mathcal{B}\left(D^0 \to e^\pm\mu^\mp\right)$, whereas the PID requirements applied to the $D^0 \to K^-\pi^+$ candidates were chosen to reduce the amount of background. After the above selections, 2114 candidates remain in the signal $m(e\mu)$ region ($1815\,\text{MeV}\,\text{c}^{-2}$ to $1915\,\text{MeV}\,\text{c}^{-2}$) for $D^0 \to e^\pm\mu^\mp$ and 330 359 candidates remain in the $D^0 \to K^-\pi^+$ dataset.

Both the data and simulated events are selected using the same process. In this analysis we use large samples of simulated $D^0 \to e^\pm\mu^\mp$ decays, $D^0 \to K^-\pi^+$ decays and various other background decays.

5.2 Backgrounds

The dominant background is combinatorial in nature; it is formed from randomly combined reconstructed particles. This background varies smoothly as a function of $m(e\mu)$, which makes it easy to model. The most dangerous background is $D^0 \to \pi^-\pi^+$ that has been misidentified as $D^0 \to e^\pm\mu^\mp$. This background is particularly important as the $m(e\mu)$ distribution peaks just $20\,\text{MeV}\,\text{c}^{-2}$ below that of the signal $D^0 \to e^\pm\mu^\mp$ mass distribution. The $m(e\mu)$ distributions from the simulations of both $D^0 \to e^\pm\mu^\mp$ decays and $D^0 \to \pi^-\pi^+$ decays reconstructed as $D^0 \to e^\pm\mu^\mp$ are shown in Fig. 5.1. The large low-mass tail of the $D^0 \to e^\pm\mu^\mp$ decays is caused by the electron losing energy through bremsstrahlung, while the high-mass tails in both $m(e\mu)$

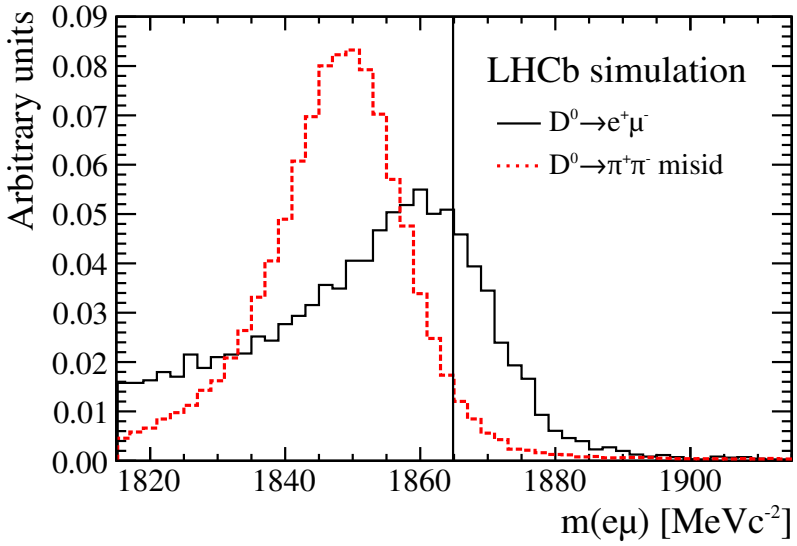

Figure 5.1: Mass spectra from simulations of $D^0 \to e^\pm\mu^\mp$ (solid black line) and $D^0 \to \pi^-\pi^+$ reconstructed as $D^0 \to e^\pm\mu^\mp$ (dashed red line). The vertical black line indicates the known mass of the D^0 meson [1]. Both mass spectra are normalised to an area of one.

distributions are caused by the incorrect assignment of a bremsstrahlung photon to the electron candidate.

Other misidentified backgrounds considered include both semileptonic D^0 decays ($D^0 \to \pi^\pm e^\mp \nu_e$, $D^0 \to K^\pm e^\mp \nu_e$, $D^0 \to \pi^\pm \mu^\mp \nu_\mu$ and $D^0 \to K^\pm \mu^\mp \nu_\mu$) and hadronic decays ($D^0 \to K^- K^+$ and $D^0 \to K^- \pi^+$). There is a large difference between the $K^\pm$ and $\mu^\pm$ masses and a larger difference between the $K^\pm$ and $e^\pm$ masses; this pushes all decays involving $K^\pm$ below the signal $m(e\mu)$ region. Similarly for decays involving a ν, the reconstructed D^0 mass is often far below the signal mass region. During reconstruction, ECAL energy deposits are added to the momentum of the electron candidate and so an energy deposit incorrectly assigned to one of these misidentified backgrounds could alter the D^0 candidate mass to place it in the signal region. In this case, the background varies smoothly in the $m(e\mu)$ signal region; given the branching fractions, the low misidentification rates, the low fraction of events that will be in the $m(e\mu)$ signal region and the mass distribution of these backgrounds in the $m(e\mu)$ signal region, these are not important backgrounds and so are considered part of the combinatorial background for this analysis.

5.3 Selection efficiencies

The selection efficiency of $D^0 \to e^\pm \mu^\mp$, $D^0 \to K^- \pi^+$ and misidentified $D^0 \to \pi^- \pi^+$ are measured in this section using a mixture of simulation and calibration samples. However, there is one problem with all of the simulated datasets: the track multiplicity of simulated events is consistently lower in the simulation than in the data. Because the detectors have higher occu-

pancies in higher multiplicity events, they generally perform less well; this significantly affects how well particles can be identified and thus alters their selection efficiencies.

The L0 trigger efficiency, the HLT1 trigger efficiency and the PID efficiencies are measured in data using the PIDCalib package, while the remaining selection efficiencies are measured using simulated events. When measuring efficiencies using simulated events, the efficiencies are measured in three bins of the number of hits in the SPD detector: $N_{\text{SPD}} \leq 150$, $150 < N_{\text{SPD}} \leq 300$ and $N_{\text{SPD}} > 300$. The efficiencies in each bin are then reweighted by the data distribution of hits in the SPD detector to find the selection efficiency. This reweights the simulation to more closely match the observed track multiplicity in data.

To measure the trigger and PID efficiencies in data, the method outlined in Chapter 4 is followed to create the selection efficiency maps. Then simulated events are used to reweight the efficiency maps and find the selection efficiencies for the datasets in this analysis. However, the problem with the simulation arises again; the data and simulation track multiplicities do not match. A different approach is taken than for the efficiencies measured purely in simulation. When finding the efficiency of a certain event in the reference sample, instead of directly using the number of tracks found in the simulation, the number of tracks is multiplied by a scale factor of 1.334; this simply chooses a different bin in the efficiency map with a larger multiplicity and with an efficiency that is, on average, lower. The scale factor was found by comparing the scaled track multiplicity distribution from simulation and the data track multiplicity distribution; the scale factor which produced the

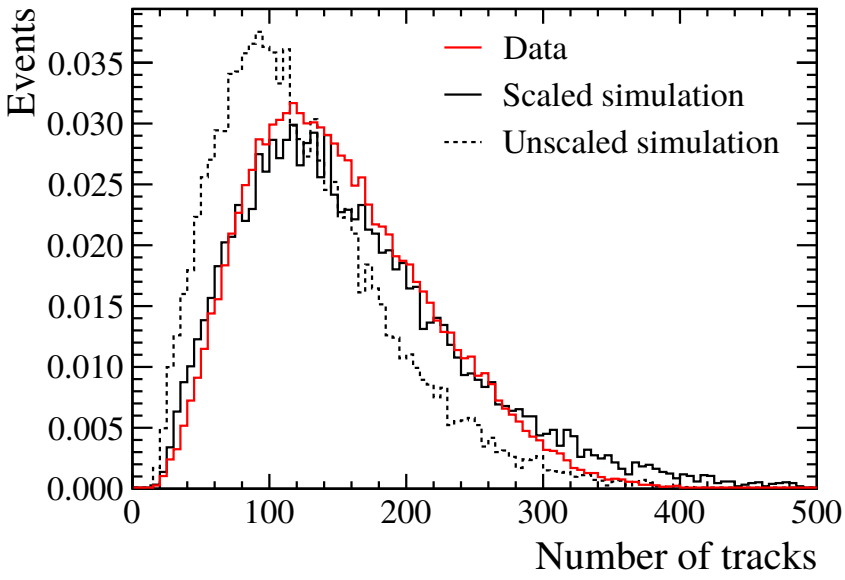

Figure 5.2: Track multiplicity distribution from 2011 simulation (black lines) and the 2011 dataset (red line). The dashed black line is the original distribution from the simulation, while the solid black line shows the simulated distribution scaled by 1.334.

lowest χ^2 between the two distributions was then selected. The distributions of track multiplicities are shown in Fig. 5.2.

Using this mixture of calibration samples and simulated events the probability to select a $D^0 \to e^\pm\mu^\mp$ decay is $(4.3 \pm 0.2) \times 10^{-4}$ in the 7 TeV data and $(4.5 \pm 0.3) \times 10^{-4}$ in the 8 TeV data. The selection efficiency of $D^0 \to K^-\pi^+$ is much lower as only 1 in 100 candidates are actually saved; this makes the number of these events more manageable by the LHCb computing infrastructure. The selection efficiency of $D^0 \to K^-\pi^+$ decays is $(2.75 \pm 0.09) \times 10^{-6}$ in the 7 TeV data and $(2.44 \pm 0.09) \times 10^{-6}$ in the 8 TeV data. The probability that a $D^0 \to \pi^-\pi^+$ decay is selected in the final $D^0 \to e^\pm\mu^\mp$ dataset is $(1.0 \pm 0.6) \times 10^{-8}$ in the 7 TeV data and $(1.8 \pm 0.4) \times 10^{-8}$ in the 8 TeV data. All of these selection efficiencies include both statistical and systematic

uncertainties, which are discussed in Section 5.6. The tables in Appendix B show the individual components of these selection efficiencies.

5.4 Multivariate classifier

A BDT that utilises gradient boosting [109, 110] is applied to the selected $D^0 \to e^{\pm}\mu^{\mp}$ candidates in order to classify those events which are more signal or background like. The BDT was designed to separate genuine $D^{*+} \to D^0\pi^+$ decays from combinatorial backgrounds and so none of the input variables contain PID information. Because of this, it is expected to perform equally well for any fully reconstructed $D^{*+} \to D^0\pi^+$ decay. The variables that are given as an input to the BDT are: the smallest distance of closest approach of the D^0 candidate to any PV; the sum of the transverse momenta of all other particles in a cone of radius 1 rad around the candidate D^* momentum direction as a fraction of the D^* transverse momentum (referred to as the isolation variable); the χ^2 value of the geometric fit; and the impact parameter χ^2 (IPχ^2) values with respect to the associated PV for each of the D^* and D^0, and for the two D^0 decay tracks. The IPχ^2 is defined as the difference in vertex fit χ^2 with and without the particle considered.

When the center-of-mass energy was increased from 7 TeV to 8 TeV the distributions of several variables were slightly altered and so the BDT is trained separately for the 7 TeV and 8 TeV datasets. An example of one distribution which was altered was the isolation variable; because there was higher track multiplicity in the 8 TeV data, there were on average more tracks with greater energies in the 1 rad cone, which changed the distribution of the isolation variable. The BDT was given a signal training sample that consisted

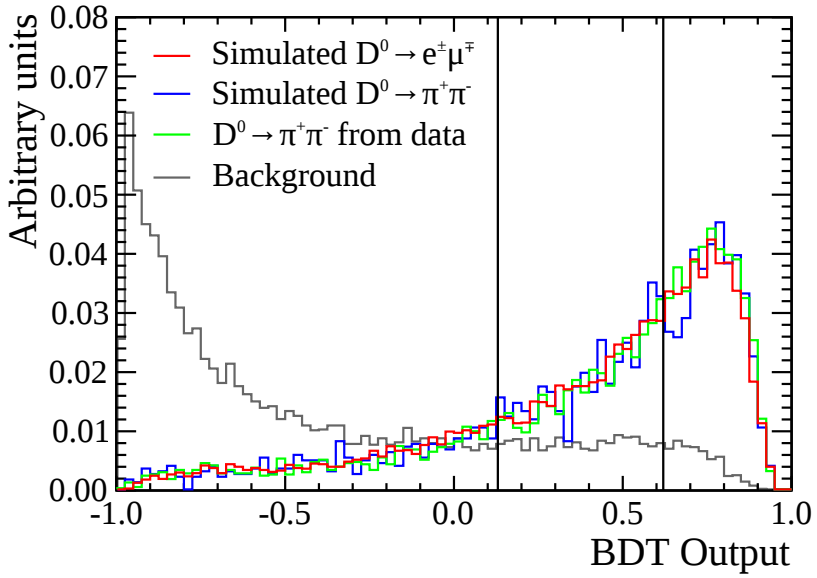

Figure 5.3: Comparison of BDT output for four 7 TeV datasets: background data (light grey), simulated $D^0 \to e^{\pm}\mu^{\mp}$ (red), simulated $D^0 \to \pi^-\pi^+$ (blue) and well-identified $D^0 \to \pi^-\pi^+$ decays from data (green). The vertical black lines indicate the boundaries that separate the BDT output into the three subsamples. All distributions are normalised to an area of one.

of simulated $D^0 \to e^{\pm}\mu^{\mp}$ events. The background training sample consists of $D^0 \to e^{\pm}\mu^{\mp}$ candidates from data within 300 MeV c^{-2} of the known D^0 mass, but excluding those in the signal region (1815 MeV c^{-2} to 1915 MeV c^{-2}). Half of the background and signal samples were used for training and the other half were used to ensure that the BDT was not over-trained.

The BDT returns a value between zero (most background-like) and one (most signal-like), as shown in Fig. 5.3, and this is used to separate the data into three subsamples: $-1.00 - 0.13$, $0.13 - 0.62$ and $0.62 - 1.00$. These ranges were chosen by a method which optimises the binning of discriminating variables used in the computation of confidence levels with the modified

frequentist approach [111]. The figure of merit for this optimisation is

$$\Delta LQ = 2\ln\left(\prod \frac{P(s_i+b_i, s_i+b_i)}{P(s_i+b_i, b_i)}\right) - 2\ln\left(\prod \frac{P(b_i, s_i+b_i)}{P(b_i, b_i)}\right), \tag{5.1}$$

where $P(k_i, \lambda_i)$ is the Poisson probability to observe k_i events in bin i when λ_i events are expected.

5.5 Fits to mass spectra

In this section several unbinned multidimensional log-likelihood maximization fits are performed using programs created with the RooFit framework [99, 100]. The data are fitted as a function of the D^0 candidate mass and the difference between the D^0 and D^* candidate masses, δm. When fitting $D^0 \to e^\pm\mu^\mp$ decay candidates, the fit is performed simultaneously to the three subsamples of BDT output.

First two fits are performed to the simulated $D^0 \to e^\pm\mu^\mp$ dataset and the simulated $D^0 \to \pi^-\pi^+$ dataset when misidentified as $D^0 \to e^\pm\mu^\mp$. The PDFs used to model the mass distributions in these two fits are the same, but with different values of the fitted parameters. The D^0 candidate mass distribution is modelled using two CB functions [101] with a common mean; one has a low-mass tail to fit the bremsstrahlung energy loss and the other has a high-mass tail to fit for candidates with an incorrectly assigned bremsstrahlung photon. Most parameters are allowed to vary independently across the three subsamples of BDT output, but the α parameters of the CB functions are constrained to be equal for all three BDT regions and all n parameters are fixed to 2. The sum of three Gaussian functions is used to model the δm distributions; of these three Gaussian functions, the two dominant Gaussian

functions share a common mean, while the third has an independent mean and accounts for less than 5 % of events. As previously stated, the track multiplicity of the simulated events is too low when compared to the data, which alters the bremsstrahlung photon reconstruction and in turn affects the D^0 candidate mass distribution. Because of this, the PDF parameters are determined in three bins of the number of hits in the SPD detector and then reweighted to the data distribution.

When determining the signal $D^0 \to e^\pm\mu^\mp$ PDFs, the full selection is applied to the simulation, but, if this were done for the simulated $D^0 \to \pi^-\pi^+$ events, no events would pass the full selection due to the low probability to misidentify a $D^0 \to \pi^-\pi^+$ decay $\mathcal{O}(10^{-8})$ and so only a subset of the selection requirements are applied. The efficiency as a function of $m(e\mu)$ was studied for the selection requirements which were not applied and it was found that these selection requirements had a small linear dependence on the reconstructed D^0 candidate mass. This dependence is fitted and applied as a correction to the $m(e\mu)$ PDF in the fit to the data.

The next fit is used to determine the number of $D^0 \to K^-\pi^+$ decays in the normalisation channel. There are many $D^0 \to K^-\pi^+$ candidates and so binned fits are performed to this data separately for the 7 TeV and 8 TeV samples. Two Gaussian functions with a common mean are used to model the $m(K^-\pi^+)$ distribution. To model the δm distribution the same function is used as in the $D^0 \to e^\pm\mu^\mp$ δm PDF; this consists of the sum of three Gaussian functions with two sharing a common mean. There are two types of background in this dataset: the purely combinatorial background (peaking in neither $m(K^-\pi^+)$ nor δm), and D^0 decays that originate from the PV and are combined with a random pion (peaking in $m(K^-\pi^+)$ and not in δm). The

purely combinatorial background is modelled using a Chebyshev polynomial to fit the $m(K^-\pi^+)$ distribution and an empirical function is used to fit the δm distribution. This has the form

$$f(\delta m) = N\left[\left(1-e^{-(\delta m - m_0)/c}\right)\left(\frac{\delta m}{m_0}\right)^a + b\left(\frac{\delta m}{m_0}-1\right)\right], \tag{5.2}$$

where N is a normalisation factor, m_0 is the threshold mass, and a, b and c are parameters. The prompt D^0 background is modelled using the signal $m(K^-\pi^+)$ PDF (consisting of two Gaussian functions) and the empirical formula (Eq. (5.2)) is again used to model the δm distribution. Figure 5.4 shows the fitted $m(K^-\pi^+)$ and δm distributions from the $D^0 \to K^-\pi^+$ 8 TeV dataset. There is a total of 80 000 $D^0 \to K^-\pi^+$ events observed in the 7 TeV dataset and 182 000 $D^0 \to K^-\pi^+$ events are observed in the 8 TeV dataset.

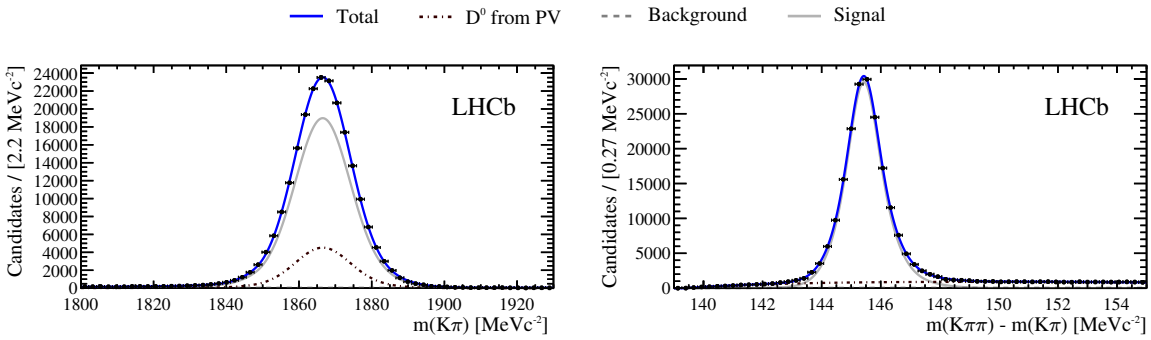

Figure 5.4: Fitted distributions of the normalisation channel $D^0 \to K^-\pi^+$: D^0 candidate mass, $m(K^-\pi^+)$, (left) and the difference between the D^* and D^0 candidate masses, $\delta m = m(K^-\pi^+\pi^-) - m(K^-\pi^+)$, (right) for the 8 TeV data. The dark (blue) line shows the overall fit, the lighter grey line shows the signal, the dotted line shows genuine D^0 events where the soft pion does not come from a D^{*+} decay, and the dashed line shows the combinatorial background.

The last fit in this section is used to determine both the number of $D^0 \to e^{\pm}\mu^{\mp}$ candidates and the branching fraction of $D^0 \to e^{\pm}\mu^{\mp}$. The reweighted results of the fits to the simulated $D^0 \to e^{\pm}\mu^{\mp}$ decays and the simulated misidentified $D^0 \to \pi^-\pi^+$ decays are input to this fit, along with the number of fitted $D^0 \to K^-\pi^+$ decays. In addition to the fit results, several other parameters are input: the selection efficiencies of $D^0 \to e^{\pm}\mu^{\mp}$ and $D^0 \to K^-\pi^+$, the misidentification rate of $D^0 \to \pi^-\pi^+$, and the branching fractions from Ref. [1] of $D^0 \to K^-\pi^+$ and $D^0 \to \pi^-\pi^+$. Gaussian constraints are applied to all of the input parameters such that they can vary within their uncertainties; both the statistical and systematic uncertainties (discussed in Section 5.6) are included. Using the input information, the expected number of misidentified $D^0 \to \pi^-\pi^+$ can be calculated and is used to constrain the number of $D^0 \to \pi^-\pi^+$ in the fit.

The combinatorial backgrounds are modelled using a second-order Chebyshev polynomial in the $m(e\mu)$ distribution, while in the δm distribution the empirical formula (Eq. (5.2)) is used again. Due to the small number of background events in this sample, the background PDF is chosen to be the same in each BDT region. The only parameters without constraints applied are the number of combinatorial background events in each region of BDT output, the shape of the background and the branching fraction of $D^0 \to e^{\pm}\mu^{\mp}$.

The final fit results are shown in Fig. 5.5 for the combined 7 TeV and 8 TeV datasets and separately for the three subsamples of BDT output. The peaks seen in the fitted PDF are due to misidentified $D^0 \to \pi^-\pi^+$ events; there is no evidence seen for $D^0 \to e^{\pm}\mu^{\mp}$ decays. The number of signal $D^0 \to e^{\pm}\mu^{\mp}$ decays determined from the fit was -7 ± 15.

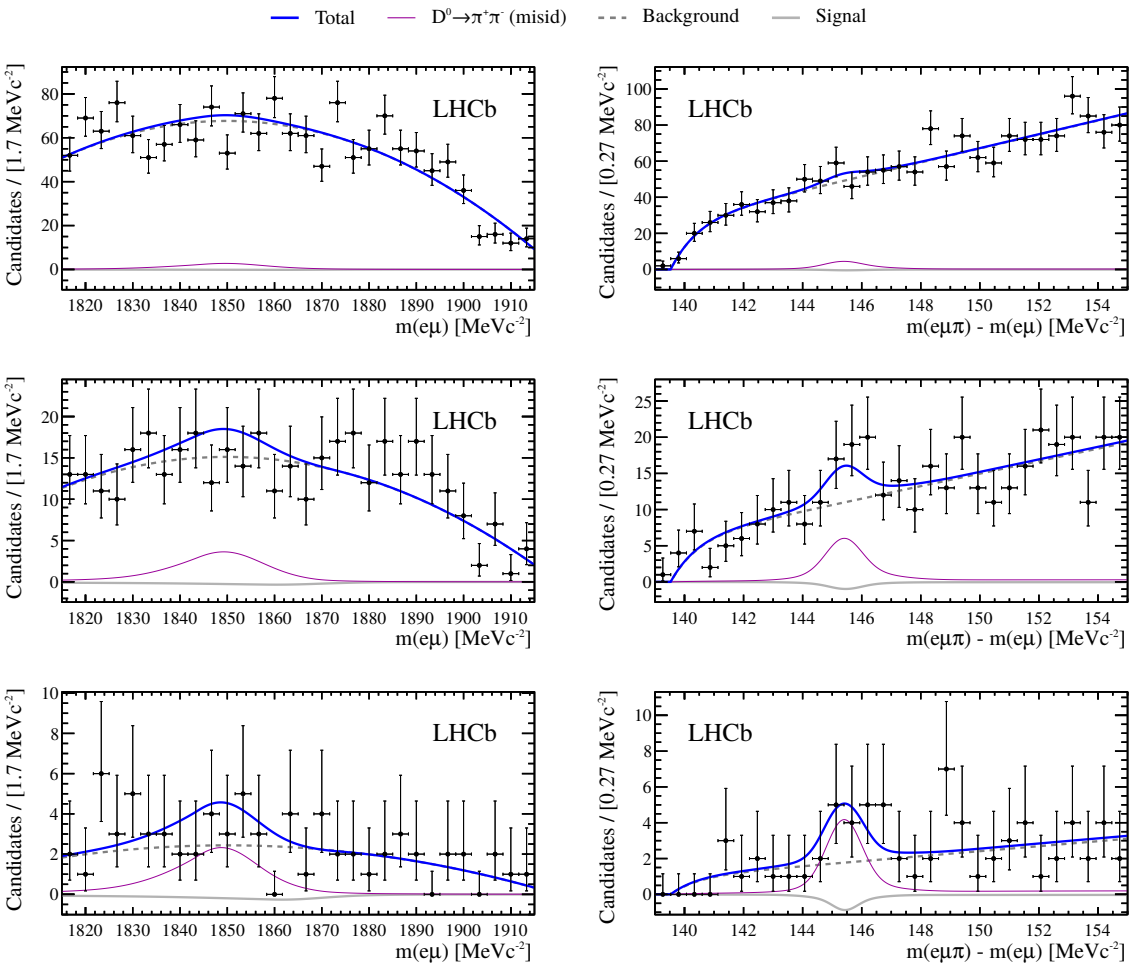

Figure 5.5: Fitted distributions of the D^0 candidate mass, $m(e\mu)$, (left) and the difference between the D^* and D^0 candidate masses, $\delta m = m(e\mu\pi) - m(e\mu)$, for the combined 7 TeV and 8 TeV data. The three rows are for the three subsamples of BDT output, with the top row for the most background-like and the bottom row for most signal-like. The solid (blue) lines show the total fit results, while the thick (grey) lines show the signal, the thin (purple) lines show the misidentified $D^0 \to \pi^-\pi^+$ and the dashed (grey) lines indicate the combinatorial background.

5.6 Systematic uncertainties

The uncertainty on the fitted branching fraction, $\mathcal{B}\left(D^0 \to e^\pm\mu^\mp\right)$, is domi-nated by statistical errors. However, many sources of systematic uncertain-

ties are considered, including from the selection efficiencies, the $D^0 \to e^\pm\mu^\mp$ and misidentified $D^0 \to \pi^-\pi^+$ PDFs, the number of normalisation candidates and the BDT performance. However, the largest of the systematic uncertainties comes from the efficiency measurements.

Of the efficiencies determined using only simulated events, the largest systematic uncertainty is caused by the difference in track multiplicities between simulation and data. To estimate the uncertainty of the reweighting procedure, half of the difference between the weighted and unweighted efficiencies is taken as the systematic uncertainty. This is typically 4.2 % and 1.8 % for signal $D^0 \to e^\pm\mu^\mp$ events and normalisation $D^0 \to K^-\pi^+$ events. The LHCb tracking group recommend a 0.4 % systematic uncertainty per track due to the difference between the track reconstruction in data and simulation; since the material in the detector is only known to 10 %, a further 1.1 %, 1.4 % and 2.0 % is assigned per $K^\pm$, $\pi^\pm$ and $e^\pm$ due to interaction uncertainties. In total around 2.5 % is applied to each decay mode for tracking uncertainties. The D^0 candidate mass is reconstructed more accurately in the simulated trigger than in actual triggered data; the simulation is reweighted to the data distribution and a change in selection efficiency of 1.0 % is applied as a systematic uncertainty for the 7 TeV data and 0.5 % for the 8 TeV data. The different components of the systematic uncertainties are tabulated and compared to the statistical errors in Appendix B.

For the efficiencies determined using the PIDCalib software package, different systematic uncertainties are considered. As discussed in Section 4.3, the calibration dataset is binned such that each bin has a roughly equal number of events. The number of bins in each dimension was generally chosen to be 7; in order to find the systematic uncertainty associated to

the choice of binning scheme, the number of bins was altered by ±1 and the largest change in the selection efficiency was treated as the systematic uncertainty. The scale factor used to scale the number of simulated tracks was also altered by ±5 % and the change in efficiency taken as the systematic uncertainty. The last systematic uncertainty considered was from the ${}_s$weighting procedure; as suggested by the PID group a 0.1 % uncertainty is applied for this.

The total systematic uncertainties are 6 % on the $D^0 \to e^\pm\mu^\mp$ selection efficiency, 30 % on the $D^0 \to \pi^-\pi^+$ misidentification probability and 3 % on the $D^0 \to K^-\pi^+$ selection efficiency. Tables B.1 to B.6 in Appendix B show the individual parts of the selection efficiency and break down their systematic uncertainties into the constituent parts.

For the signal $D^0 \to e^\pm\mu^\mp$ PDF it is possible that there could be a difference in the bremsstrahlung reconstruction between the data and simulated events. A sample of $B^+ \to J/\psi\left(e^+e^-\right)K^+$ is taken from simulation and data; it is fitted with a PDF similar to that used to model the $D^0 \to e^\pm\mu^\mp$ signal. The fractional differences between the fit results to the simulated events (after reweighting for the track multiplicity) and the data is taken as an additional systematic uncertainty. These uncertainties are included in the Gaussian constraints placed on the PDF parameters in the fit to the data.

When determining the misidentified $D^0 \to \pi^-\pi^+$ PDF, the selection requirements not applied to the simulated dataset show a dependence on $m(e\mu)$. The efficiency changes by 9.3 % from the centre of the signal region to the edge; half of this change (4.7 %) is applied as a systematic uncertainty.

The binned fit that determines the number of $D^0 \to K^-\pi^+$ normalisation events has the number of bins altered by $-20\,\%$ and $+50\,\%$. Further, the order of the Chebyshev polynomial is increased by one, any fixed parameters are varied and the Gaussian mean values in the δm fits are constrained to be equal. From these studies a systematic uncertainty of $1\,\%$ is assigned to the number of $D^0 \to K^-\pi^+$ normalisation events.

Since the BDT used in this analysis contains no PID information and primarily separates real $D^{*+} \to D^0\pi^+$ decays from combinatorial backgrounds, both simulated $D^0 \to e^\pm\mu^\mp$ decays and well identified $D^0 \to \pi^-\pi^+$ decays selected in data are analysed by the BDT. The fraction of these datasets in each BDT bin is compared and small variations are seen (as demonstrated in Fig. 5.3). The difference in the fraction of events in each bin between the two samples is taken as a systematic uncertainty; this is typically $2.5\,\%$.

5.7 Results

The measured branching fraction of $D^0 \to e^\pm\mu^\mp$ is given by

$$\mathcal{B}\left(D^0 \to e^\pm\mu^\mp\right) = \frac{N_{e\mu}/\epsilon_{e\mu}}{N_{K\pi}/\epsilon_{K\pi}} \times \mathcal{B}\left(D^0 \to K^-\pi^+\right) \tag{5.3}$$

where $N_{e\mu}$ and $N_{K\pi}$ are the fitted numbers of $D^0 \to e^\pm\mu^\mp$ and $D^0 \to K^-\pi^+$ events, the corresponding ϵ parameters are the overall efficiencies and $\mathcal{B}\left(D^0 \to K^-\pi^+\right)$ is the branching fraction for the normalisation channel, $\mathcal{B}\left(D^0 \to K^-\pi^+\right) = 3.88 \pm 0.05\,\%$, taken from Ref. [1].

No evidence is seen of $D^0 \to e^\pm\mu^\mp$ decays in any bin of BDT output or in the overall mass spectra. The fit gave a total of -7 ± 15 signal events, which results in a measured branching fraction of $\mathcal{B}\left(D^0 \to e^\pm\mu^\mp\right) =$

$(-0.6 \pm 1.2) \times 10^{-8}$. An upper limit is placed on the branching fraction as the data are consistent with the background-only hypothesis. The calculation of the upper limit assumes Wilks's Theorem [112] and uses the CL_s method [113], where the p-value for the signal-plus-background hypothesis is compared to that for the background-only hypothesis. The expected and observed CL_s values are shown in Fig. 5.6 as a function of the assumed branching fraction, where the expected CL_s values are obtained using an Asimov dataset [114] as implemented in the RooStats software framework [115]. There is a good agreement between the expected and observed CL_s values and calculating an expected limit using pseudo-experiments gives similar results. An upper limit is set on the

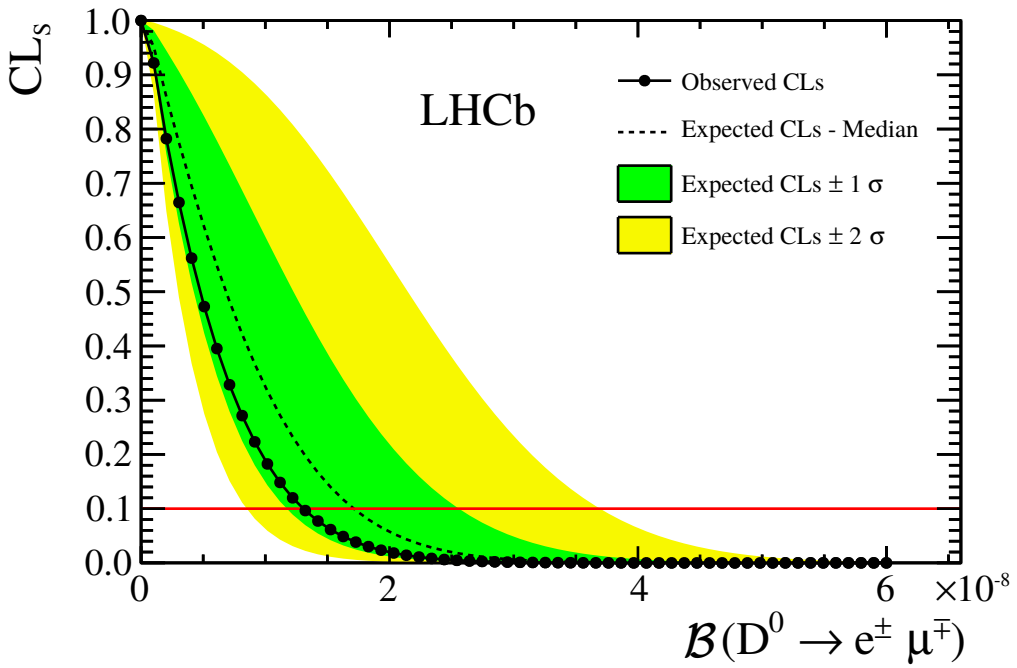

Figure 5.6: Expected and observed CL_s values. The expected values are shown by the dashed line, with $\pm 1\sigma$ and $\pm 2\sigma$ regions shaded. The observed values are shown by the solid line connecting the data points. The horizontal (red) line indicates the 90 % CL.

branching fraction, $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right) < 1.3 \times 10^{-8}$ at a 90 % CL [116]; this is an order of magnitude lower than the previous lowest limit. Using this limit, stringent constraints can be placed on the product of couplings in the RPV-MSSM [33, 42, 43].

CHAPTER 6

B_s^0 and B_d^0 mixing frequencies

The first observation of B_d^0 mixing was made by the ALEPH collaboration [117] in 1993; since then many measurements have been made of the B_d^0 mixing frequency (Δm_d) and these are averaged to determine $\Delta m_d = 0.510 \pm 0.003\,\text{ps}^{-1}$ [1]. However, the first observation of B_s^0 mixing was not until 2006, when the CDF collaboration published a measurement where the B_s^0 mixing frequency (Δm_s) was determined to be $\Delta m_s = 17.77 \pm 0.10(\text{stat.}) \pm 0.07(\text{syst.})\,\text{ps}^{-1}$ [118]. The delay between the observation of B_d^0 and B_s^0 mixing is due to the fast oscillations of the B_s^0 mixing, requiring a high decay-time resolution in order to observe the oscillations. Since the measurement from CDF, the LHCb experiment has made a measurement of $\Delta m_s = 17.768 \pm 0.023(\text{stat.}) \pm 0.006(\text{syst.})\,\text{ps}^{-1}$ [119] using fully-reconstructed hadronic B_s^0 decays.

The analysis that follows measures both the B_d^0 and B_s^0 mixing frequencies. This is done by selecting semileptonic B_d^0 and B_s^0 decays of the type

$B_q^0 \to D_q^- \mu^+ \nu X$, where $D_q^- \to K^- K^+ \pi^-$. Since this decay involves a neutrino, which cannot be detected by LHCb, the momentum of the B_q^0 is not fully reconstructed and so the missing momentum is corrected for using the k-factor method. The lifetime and invariant mass distributions are then fitted and Δm_d and Δm_s are determined. A Fourier analysis of the decay-time distribution is also presented as a cross-check and yields results consistent with the multivariate fitting method. This analysis demonstrates that the LHCb experiment can make time-dependent measurements using semileptonic B_q^0 decays from pp-collisions without a hermetic detector. This is particularly interesting as time-dependent semileptonic measurements are needed to enable measurements of the $\mathcal{CP}$-violation in B_q^0 mixing (for example the semileptonic asymmetries a_{sl}^d and a_{sl}^s [9, 11, 70]).

6.1 Data selection and reconstruction

The LHCb data used in this analysis corresponds to an integrated luminosity of $1.0\,\text{fb}^{-1}$ and was collected in pp-collisions at $\sqrt{s} = 7\,\text{TeV}$ during the 2011 physics run of the LHC.

Events are selected when a $D_q^- \to K^- K^+ \pi^-$ candidate and a muon candidate form a vertex. A cut-based selection is then applied to the data to enhance the fraction of D_q^- mesons which come from semileptonic B_q^0 decays. The vertex reconstruction quality, track reconstruction quality, momenta, invariant masses, flight distances and PID variables are used in the selection. The selection was initially optimised on simulated data to maximise the signal significance, $S/\sqrt{(S+B)}$, where S represents the number of signal candidates and B represents the number of background

Table 6.1: Selection requirements used to select $B_q^0 \to D_q^- \mu^+ \nu$ candidates.

Particle(s)	Variable	Description		Cut
$(\pi^+, K^\pm)$	p	Momentum	>	(2,2) [GeV c^{-1}]
$(\pi^+, K^\pm, \mu^+)$	p_T	Transverse momentum	>	(0.3, 0.4, 1.0) [GeV c^{-1}]
$(\pi^+, K^\pm, \mu^+)$	χ^2_{IP}	χ^2 inconsistency with primary vertex	>	(7, 4, 2)
$(\pi^+, K^\pm, \mu^+)$	Δ_{IP}	Distance inconsistency with primary vertex	>	(0.04,0.01,0.01) [mm]
$(\pi^+, K^\pm, \mu^+)$	$\chi^2_{tr}/N_{\text{dof}}$	Track fit quality	<	(10, 10, 5)
π^+	$\Delta_{LL}(\pi-\mu)$	Particle ID w.r.t. muon	>	-5
π^+	$\Delta_{LL}(\pi-K)$	Particle ID w.r.t. kaon	>	-10
$K^\pm$	$\Delta_{LL}(K-\pi)$	Particle ID w.r.t. pion	>	(0, 13)[a]
$K^\pm$	$\Delta_{LL}(K-\mu)$	Particle ID w.r.t. muon	>	5
$K^\pm$	$\Delta_{LL}(K-p)$	Particle ID w.r.t. proton	>	-10
μ^+	$\Delta_{LL}(\mu-\pi)$	Particle ID w.r.t. pion	>	1
μ^+	$\Delta_{LL}(\mu-K)$	Particle ID w.r.t. kaon	>	-5
μ^+	IsMuon	Standard muon ID, hits in three stations		true
(D_q^-, B_q^0)	DOCA	Distance of closest approach of daughters	<	(0.3, 0.3) [mm]
(D_q^-, B_q^0)	P_{vtx}	Vertex probability	>	(0.01, 0.015)
(D_q^-, B_q^0)	χ^2_{pv}	Vertex separation χ^2	>	(144, 1)
(D_q^-, B_q^0)	$\cos\theta_{pv}$	Cosine of momentum w.r.t. flight vector	>	(0.99, 0.998)
D_q^-	$\sum_{PT}$	Scalar sum of daughter P_T	>	1.2 [GeV c^{-1}]
D_q^-	P_T	Transverse momentum	>	1.5 [GeV c^{-1}]
D_q^-	$\Delta_{pv,z}$	Vertex separation distance in z-direction	>	1 [mm]
D_q^-	Δ_z	Flight distance in z-direction	>	0 [mm]
K^+K^-	M	D^0-veto in mass of KK-pair	<	1.84 [GeV c^{-2}]
D_q^-	$M-M(KK)-M(\pi)$	Threshold-veto, removes slow pions	>	15 [MeV c^{-2}]
D_q^-	$M-M_{PDG}$	D_s mass window		-0.2, 0.1 [GeV c^{-2}]
B_q^0	n	Normalised mass window		0.24, 1.0

[a] A hard cut is applied to the Kaon which shares a charge with the Pion, to suppress the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ background.

candidates; however, the PID requirements were subsequently tightened to reduce the number of misidentified particles and the requirements on the vertex reconstruction qualities were reduced as simulations predicted better performance of the vertexing than could be achieved in the real data. Finally the track quality variables were set to those recommended by the LHCb tracking group. The final list of all selection requirements is shown in Table 6.1.

In the following section various requirements on the invariant masses are discussed; m_0 represents the known mass of a particle, measured from previous experiments and averaged by Ref. [1], while m represents the mass of a particular candidate or combination of a group of candidates measured in the data or simulation. To allow both a measurement of Δm_d and Δm_s, a broad mass window for the D_q^- candidate is used, which covers both the D^- and D_s^- masses, $-200\,\mathrm{MeV\,c^{-2}} < m(K^-K^+\pi^-) - m_0(D_s^-) < 100\,\mathrm{MeV\,c^{-2}}$. Decays of the type $D^{*+} \to D^0\pi^+$ (where $D^0 \to K^-K^+$) can mimic the $K^-K^+\pi^-$ final state of the D_q^-. These events are suppressed by requiring that $m(K^-K^+\pi^-) - m(K^+K^-) - m_0(\pi^+) > 15\,\mathrm{MeV\,c^{-2}}$ and $m(K^+K^-) < 1840\,\mathrm{MeV\,c^{-2}}$. The first requirement causes the pion to have a reasonable fraction of the B_q^0 momentum; because of the small mass difference between the D^* and D^0, any pions from a $D^{*+} \to D^0\pi^+$ decay have low momenta and these are removed with this requirement. The second requirement attempts to remove any $D^0 \to K^-K^+$, $D^0 \to K^-\pi^+$ or $D^0 \to \pi^-\pi^+$ decays; when any of these decays are reconstructed as two kaons their mass will be approximately equal to or greater than the nominal D^0 mass and $1840\,\mathrm{MeV\,c^{-2}}$ is just less than the nominal value [1]. Simulations indicate that the selected sample is dominated by $B_s^0 \to D_s^-\mu^+\nu X$, $B_d^0 \to D^-\mu^+\nu X$, $B^+ \to D^-\mu^+\nu\pi^+ X$ and $B^+ \to D_s^-\mu^+\nu K^+ X$ decays, where the X represents any number of ν, π^0 and γ. The ν, K^+ and π^+ from the B^+ decays and the possible additional ν, π^0 and γ particles are ignored in the reconstruction and so a clear B_q^0 mass peak is not reconstructed. A normalised mass (n) is defined as

$$n = \frac{m(K^-K^+\pi^-\mu^+) - m_0(D_q^-) - m_0(\mu^+)}{m_0(B_q^0) - m_0(D_q^-) - m_0(\mu^+)}; \tag{6.1}$$

this is used as a proxy for the B_q^0 mass in most parts of this analysis. To calculate n, an assumption must be made as to whether the candidate is

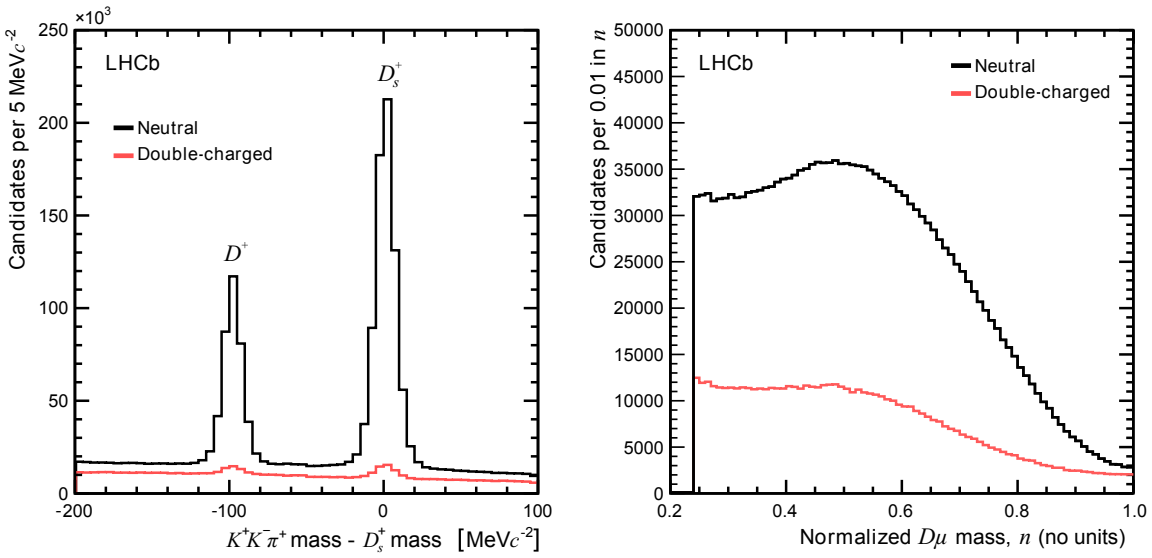

Figure 6.1: Mass distributions for the selected signal candidates. The left plot shows the $K^-K^+\pi^-$ invariant mass, where the known mass of the D_s^- has been subtracted. The right plot shows the normalised mass as defined in Eq. (6.1). Neutral candidates ($D_q^-\mu^+$) are possible signal decays, while the double-charged candidates ($D_q^-\mu^-$) are background.

a B_d^0 or B_s^0 decay. If $m(K^-K^+\pi^-)$ is greater than the midpoint between the D^- and D_s^- mass peaks ($m(K^-K^+\pi^-) - m(D_s^-) > -50\,\text{MeV}\,\text{c}^{-2}$) the candidate is assumed to be a B_s^0 decay, otherwise it is assumed to be a B_d^0 decay. The value of n is required to be in the range $0.24 < n < 1.0$; the lower cut removes mostly combinatorial background while the upper cut removes unphysical B_q^0 masses.

The $K^-K^+\pi^-$ invariant mass distribution and the n distribution of the selected candidates are shown in Fig. 6.1, where the D_q^- peaks can clearly be seen over the background. In addition to the selected dataset where the B_q^0 candidate is formed from a D_q^- candidate and a μ^+ candidate (with opposite charges), another dataset is created where the B_q^0 candidate is formed from

a D_q^- candidate and a μ^- candidate (with the same charge); the former dataset is a possible signal decay, whereas the latter dataset is background. The analysis is performed on the former dataset, but both datasets are shown in Fig. 6.1 so that the peaking background from the prompt D_q^- decays can be seen in the double-charged dataset; the backgrounds are discussed further in Section 6.4. A clear B_q^0 peak cannot be seen (in the n distribution) as the neutrino (and other possibly missing or ignored particles) in the decay $B_q^0 \to D_q^- \mu^+ \nu$ remove a large fraction of the momentum of the B_q^0. However, the n distribution does show that the neutral data has different shape to that of the double-charged dataset. The neutral dataset has a larger hump at around $n = 0.5$ than the double-charged dataset; this is expected to be caused by the signal.

6.2 Flavour tagging

To measure mixing, knowledge about the flavour of the B_q^0 meson is needed at two points in time. In this analysis these two points are at production and decay of the B_q^0 meson. The flavour state of the B_q^0 at decay is easy to determine when using a decay like $B_q^0 \to D_q^- \mu^+ \nu$, since it is flavour specific; the B_q^0 always decays to $D_q^- \mu^+ \nu$ and the $\overline{B}_q^0$ always decays to $D_q^+ \mu^- \nu$. Determination of the initial flavour state is performed using the standard LHCb flavour-tagging algorithms [120–122].

The flavour tagging algorithms work by searching for particles which are produced with, and are flavour correlated to, the signal B_q^0 meson. When the signal B_q^0 is created, it is usually formed from a $b\overline{b}$-pair and so there is a b-quark of opposite flavour in the same event which can be searched

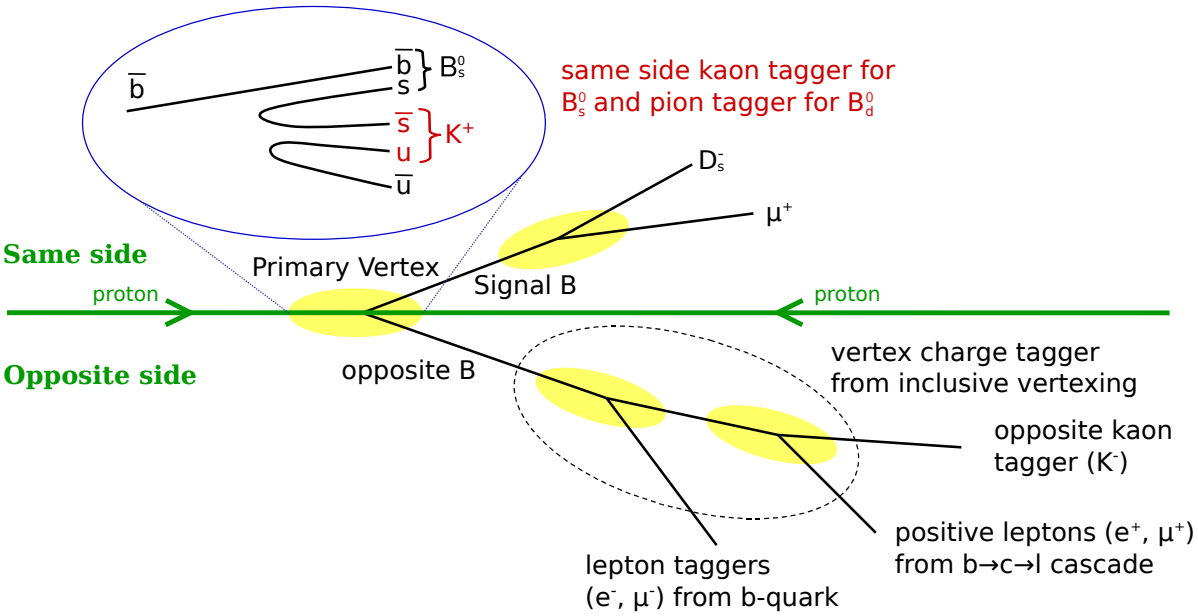

Figure 6.2: Diagram demonstrating the different methods which can be used to determine the flavour of the B_q^0 candidate. Image taken from Ref. [123].

for. This is referred to as the opposite-side (OS) quark and algorithms which search for this are OS tagging algorithms. The OS quark and all tagging methods are illustrated in Fig. 6.2. The OS tagging algorithms use the charge of various particles to tag the flavour of the OS b-quark: the meson that was formed from the OS b-quark (by inclusively vertexing the OS b-quark decay), the lepton from a semileptonic decay of the b-quark, the lepton from a $b \to c \to \ell$ transition and the kaon from a $b \to c \to s$ transition. Alternatively algorithms can also look at particles which were created in the decay which produced the signal B_q^0 meson; these are known as same-side (SS) tagging algorithms. The SS pion and kaon tagging algorithms (used to tag B_d^0 and B_s^0 mesons respectively) use the fact that during hadronisation an additional pion or kaon could be created in association with the signal meson using one of the $d\bar{d}$-pair or $s\bar{s}$-pair that was created in association with the B_q^0

meson. The results of the various tagging algorithms are combined within a neural network [121]. Despite the use of all these tagging algorithms, these features which reveal the initial flavour state can only be found in around a third of events (either because they are not there or cannot be found). The fraction of events which have any tagging decision attributed to them is called the tagging efficiency (ε_{tag}). For tagged events, the neural network provides both a tag and an event-by-event probability that the tag was incorrectly assigned. The average probability that the tag was incorrectly assigned is known as the mis-tag probability (ω) and is also usually around a third. The tagging power is equivalent to the tagging efficiency if all events were correctly tagged and is defined as $\varepsilon_{\text{tag}}(1-2\omega)^2$. In order to maximise the tagging power, it is required that the event-by-event mistag probability from the neural network is < 0.42; this was recommended by the LHCb flavour tagging group. The selected dataset comprises 594 845 flavour-tagged candidates.

In the same way that the calculation of n requires a choice of whether a candidate is a B_d^0 or B_s^0 decay, the tagging algorithms need to assume a certain candidate is a B_d^0 or B_s^0 decay; the midpoint between the D^- and D_s^- mass peaks ($m(K^-K^+\pi^-) - m(D_s^-) = -50\,\text{MeV}\,\text{c}^{-2}$) is again used for this purpose.

In order to measure B_q^0 mixing, the change in flavour must be studied and so odd and even flavour tags are defined. Candidates with an odd flavour tag have been reconstructed to have different initial and final flavour states, while candidates with an even flavour tag have been reconstructed with the same initial and final flavour states. It is the tagging asymmetry of the odd and even tags $\left(\frac{N_+ - N_-}{N_+ + N_-}\right)$ that will show oscillations as a function of

the measured decay-time, where $N_\pm$ is the number of even (+) or odd (−) flavour tagged events.

6.3 Decay-time distributions

The following is used to calculate the reconstructed proper decay-time (τ) of the B_q^0-meson candidates,

$$\tau = \frac{m \times \left(\vec{F} \cdot \vec{p}\right)}{|\vec{p}|^2}\,, \tag{6.2}$$

where m is the mass of the B_q^0 candidate, $\vec{F}$ is the flight distance vector and $\vec{p}$ is the momentum vector. The flight distance vector can be calculated due to the excellent spacial resolution of the VELO, but the B_q^0 candidate momentum and mass cannot be fully reconstructed due to the missing or ignored particles from the B_q^0 decay. A statistical simulation-based correction [25] is applied to the reconstructed decay-time, called the k-factor; this method has been successfully used in previous experiments [124, 125]. Using simulations, the ratio of reconstructed momentum to true momentum (k) is plotted as a function of n; an average k is then taken and fitted as a function of n using a fourth-order polynomial The fitted function is shown in Fig. 6.3. Since this polynomial is parametrised as a function of n, the same correction can be used for both B_d^0 and B_s^0 candidates. The measured decay-time (t') can then be calculated by dividing the reconstructed proper decay-time by the k-factor $t' = \tau/k(n)$. Using this decay-time correction method Δm_d and Δm_s both have small biases of $\sim 1\,\%$; these biases are measured with simulated events and a correction is applied to the final result.

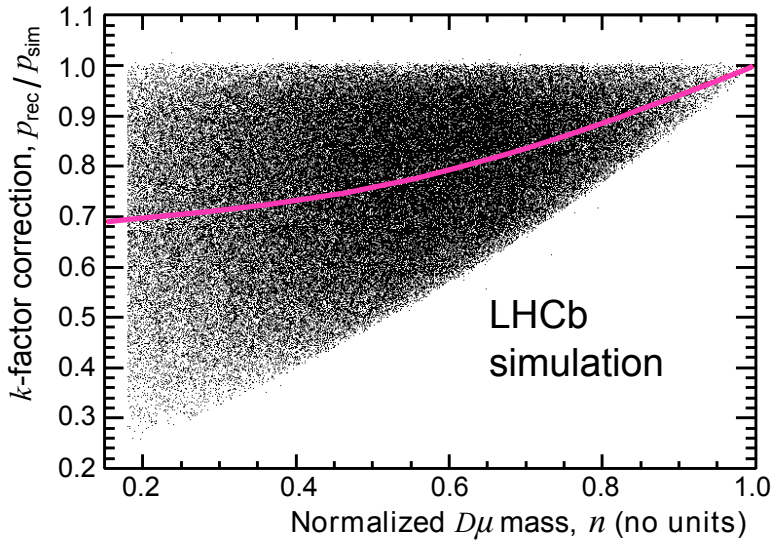

Figure 6.3: Ratio of reconstructed momentum to true momentum as a function of normalised mass from simulated B_s^0 decays. The magenta curve shows the result of a fit to the mean of the distribution as a function of n using a fourth-order polynomial.

Unlike the theoretical prediction of the B_q^0 proper decay-time distribution, the real distribution has imperfect tagging and is smeared by a resolution function $R(t, t'-t)$, where t is the true time. Further the selection contains requirements (such as those on the IP) which are highly correlated to the decay-time of a candidate; this introduces a time-dependent selection efficiency $\varepsilon(t')$. Given these experimental effects, the decay-time distribution of reconstructed B_q^0 mesons is given by

$$N_{\pm}(t') = N\,\varepsilon_{\rm tag}\,\frac{e^{-\Gamma t}}{2}\Big[\cosh(\Delta\Gamma\, t/2) \pm (1-2\omega)\cos(\Delta m\, t)\Big] \otimes R(t, t'-t) \times \varepsilon(t')\,, \tag{6.3}$$

where the subscript ± represents even (+) and odd (−) tagged events and N is the number of signal B_q^0 candidates. The time-dependent efficiency is

parametrised using an empirical acceptance function and consists of the sum of several Gaussian functions, $G(t';\mu,\sigma)$, where μ is the mean of the Gaussian and σ is the width. The acceptance function is given by:

$$\varepsilon(t') = \begin{cases} 0\,, & \text{if } t' < \mu_0 \\ 1 - f\,G(t';\mu_0,\sigma_1) - (1-f)\,G(t';\mu_0,\sigma_2)\,, & \text{otherwise} \end{cases} \tag{6.4}$$

This function is motivated by studies on data and simulated events [25]. Typical values of the Gaussian parameters are $\sigma_{1,2} < 1\,\text{ps}$ and $\mu_0 \approx 0.01\,\text{ps}$.

As can be seen in Fig. 6.3 there is a large spread in the k-factor, especially at low values of n. This uncertainty on the momentum dominates the

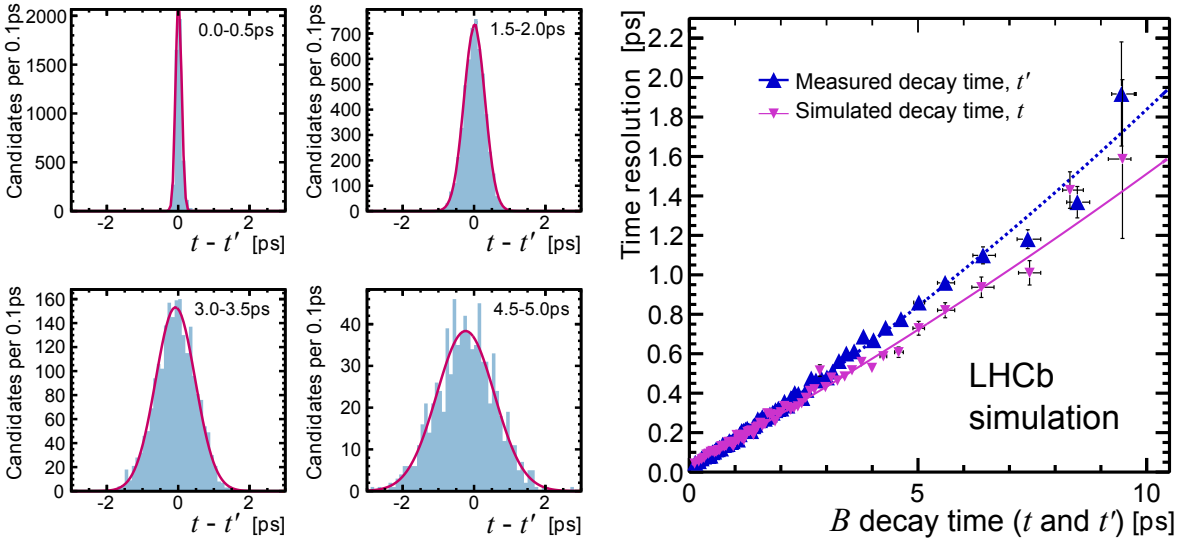

Figure 6.4: Decay-time resolution as measured with simulated B_s^0 decays. The left four plots show the decay-time resolution (light blue) in bins of measured decay-time fitted with a Gaussian function (dark pink). The right plot shows the sigma of Gaussian fits plotted as a function of the measured (blue) and simulated (purple) decay-time; this is fitted with a quadratic or cubic function for measured and simulated decay-times respectively.

uncertainty on the decay-time resolution $\sigma(t')/t' \approx \sigma(p)/p$ and so $\sigma(t')$ is approximately proportional to t'; this is shown in Fig. 6.4, where $\sigma(t')$ shows an almost linear relation to t'. There are two approaches used in the multivariate fit to parametrise the dependence of the decay-time resolution and both use simulated events. The first method parametrises the resolution as a function of the true decay-time using a computationally expensive numerical convolution, while the second and default method parametrises the resolution as a function of the measured decay-time using an analytical function. The dependence of the decay-time resolution on the measured or true decay-time is fitted with a polynomial function.

6.4 Backgrounds

There are three main types of background expected to be selected in this analysis: combinatorial backgrounds, prompt D_q^- decays and semileptonic B^+ decays.

Combinatorial backgrounds are formed when three random reconstructed tracks are combined to form a D_q^- candidate. Because this background is randomly formed, it is non-peaking in the $K^-K^+\pi^-$ mass distribution. The combinatorial background comes in two types: detached and prompt. The prompt combinatorial backgrounds peak around zero measured decay-time because the reconstructed particles in the D_q^- candidate came directly from the PV and not a B_q^0 decay. Conversely, a detached background candidate contains at least one reconstructed track from a SV and so these background B_q^0 candidates tend to have larger lifetimes.

The prompt D_q^- background is so called because the D_q^- mesons originate from the PV. These prompt D_q^- mesons are combined with a μ^+ (again from the PV) as if they were both produced in a decay from a B_q^0. This background peaks under the D_q^- mass signal region, but because of the large difference between the lifetime of the D_q^- (0.4 ps) and B_q^0 ($\sim$1.5 ps) this background can be separated from the signal in the multivariate fit. This background can be seen as a peak in the double charged data shown in Fig. 6.1.

For an event to be considered signal, it must peak in $K^-K^+\pi^-$ mass and be detached (not peak around zero in measured decay-time). The B^+ background is both peaking in $K^-K^+\pi^-$ mass and detached, which makes it impossible to distinguish from signal in the multivariate fit and so is a very important background to understand. This B^+ background is readily selected as it fulfils many of the selection criterion for a B_q^0 decay; the only difference is that B^+ events will have an additional charged particle which is ignored by the selection, which only searches for 4 final state particles. It should be noted that B^+-mesons do not mix, unlike the signal B_q^0-mesons and this has important consequences when trying to measure B_d^0 mixing.

6.5 Mixing frequency extraction methods

Two methods are used to measure the mixing frequencies: a binned multivariate log-likelihood maximization fit of the data and a model-independent analysis of the Fourier spectrum. The Fourier analysis provides a cross-check to the primary multivariate fit method. Both methods reconstruct the data and correct for missing momentum using the same techniques and so share the systematic effects of the detector alignment and the k-factor correction.

6.5.1 Multivariate fits to the data

The data are fitted using a binned multivariate log-likelihood maximization program created with the RooFit framework [99, 100]. Because the decay-time resolution is heavily correlated to the normalised mass, the dataset is split into two regions about the point $n = 0.56$ (this value was determined using fast-simulations) and the fit is performed simultaneously to the two subsamples. The events with higher normalised mass have a better decay-time resolution and this helps to ensure that the fast oscillations of the B_s^0-meson can be resolved. The fit is performed in 101,000 bins of the two continuous dimensions ($K^-K^+\pi^-$ mass and measured decay-time) and the two discrete dimensions (normalised mass, $n < 0.56$ or $n > 0.56$, and the tagging result, either odd or even). There are seven types of events which are described with PDFs in the fitting program: prompt and detached combinatorial backgrounds, prompt D^- and D_s^- backgrounds, B^+ in the D^- mass peak and most importantly the B_d^0 and B_s^0 signal; these PDFs can be seen in Figs. 6.5 to 6.7.

The $m(K^-K^+\pi^-)$ distribution is shown in Fig. 6.5, with all components of the fit overlaid. The two combinatorial backgrounds consist of a second-order polynomial that models the $m(K^-K^+\pi^-)$ distribution and the decay time distribution consists of an exponential decay multiplied by an acceptance function of the form shown in Eq. (6.4). The prompt combinatorial background has a shorter lifetime and sharper acceptance function than the detached combinatorial background and the ratio of the two combinatorial backgrounds varies linearly as a function of $m(K^-K^+\pi^-)$. For the prompt background that is known not to mix, the tagging asymmetry is set to be constant as a function of time. Generic detached backgrounds are allowed

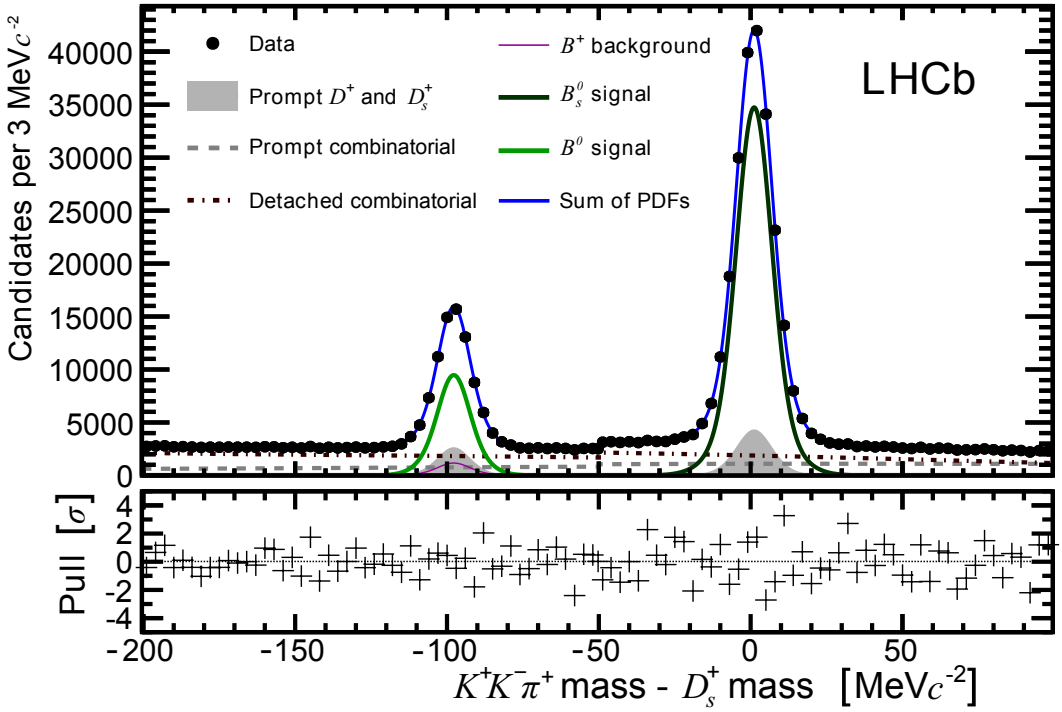

Figure 6.5: Invariant mass distribution of the $K^-K^+\pi^-$ combination, where the known mass of the D_s^- has been subtracted. The points show the data while the lines shows the components of the fit. The small step at $-50\,\text{MeV}\,\text{c}^{-2}$ is the result of differences in tagging efficiency for the B_d^0 and B_s^0 hypotheses. The figure under the main plot shows the difference between the data point and the fitted line divided by the uncertainty on the data point.

to have a time-dependent asymmetry varying as an arbitrary quadratic polynomial.

The two signal PDFs consist of a sum of two Gaussians that model the $m(K^-K^+\pi^-)$ distribution and the decay-time PDF is described in Eq. (6.3), with two separate copies to model the B_d^0 and B_s^0. The only difference between the two PDFs is that the B_s^0 PDF has an additional tagging asymmetry

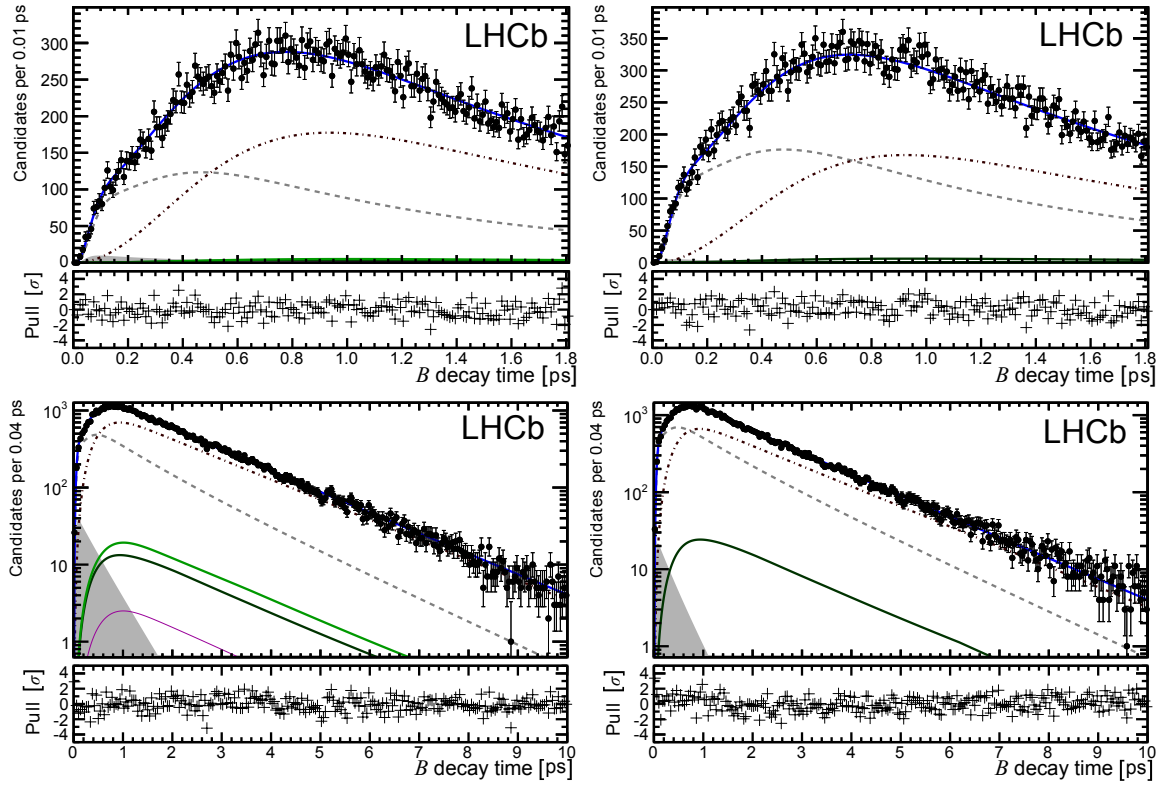

Figure 6.6: Measured B_q^0 candidate decay-time distributions, overlaid with projections of the fit, for background-dominated regions. Top left: a region between the two signal peaks, $-80\,\text{MeV}\,\text{c}^{-2}$ to $-20\,\text{MeV}\,\text{c}^{-2}$ (with respect to the known mass of the D_s^-), showing only low decay-times. Top right: a region to the right of the signal peaks $20\,\text{MeV}\,\text{c}^{-2}$ to $100\,\text{MeV}\,\text{c}^{-2}$, showing only low decay-times. Bottom row: the same as above, but with an extended decay-time scale and a logarithmic y-axis. The legend is the same as in Fig. 6.5.

that is flat as a function of proper time; this is added so that the B^+ in the B_s^0 signal region is accounted for. The value of $\Delta\Gamma_d$ is fixed to zero, and $\Delta\Gamma_s$ is fixed using the result of an LHCb analysis [126]. Figures 6.6 and 6.7 show the measured decay-time distributions; the former shows the background-dominated region either side of the D_s^- peak, while the latter shows the two signal regions.

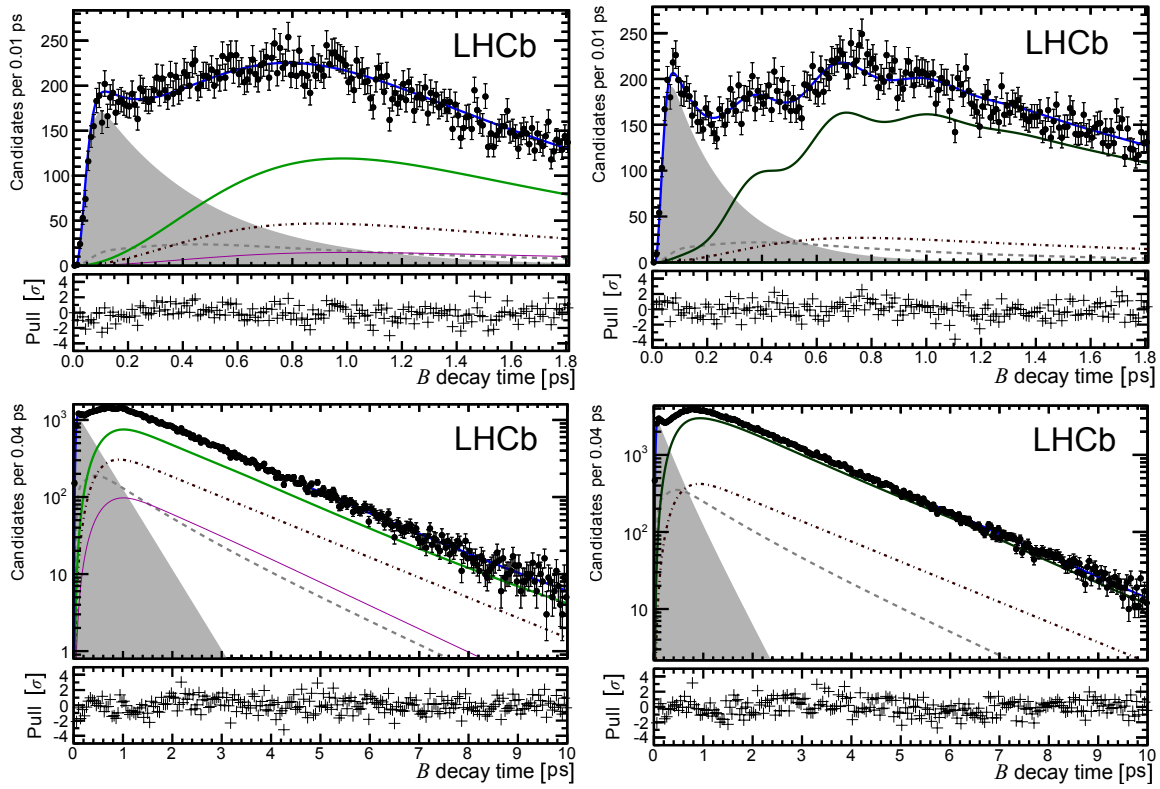

Figure 6.7: Measured B_q^0 candidate decay-time distributions, overlaid with projections of the fit in signal-dominated regions. Top left: odd-tagged, all n regions and within a mass window $\pm 20\,\mathrm{MeV\,c^{-2}}$ around the D^- mass peak, showing only low decay-times. Top right: odd-tagged, high-n region within $\pm 20\,\mathrm{MeV\,c^{-2}}$ of the D_s^- mass peak, showing only low decay-times, where B_s^0 oscillations can be clearly seen. Bottom row: both tags, all n regions and within a mass window of $\pm 20\,\mathrm{MeV\,c^{-2}}$ around the D^- (left) and D_s^- (right) mass peaks. The legend is the same as in Fig. 6.5.

The two prompt D^- and D_s^- backgrounds share the same mass PDFs as the signal B_d^0 and B_s^0, respectively. The proper time model of the prompt D^- and D_s^- background PDFs is the same as the prompt combinatorial background, but it is allowed to float independently.

Only the signal mass shapes and the parameters of interest, Δm_s and Δm_d, are shared between the two subsamples in n, which are fitted simultaneously; all other parameters of the PDFs described above are allowed to float independently between the two n subsamples.

The final components of the fit are the B^+ backgrounds in the D^- and D_s^- peaks. The B^+ backgrounds cannot be distinguished from the signal; several assumptions must be made and multiple variables must be fixed in order to fit these backgrounds. Simulation studies show that the measured decay-time and D_q^- candidate mass distributions are very similar for the B_d^0 and B^+ and so the shape of the B^+ PDF is fixed to be the same as the B_d^0 signal in both $K^-K^+\pi^-$ mass and measured decay-time, but with no mixing in the tagging asymmetry. The B^+ fraction is fixed in the fit to 11 % (relative to the amount of B_d^0) and it is later varied within its uncertainty of 2 %. This relative fraction of B^+ is calculated using the selection efficiencies from simulation, branching fractions and the known ratio of the relative production rates of the B^+ and B_d^0 known as the fragmentation fractions ($f_u/f_d \sim 1$); see Appendix C for further information.

Since the B^+ meson does not mix, in an ideal world all events would be even tagged and so the PDF would have a tagging asymmetry of 1. However, ω is not zero and so this results in a constant tagging asymmetry for the B^+ PDF of $\big(1-2\omega(B^+)\big)$. It is initially assumed that $\omega(B_d^0) = \omega(B^+)$ in order to fix the tagging asymmetry of the B^+ PDF; this is particularly useful as $\omega(B_d^0)$ is measured in the fit to the data. To minimise any differences in the performance of the tagging algorithms when studying B_d^0 or B^+ decays, the tagging algorithms that are sensitive to the signal B are not used; this means that only OS tagging algorithms are used for the B_d^0 hypothesis.

However, simulated datasets show that the OS tagging algorithms do not have identical performances on B_d^0 and B^+ decays; this means that the assumption that $\omega(B_d^0) = \omega(B^+)$ is incorrect and so correction factors are applied. This is mostly caused by the inclusive vertex charge tagging algorithm which often includes the pion from the B^+ decay when inclusively vertexing the OS b-decay. Due to differences in tagging efficiency the B^+ fraction is enhanced by $\varepsilon_{\text{tag}}(B^+)/\varepsilon_{\text{tag}}(B_d^0) = (38.04 \pm 0.17\,\%)/(33.88 \pm 0.14\,\%)$ or $12.28 \pm 0.68\,\%$ and the B^+ mistag fraction is worsened by $\omega(B^+)/\omega(B_d^0) = (39.8 \pm 0.3\,\%)/(34.6 \pm 0.2\,\%)$ or $15.0 \pm 1.1\,\%$. Since the overall tagging asymmetry is proportional to $\varepsilon_{\text{tag}}(1 - 2\omega)$, these correction factors are equivalent to an overall drop in the amount of B^+ by $25.6 \pm 2.4\,\%$. The fraction of B^+ relative to the amount of B_d^0 is corrected by 25 % and the full value of this correction is taken as a systematic uncertainty. There is enough information to include a B^+ PDF in the fit given that: only the OS tagging algorithms are used for B_d^0 candidates, the relative fraction of B^+ is fixed, the B_d^0 and B^+ have similar tagging performances (after correction factors have been applied) and the shape of the B^+ PDF is the same as the signal B_d^0 PDF. All PDF parameters of the B^+ background in the D^- peak are fixed.

While a PDF is included for B^+ decays contributing to the D^- mass peak, none is included for those contributing to the D_s^- peak. Instead the distribution for the background B^+ candidates is included in the B_s^0 signal PDF; this is done by introducing an offset in the tagging asymmetry which is allowed to vary in the data fit. The B_d^0 mixing oscillation has a period of 12.320 ± 0.072 ps and due to the limited number of events at high decay-times only the region from 0 ps to 10 ps is studied; this means that it is not possible to see a whole B_d^0 oscillation and so the measurement becomes sensitive to the overall tagging asymmetry in the signal region. This is the

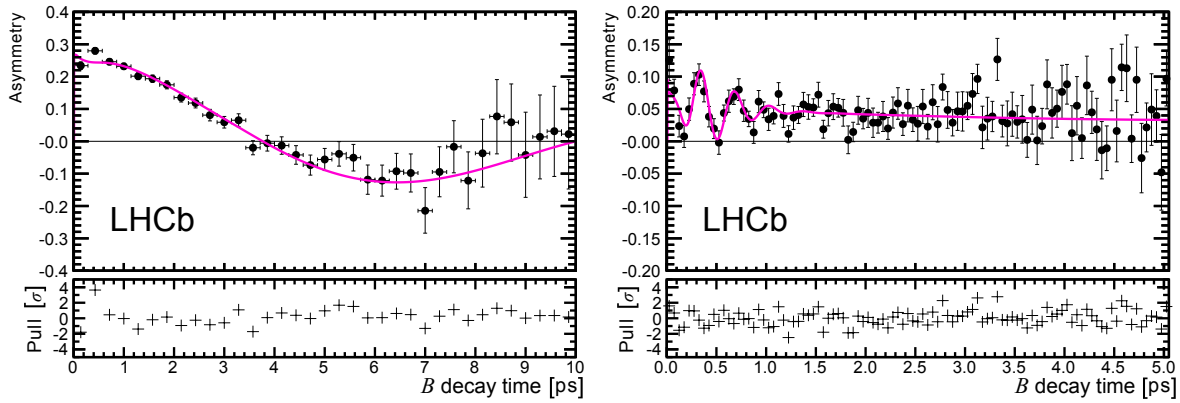

Figure 6.8: Tagged (mixing) asymmetry, $(N_+ - N_-)/(N_+ + N_-)$, as a function of measured decay-time. The left plot shows the asymmetry for events within a mass window of $\pm 20\,\mathrm{MeV\,c^{-2}}$ around the D^- mass peak, and the right plot shows the corresponding asymmetry around the D_s^- mass peak. The black points show the data, and the curves are projections of the fitted PDF. On the right plot the fast oscillations of B_s^0 are gradually washed out by the worsening decay-time resolution.

reason that an offset in the tagging asymmetry cannot be allowed to float in the B_d^0 signal PDF; doing so results in a 50 % reduction of the statistical power of the fit for Δm_d.

In Fig. 6.8 the tagging asymmetry is plotted, and the fitted asymmetry curve is overlaid. The large period of the B_d^0 oscillations is visible in the left figure, where almost a whole oscillation can be seen. The figure also shows that the B_d^0 oscillations have a positive offset (caused by the B^+ background), but this is much clearer in the case of the B_s^0 oscillations in the plot on the right of this figure. Around 3 clear B_s^0 oscillations can be seen, the magnitude of which initially increases with the efficiency of the acceptance function, but then decreases as the decay-time resolution becomes poorer. The goodness

of the multivariate fit is verified with a local density method [127], which finds a p-value of 19.6 %.

Results and systematic uncertainties

Table 6.2 gives the fitted results for some quantities in each of the regions of normalised mass; this includes the fraction of both prompt and detached, D^- and D_s^- candidates along with the mistag probabilities. The signal lifetimes are also measured, but these have very large systematic uncertainties associated to them; due to the large size of these uncertainties they are not analysed in detail and so the results are not quoted. There are several

Table 6.2: A selection of fitted parameter values; these only include statistical uncertainties. The B_s^0 signal fraction includes contributions from any detached D_s^- production. When the omitted fractions (of combinatorial background components) are included, the total fraction sums to unity within each n region separately.

Quantity	Normalised mass region	
	Low-n	High-n
Fit fraction of:		
- B_s^0 signal	0.3247 ± 0.0029	0.3604 ± 0.0023
- B_d^0 signal	0.0781 ± 0.0017	0.0968 ± 0.0022
- prompt D_s^-	0.0410 ± 0.0026	0.0444 ± 0.0018
- prompt D^-	0.0196 ± 0.0018	0.0311 ± 0.0024
Mistag probability ω:		
- B_s^0 signal	0.347 ± 0.054	0.333 ± 0.021
- B_d^0 signal	0.3567 ± 0.0063	0.3319 ± 0.0065
Total candidates	368 965	225 880

sources of systematic uncertainty on Δm_s and Δm_d, which are summarised in Table 6.3 and discussed in detail below.

Assumptions on B^+ decays

The measurement of Δm_d is sensitive to the integrated mixing probability (χ_d) and so it is sensitive to the amount and the shape of the non-mixing B^+-background. Most aspects of this background are fixed from external sources and so these are varied and the largest change in Δm_d is assigned as a systematic uncertainty. The lifetime of the B^+ is about 7 % greater than the B_d^0 lifetime and so the relative lifetime of the B^+-background is varied by ±7 %. The relative amount of B^+ is also varied within its uncertainty to 9 % and 13 %. The tagging is varied in several ways: the tagging performance correction factors are removed, all tagging algorithms (SS and OS) are used, a tagging asymmetry off-

Table 6.3: Sources of systematic uncertainty on Δm_s and Δm_d. Simulation implies a combination of full LHCb simulation and pseudo-experiment studies.

Source of uncertainty	Method	Systematic uncertainty	
		Δm_s [ps^{-1}]	Δm_d [ps^{-1}]
B^+ ($\mathcal{B}$, efficiency, tagging)	Data refit	n/a	0.008
Signal decay-time model	Data refit	0.09	0.007
Residual correction bias	Simulation	0.09	0.0055
k-factor simulation	Simulation	0.06	0.0052
Mass models and binning	Data refit	0.05	0.001
Detector alignment	Calibration	0.03	0.0008
Values of $\Delta\Gamma$	Data refit	0.00	0.0004
Total	Sum in quadrature	0.15	0.013

set is added (as is done in the B_s^0 signal PDF) and finally odd and even tagged events are allowed to have different mistag rates. The changes to the tagging produce the largest variation in the fitted value of Δm_d and a systematic uncertainty of $0.008\,\text{ps}^{-1}$ is assigned to Δm_d. The B^+-background does not effect Δm_s.

Signal measured decay-time model

The signal PDF is altered in several ways to investigate any bias or systematic uncertainties that could arise from the use of the signal PDF with a resolution that is parametrised as a function of the measured decay-time. The fit is repeated with the resolution parametrised as a function of the true decay-time. The B_d^0 signal candidate PDF is replaced with a PDF that contains no time-dependence; this cannot be done for the B_s^0 signal candidate PDF, because at larger decay-times the decay-time resolution quickly becomes greater than the period of the B_s^0 oscillation and obscures any visible oscillations. Further, any parameters of the PDF which are held constant have their values altered by ± 1 estimated uncertainty. These changes result in small variations in the fitted values of Δm_d and Δm_s and so systematic uncertainties of $0.007\,\text{ps}^{-1}$ and $0.09\,\text{ps}^{-1}$ are assigned respectively.

The *k*-factor simulation

The *k*-factor is a simulation-based correction method and so differences between the simulation and reality that alter the momentum spectra of the reconstructed or ignored particles could invalidate the correction. For example, if the D^{**} branching fractions or form factors were incorrectly simulated the number of missing or ignored particles would be altered and thus the reconstructed momentum spectrum would

be altered as well. A large number of pseudo-datasets are generated using various underlying k-factor distributions (determined from full-simulations); these datasets are then fitted with the nominal PDF to determine the effect of an incorrectly modelled k-factor distribution. Several k-factor distributions were altered to increase the fraction of $D^{*(*)}$ decays by up to a factor of 10, which makes this study sensitive to the effects of incorrectly modelled form factors and branching fractions of the $D^{*(*)}$ decays. By comparing k-factor distributions from the nominal 2010 simulations to the simulations from 2009, this study is also sensitive to significant changes in the beam parameters. Two groups of datasets were also created where the fitted k-factor was not used, but instead each coefficient of the fitted polynomial was altered by ±1 uncertainty from the fit; this makes the study sensitive to the polynomial fitting procedure. The last group of datasets which were created used the B_d^0 simulation instead of the nominal B_s^0 simulation; this makes the study sensitive to non-flavour symmetric effects and mass-sensitive effects. The decay-time resolution (also measured using simulation) is a direct measure of how well the k-factor corrects the decay-time distribution. If the simulation predicted a better or worse resolution than reality, this would also affect the fitted values of Δm_q. The constant term in the time-dependence of the resolution is altered by ±40 % and rotated about the point $t = 1\,\text{ps}$ until the linear term is altered by ±10 %. Overall these changes caused variations in the fitted values of Δm_d and Δm_s of $0.0052\,\text{ps}^{-1}$ and $0.06\,\text{ps}^{-1}$, respectively; these variations are assigned as systematic uncertainties. This currently represents the limit of this technique without more precise knowledge of the various sub-decays, as listed in Appendix C.

Residual correction bias

The k-factor correction is not perfect and causes a small bias to the measured decay-time distribution. The effect of this on Δm_q is measured using the full simulation, where the true flavour is used to tag the B_q^0 at production instead of the flavour tagging algorithms; this reduces the statistical uncertainty on this correction. Biases of around 1 % are found for both Δm_d and Δm_s; these biases are then corrected for. Similar biases are found using pseudo-experiment studies and the Fourier analysis of the full-simulation. Half of the correction is assigned as a systematic uncertainty.

Mass models and binning

The mass PDFs of the signal and background components are altered to ensure that they do not affect the values of Δm_q. The order of the Chebychev polynomial is decreased to 3 and increased to 5, while some of the Gaussian functions in the mass peaks are exchanged for CB functions [101]. The fits are nominally performed to binned datasets; this binning is removed and the fit performed unbinned. These changes to the mass PDF and binning result in changes of the values of Δm_d and Δm_s by $0.001\,\text{ps}^{-1}$ and $0.05\,\text{ps}^{-1}$ and so these are assigned as systematic uncertainties.

Detector alignment

Uncertainties in the momentum scale, decay-length scale and track position arise from the uncertainty on the alignment of the detector; this produces changes in the reconstructed mass and lifetime as function of the opening angle decay products. The alignment has been studied with surveys of the detector, and the effects of residual misalignments have

been studied using several control modes [128]; the systematic uncertainties on Δm_d and Δm_s caused by the uncertainty of the alignment are $0.0008\,\text{ps}^{-1}$ and $0.03\,\text{ps}^{-1}$ respectively.

Values of $\Delta\Gamma$

Both $\Delta\Gamma_d$ and $\Delta\Gamma_s$ are held constant in the nominal fit to the data. When the value of $\Delta\Gamma_d$ is varied by $\pm 5\,\%$ (chosen to cover well the experimental uncertainty, given the lack of information on its sign [129]) and $\Delta\Gamma_s$ is varied within its uncertainty [126]. The fitted results of Δm_q are altered by a small fraction and this change is assigned as a systematic uncertainty.

To ensure there is no unknown dependence on trigger variables or magnet polarity the data are split into two subsamples and the fit results compared to the nominal values. Both when the data are split by magnet polarity and when the data are split depending on how the event was triggered, the two fitted results are compatible with the nominal result. The B_q^0 mixing frequencies have been determined to be

$$\Delta m_d = 0.503 \pm 0.011\,(\text{stat}) \pm 0.013\,(\text{syst})\,\text{ps}^{-1}\,,$$
$$\Delta m_s = 17.93 \pm 0.22\,(\text{stat}) \pm 0.15\,(\text{syst})\,\text{ps}^{-1}\,.$$

The value of the fitted likelihood is studied in order to measure the significance of the observed oscillations. The nominal value of the likelihood is compared to that at the extrema of the search window $\Delta m_q = 0\,\text{ps}$ or $\Delta m_q \geq 50\,\text{ps}$; both of these values result in no oscillations with flat mixing asymmetry curves (cf. Fig. 6.8). The null hypotheses of no oscillations is rejected with a p-value equivalent to $13.0\,\sigma$ for B_d^0 oscillations and $5.8\,\sigma$ for B_s^0 oscillations.

6.5.2 Fourier analysis

In a addition to the nominal multivariate fitting approach the B_q^0 mixing frequencies can also be extracted using the well-established Fourier transform techniques [130–132]. When the data are split by odd and even tags the cosine component of the signal true decay-time distribution (cf. Eq. (6.3)) is positive for one tag and negative for the other; this feature simplifies the Fourier analysis of the data. The even-tagged Fourier spectrum of the measured decay-time can be subtracted from the odd-tagged Fourier spectrum to suppress the underlying background, while simultaneously enhancing the signal distribution. This allows a model-independent measurement of the B_q^0 mixing frequencies. The Fourier spectrum in a $25\,\mathrm{MeV\,c^{-2}}$ wide bin around the D_s^- mass is shown in Fig. 6.9. The value and uncertainty of Δm_s is measured using this plot and is found to be $\Delta m_s = 17.95 \pm 0.40\,\mathrm{ps^{-1}}$. The RMS of this distribution is used as a model-independent proxy for the statistical uncertainty.

Pseudo-experiment studies were conducted where the measured decay-time was generated with known values of Δm_d and Δm_s. These decay-time distributions underwent the Fourier analysis and it was found that the statistical variation on Δm_s was $0.32\,\mathrm{ps^{-1}}$; this value is slightly smaller than given by the RMS. The pseudo-experiment studies show that Δm_s can be accurately extracted with reasonable precision using the Fourier transform, but that the value of Δm_d is highly biased and has a low precision. The magnitude of the Δm_s peak in the Fourier spectrum is significantly larger than the Δm_d peak; this is because several B_s^0 oscillations are visible, whereas for Δm_d the period of the oscillation is larger than the decay-time range which is studied. Further, residual background components of the

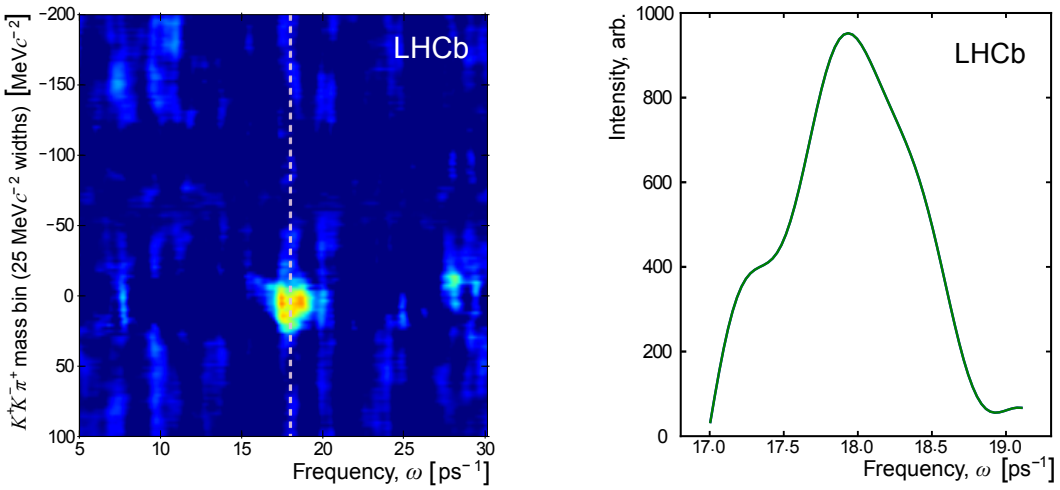

Figure 6.9: Fourier analysis of the data, searching for the B_s^0 mixing frequency. The plot on the left was produced using 25 MeV c^{-2} wide bins of $K^-K^+\pi^-$ mass, which are then analysed in steps of 5 MeV c^{-2} to produce a smooth image. The colour scale (blue-green-yellow-red) is an arbitrary linear representation of the signal intensity; dark blue is used for zero and below. The vertical dashed line is drawn at 18.0 ps^{-1}. The right plot shows the Fourier spectrum around the D_s^- signal region and it is this plot which is used to measure the central value and RMS width.

subtracted Fourier spectrum have frequencies similar to Δm_d. Given the low magnitude of the peak and the relatively large background in this area, it is hard to extract an accurate value of Δm_d without good knowledge of the background spectrum. However, the situation is the opposite for extraction of Δm_s, which has low backgrounds and a high intensity signal peak. Because of the large bias and poor precision on Δm_d, a value is not quoted with this method.

Using Fourier analysis it was determined that

$$\Delta m_s = (17.95 \pm 0.40\,(\text{rms}) \pm 0.11\,(\text{syst.}))\,\text{ps}^{-1}\,, \tag{6.5}$$

where the statistical uncertainty is taken from the RMS instead of the pseudo-experiments in order to be model-independent. The systematic uncertainties from the detector alignment and the k-factor correction (as discussed in Section 6.5.1) apply to both the multivariate fit and the Fourier analysis and so are quoted in this result.

6.6 Conclusion

The mixing frequencies of the B_q^0-mesons have been measured using flavour-specific semileptonic decays of the type $B_q^0 \to D_q^- \mu^+ \nu$. Daughter particles of the B_q^0 mesons which were missing or ignored in the selection reduce the reconstructed momentum of the B_q^0 candidate and so a simulation-based statistical correction is used to correct the reconstructed decay-time. Two methods to extract the mixing frequencies are then presented: a multivariate fit and a Fourier analysis. Although both methods produced consistent results, the first method is more precise. The B_q^0 mixing frequencies are found to be

$$\Delta m_d = (0.503 \pm 0.011\,(\text{stat}) \pm 0.013\,(\text{syst}))\,\text{ps}^{-1}\,,$$
$$\Delta m_s = (17.93 \pm 0.22\,(\text{stat}) \pm 0.15\,(\text{syst}))\,\text{ps}^{-1}\,.$$

The hypotheses of no oscillations is rejected by $13.0\,\sigma$ for B_d^0 oscillations and $5.8\,\sigma$ for B_s^0 oscillations. This is the first observation of B_s^0-$\bar{B}_s^0$ mixing to be made using only semileptonic decays [133].

CHAPTER 7

Conclusion

This thesis has covered a wide range of topics, which have contributed to both the LHCb collaboration and physics knowledge in general. The thesis included: simulations and optimisations of the VELO upgrade detector, measurements of electron PID and trigger efficiencies with LHCb, the search for the lepton flavour violating decay $D^0 \to e^{\pm}\mu^{\mp}$ and finally the measurement of the B_s^0 and B_d^0 mixing frequencies using semileptonic decays.

7.1 The VELO upgrade

The LHCb upgrade plans to run at significantly higher luminosities, which requires the VELO to be replaced. Multiple studies were performed to inform decisions about the design and to ensure that it was optimal. Several possible designs were simulated and a silicon pixel-based microchannel-cooled detector was decided upon. Further it was decided to rotate the

VELO to simplify the installation of the modules; this also has the benefit of increasing the performance of the VELO if it had to be run while open.

The design was then optimised to minimise the impact parameter resolution, while ensuring the maximum number of particles were detected within the acceptance of LHCb. This was performed using simple geometric considerations, then improved upon with a ray-tracing simulation and finally the full LHCb simulation was used. This resulted in the module positions shown in Fig. 3.27, which have been chosen by the VELO group as the final design.

These studies ensured that the VELO upgrade out-performs the current VELO despite the more difficult operating conditions. The upgrade VELO is due to be installed during LS2 and will ensure that LHCb is able to make world leading precision measurements of flavour for many more years.

7.2 Electron PID and trigger efficiencies

Studies were performed to ensure that the standard method to measure the PID efficiency would provide correct results for electrons. It was found that the electron specific variables, on which the electron identification efficiency depended, were highly correlated to the three standard PIDCalib variables and so the standard PIDCalib method could be applied to electrons. A sample of $B^+ \to J/\psi \left(e^+e^-\right) K^+$ was added to the calibration datasets; this enables the whole collaboration to easily measure electron PID efficiencies and it is currently available for use. Further, the PIDCalib software was also extended to allow trigger decisions to be efficiently stored in the datasets and so the efficiency of trigger lines can now be evaluated with PIDCalib as

well. This lays the groundwork for TrigCalib, a new software package that measures trigger efficiencies.

In addition to the physics improvements of PIDCalib, there were several technical improvements which resulted in a factor of three speed increase and a reduction in memory usage by 50 %. Many new user-facing scripts were also developed for the PIDCalib package, which simplifies its use and performs common tasks needed in many analyses, such as creating binning schemes, making efficiency plots and plotting distributions from both the calibration and reference samples. These improvements to PIDCalib will allow analyses to more quickly and easily evaluate PID and trigger efficiencies.

7.3 The search for $D^0 \to e^\pm \mu^\mp$

The lepton flavour violating decay $D^0 \to e^\pm \mu^\mp$ is searched for using D^0 decays from $D^{*+} \to D^0 \pi^+$, which reduces the combinatorial background. In order to remove the uncertainty from the measurement of the luminosity, the measurement is performed normalised to the number of $D^0 \to K^- \pi^+$ decays. Candidate $D^0 \to e^\pm \mu^\mp$ decays are classified into three subsamples based on the output of a BDT. A log-likelihood fit is performed as a function of the D^0 candidate mass distribution and the distribution of the difference between the D^* and D^0 candidate masses. This fit is performed simultaneously in the three subsamples of BDT output. No evidence was seen for $D^0 \to e^\pm \mu^\mp$ decays in any bin of BDT output or in the overall mass spectra. The fit gave a total of -7 ± 15 signal events, which results in a measured branching fraction of $\mathcal{B}\left(D^0 \to e^\pm \mu^\mp\right) = (-0.6 \pm 1.2) \times 10^{-8}$; this is consistent

with the background-only hypothesis and so an upper limit is placed on $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$.

The limit produced from this analysis can be used to constrain new physics models. This method is not yet sensitive enough to constrain new physics models involving FCNC at tree-level or theories involving additional neutral leptons, but in the RPV-MSSM this limit is able to produce the most stringent constraints on the products of the couplings that are involved in this decay [33,42,43]. Further, the new limit on $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$ will also allow constraints on certain new physics models involving leptoquarks [134]. An upper limit of $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right) < 1.3 \times 10^{-8}$ at a 90 % CL [116] is placed on the branching fraction; this is an order of magnitude lower than the previous limit from the Belle collaboration [36] and it is the world's best measurement of this quantity.

7.4 B_s^0 and B_d^0 mixing frequencies

The mixing frequencies of the B_q^0-mesons have been measured using flavour-specific semileptonic decays of the type $B_q^0 \to D_q^- \mu^+ \nu$. Daughter particles of the B_q^0 mesons which were missing or ignored in the selection reduce the reconstructed momentum of the B_q^0 candidate and so a simulation-based statistical correction is used to correct the reconstructed decay-time. Two methods to extract the mixing frequencies are then presented: a multivariate fit and a Fourier analysis. This is the first time that a Fourier analysis has been applied to measure the B_q^0 mixing frequencies. Although both methods produced consistent results, the first method is more precise.

The B_q^0 mixing frequencies are found to be

$$\Delta m_d = (0.503 \pm 0.011\,(\text{stat}) \pm 0.013\,(\text{syst}))\,\text{ps}^{-1}\,,$$
$$\Delta m_s = (17.93 \pm 0.22\,(\text{stat}) \pm 0.15\,(\text{syst}))\,\text{ps}^{-1}\,.$$

This measurement is sensitive to new particles with masses much larger than the energy of the accelerator, as they can enter into the loop as virtual particles and affect the oscillation frequency. No deviation is seen from the SM predictions or from the previous measurements performed with fully-reconstructed decays [119, 135]. The hypotheses of no oscillations is rejected by $13.0\,\sigma$ for B_d^0 oscillations and $5.8\,\sigma$ for B_s^0 oscillations. This is the first observation of B_s^0-$\bar{B}_s^0$ mixing to be made using only semileptonic decays [133].

APPENDIX A

VELO module positions

In the following section the position of all modules are listed for reference.Figure A.7 compares all of the different VELO geometries.

A.1 Layout 1: Geometric layout

Table A.1 and Fig. A.1 shows the positions of 42 modules called “Layout 1,” which is derived from simple geometric considerations.

Table A.1: Layout 1 is calculated from geometric considerations. A side is also known as the left side and is in the positive x-direction. C side is also known as the right side and is in the negative x-direction. The z-positions are given in mm from the interaction point.

Side	Module z-position [mm]						
A	-100.50	-75.50	-50.50	-25.50	-0.50	24.50	49.50
	74.50	99.50	124.50	149.50	174.50	199.50	224.50
	249.50	274.50	429.25	604.00	629.00	654.00	679.00
C	-88.50	-63.50	-38.50	-13.50	11.50	36.50	61.50
	86.50	111.50	136.50	161.50	186.50	211.50	236.50
	261.50	286.50	441.25	616.00	641.00	666.00	691.00
Mean	-94.50	-69.50	-44.50	-19.50	5.50	30.50	55.50
	80.50	105.50	130.50	155.50	180.50	205.50	230.50
	255.50	280.50	435.25	610.00	635.00	660.00	685.00

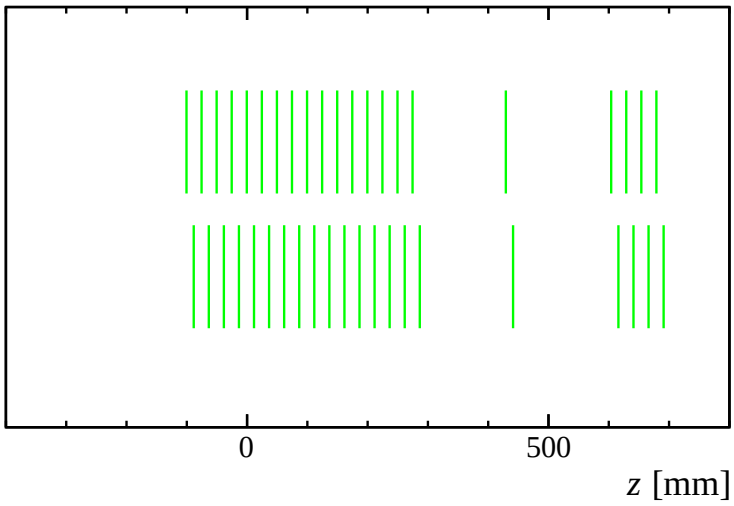

Figure A.1: Layout calculated from only simple geometric considerations.

A.2 Layout 2: Ray-tracing optimised

Table A.2 and Fig. A.2 shows the positions of 48 modules called “Layout 2,” which is optimised using the ray-tracing simulation.

Table A.2: Layout 2 is optimised using only the ray-tracing simulation. A side is also known as the left side and is in the positive x-direction. C side is also known as the right side and is in the negative x-direction. The z-positions are given in mm from the interaction point.

Side	Module z-position [mm]						
A	-288.50	-263.50	-238.50	-213.50	-143.75	-74.00	-49.00
	-24.00	1.00	26.00	51.00	76.00	101.00	126.00
	151.00	176.00	201.00	226.00	251.00	427.50	604.00
	629.00	654.00	679.00				
C	-276.50	-251.50	-226.50	-201.50	-131.75	-62.00	-37.00
	-12.00	13.00	38.00	63.00	88.00	113.00	138.00
	163.00	188.00	213.00	238.00	263.00	439.50	616.00
	641.00	666.00	691.00				
Mean	-282.50	-257.50	-232.50	-207.50	-137.75	-68.00	-43.00
	-18.00	7.00	32.00	57.00	82.00	107.00	132.00
	157.00	182.00	207.00	232.00	257.00	433.50	610.00
	635.00	660.00	685.00				

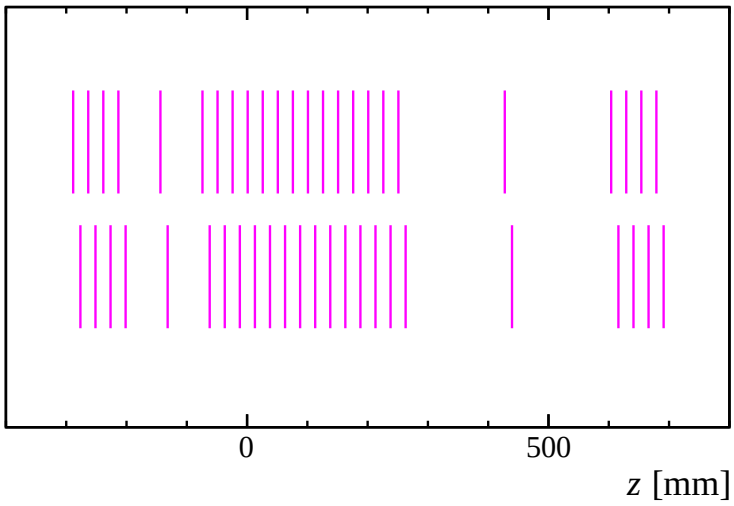

Figure A.2: Layout optimised using only the ray-tracing simulation.

A.3 Layout 3: Full simulation optimised

Table A.3 and Fig. A.3 shows the positions of 52 modules called “Layout 3,” which is optimised using the full LHCb simulation.

Table A.3: Layout 3 is optimised using the full LHCb simulation. A side is also known as the left side and is in the positive x-direction. C side is also known as the right side and is in the negative x-direction. The z-positions are given in mm from the interaction point.

Side	Module z-position [mm]						
A	-288.50	-263.50	-238.50	-213.50	-143.75	-74.00	-49.00
	-24.00	1.00	26.00	51.00	76.00	101.00	126.00
	151.00	176.00	201.00	226.00	251.00	312.99	389.94
	485.45	604.00	649.00	694.00	739.00		
C	-276.50	-251.50	-226.50	-201.50	-131.75	-62.00	-37.00
	-12.00	13.00	38.00	63.00	88.00	113.00	138.00
	163.00	188.00	213.00	238.00	263.00	324.99	401.94
	497.45	616.00	661.00	706.00	751.00		
Mean	-282.50	-257.50	-232.50	-207.50	-137.75	-68.00	-43.00
	-18.00	7.00	32.00	57.00	82.00	107.00	132.00
	157.00	182.00	207.00	232.00	257.00	318.99	395.94
	491.45	610.00	655.00	700.00	745.00		

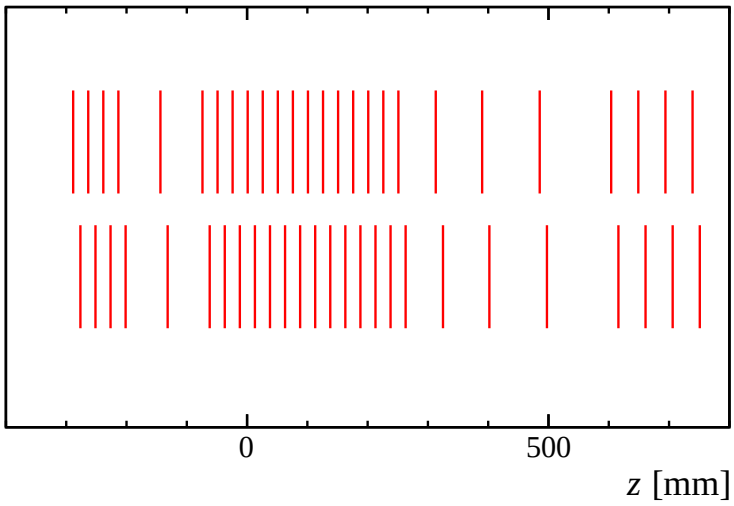

Figure A.3: Layout optimised using the full LHCb simulation.

A.4 Layout 4: Regularised layout

Table A.4 and Fig. A.4 shows the positions of 52 modules called “Layout 4,” which is optimised using the full LHCb simulation, then moved to fit on a regular spacing. The positions of the modules in the A and C sides have been swapped between Layout 3 and Layout 4; this effectively reflects the module positions $x \to -x$. This can be seen in the last two layouts of Fig. A.7.

Table A.4: Layout 4 is fully optimised then regularised. A side is also known as the left side and is in the positive x-direction. C side is also known as the right side and is in the negative x-direction. The z-positions are given in mm from the interaction point.

Side	Module z-position [mm]						
A	-275.00	-250.00	-225.00	-200.00	-125.00	-50.00	-25.00
	0.00	25.00	50.00	75.00	100.00	125.00	150.00
	175.00	200.00	225.00	250.00	275.00	325.00	400.00
	500.00	600.00	650.00	700.00	750.00		
C	-287.50	-262.50	-237.50	-212.50	-137.50	-62.50	-37.50
	-12.50	12.50	37.50	62.50	87.50	112.50	137.50
	162.50	187.50	212.50	237.50	262.50	312.50	387.50
	487.50	587.50	637.50	687.50	737.50		
Mean	-281.25	-256.25	-231.25	-206.25	-131.25	-56.25	-31.25
	-6.25	18.75	43.75	68.75	93.75	118.75	143.75
	168.75	193.75	218.75	243.75	268.75	318.75	393.75
	493.75	593.75	643.75	693.75	743.75		

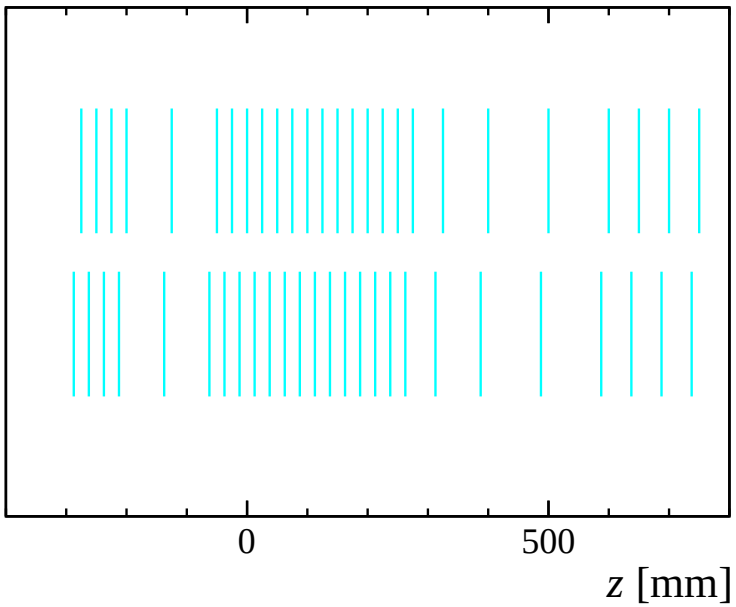

Figure A.4: Layout optimised using full simulation, then moved to fit on regular spacing.

A.5 Layout 5: Current VELO

Table A.5 and Fig. A.5 shows the positions of the 46 modules in the current VELO detector as described in [58].

Table A.5: Layout used in current strip-based VELO. A side is also known as the left side and is in the positive x-direction. C side is also known as the right side and is in the negative x-direction. The z-positions are given in mm from the interaction point.

Side	Module z-position [mm]						
A	-315.00	-235.00	-175.00	-145.00	-115.00	-85.00	-55.00
	-25.00	5.00	35.00	65.00	95.00	125.00	155.00
	185.00	215.00	245.00	275.00	435.00	585.00	635.00
	685.00	735.00					
C	-300.00	-220.00	-160.00	-130.00	-100.00	-70.00	-40.00
	-10.00	20.00	50.00	80.00	110.00	140.00	170.00
	200.00	230.00	260.00	290.00	450.00	600.00	650.00
	700.00	750.00					
Mean	-307.50	-227.50	-167.50	-137.50	-107.50	-77.50	-47.50
	-17.50	12.50	42.50	72.50	102.50	132.50	162.50
	192.50	222.50	252.50	282.50	442.50	592.50	642.50
	692.50	742.50					

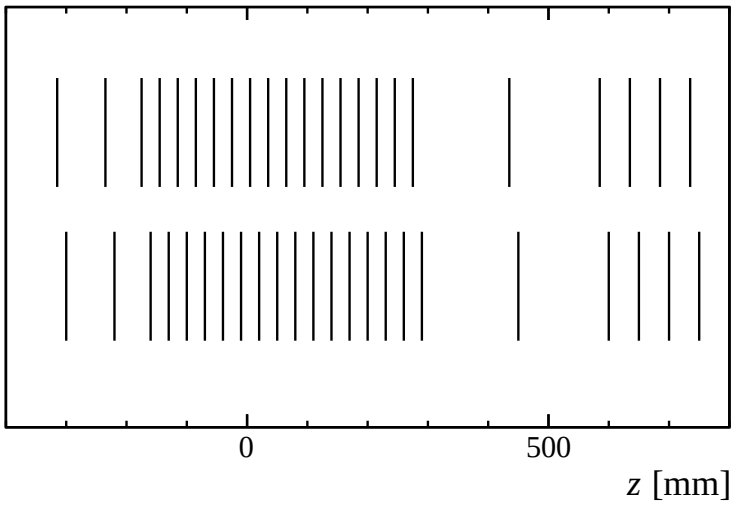

Figure A.5: Layout used in current strip-based VELO.

A.6 Layout 6: Letter of intent

Table A.6 and Fig. A.6 shows the positions of the 48 modules as proposed in the "Letter of Intent (LoI) for the LHCb Upgrade" [85].

Table A.6: Layout proposed in the LHCb upgrade LoI. A side is also known as the left side and is in the positive x-direction. C side is also known as the right side and is in the negative x-direction. The z-positions are given in mm from the interaction point.

Side	Module z-position [mm]						
A	-165.00	-141.00	-117.00	-93.00	-69.00	-45.00	-21.00
	3.00	27.00	51.00	75.00	99.00	123.00	147.00
	171.00	195.00	240.00	298.00	373.00	458.00	588.00
	638.00	688.00	738.00				
C	-153.00	-129.00	-105.00	-81.00	-57.00	-33.00	-9.00
	15.00	39.00	63.00	87.00	111.00	135.00	159.00
	183.00	207.00	252.00	310.00	385.00	470.00	600.00
	650.00	700.00	750.00				
Mean	-159.00	-135.00	-111.00	-87.00	-63.00	-39.00	-15.00
	9.00	33.00	57.00	81.00	105.00	129.00	153.00
	177.00	201.00	246.00	304.00	379.00	464.00	594.00
	644.00	694.00	744.00				

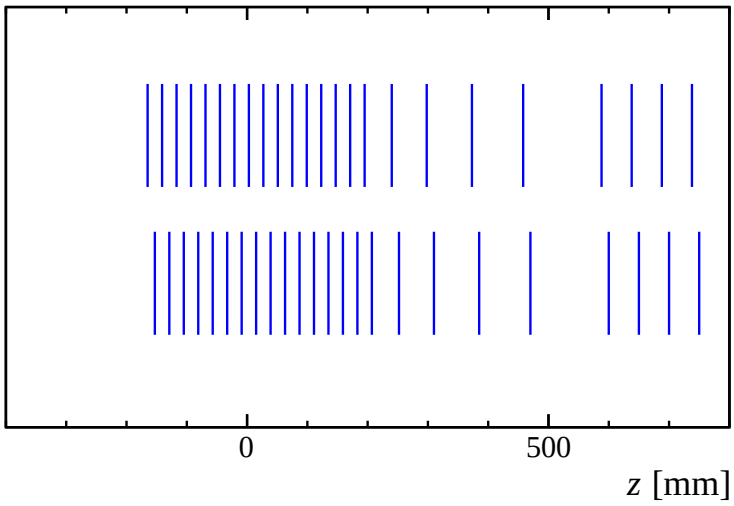

Figure A.6: Layout proposed in the letter of intent for the LHCb upgrade.

A.7 Layout comparison

Finally Fig. A.7 compares the positions of all layouts.

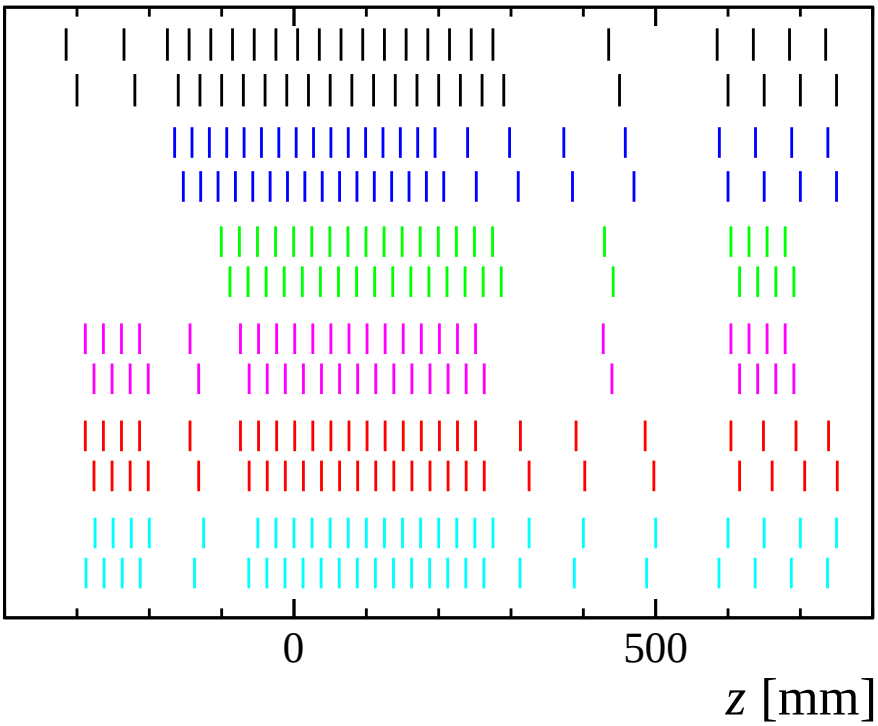

Figure A.7: The figure shows a comparison of VELO *z*-layouts. The first (top) layout is the current VELO (black). The second layout is the upgrade VELO layout proposed in the LHCb letter of intent (blue). The third layout is Layout 1, which is calculated from geometry considerations (green). The fourth layout is Layout 2 and is optimised with the ray-tracing simulation (magenta). The fifth layout is Layout 3 and is optimised using the full-simulation (red). The sixth (bottom) layout is Layout 4 and is a simplified version of Layout 3 (cyan). All module positions are listed in Appendix A.

APPENDIX B

Charm selection efficiencies

This appendix contains the selection efficiencies determined for the $D^0 \to e^{\pm}\mu^{\mp}$ analysis (Chapter 5). This includes the signal $D^0 \to e^{\pm}\mu^{\mp}$ selection efficiency, the misidentification probability of $D^0 \to \pi^-\pi^+$ and the $D^0 \to K^-\pi^+$ selection efficiency, for both the 7 TeV and 8 TeV datasets. It also separates the systematic uncertainties into their constituent parts.

B.1 Signal selection efficiency

Tables B.1 and B.2 show the constituent parts of the selection efficiency for $D^0 \to e^{\pm}\mu^{\mp}$ events and further breaks down the contributions to the systematic uncertainties. The tables are split into two sections; the upper half shows selection efficiencies determined from the simulation, while the lower half shows the selection efficiencies determined with PIDCalib. The column labelled efficiency shows the selection efficiency of the given

requirement and it's statistical uncertainty. The multiplicity, tracking, mass and total columns in the upper sections of the tables show the systematic uncertainties from: weighting the efficiency in three bins of multiplicity, the tracking group (around 2.5 %), the difference in mass reconstructed in the trigger between data and simulation and the final column is the total systematic uncertainty. In the lower half of the tables the columns binning, scaling, fitting and total show the systematic uncertainties from: the binning scheme used, scaling the number of tracks, the fit used to $_s$weight the dataset and the final column is the total systematic uncertainty. The last two lines show the fraction of simulated events which are saved and the overall efficiency; the overall efficiency contains both statistical and systematic uncertainties.

Table B.1: The table summarises the efficiencies and systematic uncertainties which are applied when measuring the efficiency of a $D^0 \to e^{\pm}\mu^{\mp}$ decay to be selected from 2011 data. See Appendix B.1 for a description of the columns.

		Measured with MC			
Cut	Efficiency	Multiplicity	Tracking	Mass	Total
Stripping	$(6.452 \pm 0.027)\%$	0.260%	0.162%	–	0.306%
DTF converged	$(99.9909 \pm 0.0045)\%$	0.0025%	–	–	0.0025%
HLT2 1815-1915	$(28.60 \pm 0.32)\%$	0.31%	–	–	0.31%
Mass 1815-1915	$(86.66 \pm 0.48)\%$	0.46%	–	1.01%	1.11%
Total	$(1.599 \pm 0.021 \pm 0.080)\%$				
		Measured with PIDCalib			
Cut	Efficiency	Binning	Scaling	Fitting	Total
L0	$(64.287 \pm 0.089)\%$	0.594%	0.399%	0.100%	0.722%
HLT1	$(78.332 \pm 0.041)\%$	0.401%	0.434%	0.100%	0.599%
Loose PID	$(98.346 \pm 0.021)\%$	0.428%	0.024%	0.100%	0.441%
Offline PID	$(56.098 \pm 0.059)\%$	0.090%	0.512%	0.100%	0.530%
Ghost cut	$(80.543 \pm 0.027)\%$	0.593%	0.441%	0.100%	0.746%
Total	$(22.376 \pm 0.042 \pm 0.436)\%$				
Simulated fraction saved		$(11.971 \pm 0.028)\%$			
Grand total (inc. syst.)		$(0.0428 \pm 0.0024)\%$			

Table B.2: The table summarises the efficiencies and systematic uncertainties which are applied when measuring the efficiency of a $D^0 \to e^{\pm}\mu^{\mp}$ decay to be selected from 2012 data. See Appendix B.1 for a description of the columns.

		Measured with MC			
Cut	Efficiency	Multiplicity	Tracking	Mass	Total
Stripping	$(5.822 \pm 0.016)\%$	0.259%	0.146%	–	0.297%
DTF converged	$(99.9927 \pm 0.0027)\%$	0.0020%	–	–	0.0020%
HLT2 1815-1915	$(33.24 \pm 0.22)\%$	0.38%	–	–	0.38%
Mass 1815-1915	$(86.16 \pm 0.29)\%$	0.26%	–	0.52%	0.58%
Total	$(1.667 \pm 0.013 \pm 0.088)\%$				
		Measured with PIDCalib			
Cut	Efficiency	Binning	Scaling	Fitting	Total
L0	$(59.059 \pm 0.059)\%$	0.528%	0.309%	0.100%	0.620%
HLT1	$(87.256 \pm 0.016)\%$	0.350%	0.139%	0.100%	0.390%
Loose PID	$(97.983 \pm 0.011)\%$	0.153%	0.027%	0.100%	0.185%
Offline PID	$(53.520 \pm 0.032)\%$	0.200%	0.534%	0.100%	0.578%
Ghost cut	$(81.539 \pm 0.014)\%$	0.421%	0.425%	0.100%	0.606%
Total	$(22.035 \pm 0.027 \pm 0.385)\%$				
Simulated fraction saved		$(12.232 \pm 0.028)\%$			
Grand total (inc. syst.)		$(0.0449 \pm 0.0025)\%$			

B.2 Misidentification probability

Tables B.3 and B.4 show the constituent parts of the misidentification probability for $D^0 \to \pi^-\pi^+$ events and further breaks down the contributions to the systematic uncertainties. The tables contain the same information as those in Appendix B.1.

Table B.3: The table summarises the efficiencies and systematic uncertainties which are applied when measuring the efficiency of a $D^0 \to \pi^-\pi^+$ decay to be selected as $D^0 \to e^\pm\mu^\mp$ in the 2011 data. See Appendix B.1 for a description of the columns.

		Measured with MC			
Cut	Efficiency	Multiplicity	Tracking	Mass	Total
Stripping	$(4.665 \pm 0.033)\%$	0.062%	0.118%	–	0.133%
DTF converged	$(100 \pm 0)\%$	0.0%	–	–	0.0%
HLT2 1815-1915	$(32 \pm 13)\%$	10.09%	–	–	10.09%
Mass 1815-1915	$(96.3 \pm 3.5)\%$	1.7%	–	0.0%	1.7%
Total	$(1.45 \pm 0.58 \pm 0.46)\%$				
		Measured with PIDCalib			
Cut	Efficiency	Binning	Scaling	Fitting	Total
L0	$(1.0754 \pm 0.0055)\%$	0.0022%	0.0251%	0.1000%	0.1031%
HLT1	$(60.73 \pm 0.24)\%$	0.43%	0.48%	0.10%	0.65%
Loose PID	$(8.682 \pm 0.057)\%$	0.106%	0.029%	0.100%	0.148%
Offline PID	$(0.7584 \pm 0.0068)\%$	0.0330%	0.0146%	0.1000%	0.1063%
Ghost cut	$(81.25 \pm 0.20)\%$	0.43%	0.45%	0.10%	0.63%
Total	$(3.494 \pm 0.046 \pm 0.598) \times 10^{-6}$				
Simulated fraction saved		$(20.205 \pm 0.026)\%$			
Grand total (inc. syst.)		$(1.02 \pm 0.55) \times 10^{-8}$			

Table B.4: The table summarises the efficiencies and systematic uncertainties which are applied when measuring the efficiency of a $D^0 \to \pi^-\pi^+$ decay to be selected as $D^0 \to e^\pm\mu^\mp$ in the 2012 data. See Appendix B.1 for a description of the columns.

		Measured with MC			
Cut	Efficiency	Multiplicity	Tracking	Mass	Total
Stripping	$(4.229 \pm 0.020)\%$	0.099%	0.107%	–	0.146%
DTF converged	$(99.9896 \pm 0.0053)\%$	0.0024%	–	–	0.0024%
HLT2 1815-1915	$(75.8 \pm 5.3)\%$	2.5%	–	–	2.5%
Mass 1815-1915	$(99.72 \pm 0.32)\%$	0.31%	–	0.37%	0.48%
Total	$(3.20 \pm 0.22 \pm 0.15)\%$				
		Measured with PIDCalib			
Cut	Efficiency	Binning	Scaling	Fitting	Total
L0	$(0.9747 \pm 0.0033)\%$	0.0039%	0.0208%	0.1000%	0.1022%
HLT1	$(70.378 \pm 0.064)\%$	0.144%	0.186%	0.100%	0.256%
Loose PID	$(8.227 \pm 0.030)\%$	0.024%	0.035%	0.100%	0.109%
Offline PID	$(0.5947 \pm 0.0035)\%$	0.0087%	0.0088%	0.1000%	0.1008%
Ghost cut	$(82.124 \pm 0.078)\%$	0.242%	0.394%	0.100%	0.473%
Total	$(2.756 \pm 0.021 \pm 0.551) \times 10^{-9}$				
Simulated fraction saved		$(20.27 \pm 0.026)\%$			
Grand total (inc. syst.)		$(1.79 \pm 0.39) \times 10^{-8}$			

B.3 Normalisation channel selection efficiency

Tables B.5 and B.6 show the constituent parts of the selection efficiency for $D^0 \to K^-\pi^+$ events and further breaks down the contributions to the systematic uncertainties. These tables contain similar information to the tables in Appendix B.1, but the PIDCalib section of Tables B.5 and B.6 is separated into several parts. The $D^0 \to K^-\pi^+$ events can be triggered by either the $K^\pm$ or the $\pi^\mp$ and so three categories of efficiencies are measured: where both the $K^\pm$ and $\pi^\mp$ are TOS, where only the $K^\pm$ is TOS and where only the $\pi^\pm$ is TOS. Since the ghost cut is applied to the π from the D^*, there is no need to include it in these three categories. Due to an overwhelming number of $D^0 \to K^-\pi^+$ decays only 1 in 100 are saved for analysis, this is known as a pre-scale and it is listed in these tables.

Table B.5: The table summarises all systematics which are applied when measuring the efficiency of a $D^0 \to K^-\pi^+$ decay to be selected from the 2011 data. See Appendices B.1 and B.3 for a description of the columns.

		Measured with MC		
Cut	Efficiency	Multiplicity	Tracking	Total
Stripping	$(4.280 \pm 0.017)\%$	0.066%	0.098%	0.118%
DTF converged	$(99.9953 \pm 0.0033)\%$	0.0017%	–	0.0017%
HLT2 trigger	$(92.61 \pm 0.37)\%$	0.15%	–	0.15%
Mass 1815-1915	$(100 \pm 0)\%$	0.0%	–	0.0%
Total	$(3.964 \pm 0.022 \pm 0.109)\%$			

		Measured with PIDCalib			
Cut	Efficiency	Binning	Scaling	Fitting	Total
L0, both TOS	$(1.1060 \pm 0.0076)\%$	0.0066%	0.0166%	0.1000%	0.1016%
HLT1, both TOS	$(27.919 \pm 0.021)\%$	0.435%	0.225%	0.100%	0.500%
PID, both TOS	$(54.951 \pm 0.059)\%$	0.849%	0.289%	0.100%	0.902%
Total	$(0.1697 \pm 0.0012 \pm 0.0161)\%$				
L0, $K^\pm$ TOS	$(8.780 \pm 0.024)\%$	0.097%	0.069%	0.100%	0.155%
HLT1, $K^\pm$ TOS	$(36.205 \pm 0.035)\%$	0.233%	0.054%	0.100%	0.259%
PID, $K^\pm$ TOS	$(58.708 \pm 0.053)\%$	0.987%	0.359%	0.100%	1.055%
Total	$(1.8661 \pm 0.0057 \pm 0.0489)\%$				
L0, $\pi^\pm$ TOS	$(8.001 \pm 0.023)\%$	0.053%	0.064%	0.100%	0.130%
HLT1, $\pi^\pm$ TOS	$(34.022 \pm 0.035)\%$	0.306%	0.047%	0.100%	0.325%
PID, $\pi^\pm$ TOS	$(75.557 \pm 0.040)\%$	0.184%	0.272%	0.100%	0.343%
Total	$(2.0568 \pm 0.0063 \pm 0.0399)\%$				
Sum over triggers	$(4.093 \pm 0.066)\%$				
Ghost cut	$(79.253 \pm 0.014)\%$	0.590%	0.428%	0.100%	0.736%
Total	$(79.253 \pm 0.014 \pm 0.736)\%$				
Trigger pre-scale		0.01			
Simulated fraction saved		$(21.415 \pm 0.027)\%$			
Grand total (inc. syst.)		$(2.754 \pm 0.093) \times 10^{-6}$			

Table B.6: The table summarises all systematics which are applied when measuring the efficiency of a $D^0 \to K^-\pi^+$ decay to be selected from the 2012 data. See Appendices B.1 and B.3 for a description of the columns.

		Measured with MC		
Cut	Efficiency	Multiplicity	Tracking	Total
Stripping	$(3.925 \pm 0.010)\%$	0.076%	0.090%	0.118%
DTF converged	$(99.9975 \pm 0.0014)\%$	0.0002%	–	0.0002%
HLT2 trigger	$(92.77 \pm 0.25)\%$	0.18%	–	0.18%
Mass 1815-1915	$(99.9982 \pm 0.0048)\%$	0.0005%	–	0.0005%
Total	$(3.641 \pm 0.014 \pm 0.110)\%$			

		Measured with PIDCalib			
Cut	Efficiency	Binning	Scaling	Fitting	Total
L0, both TOS	$(1.1353 \pm 0.0056)\%$	0.0116%	0.0169%	0.1000%	0.1021%
HLT1, both TOS	$(28.618 \pm 0.015)\%$	0.435%	0.095%	0.100%	0.456%
PID, both TOS	$(51.234 \pm 0.044)\%$	1.016%	0.252%	0.100%	1.052%
Total	$(0.16646 \pm 0.00084 \pm 0.01558)\%$				
L0, $K^\pm$ TOS	$(8.820 \pm 0.018)\%$	0.086%	0.068%	0.100%	0.149%
HLT1, $K^\pm$ TOS	$(35.231 \pm 0.026)\%$	0.302%	0.024%	0.100%	0.319%
PID, $K^\pm$ TOS	$(56.661 \pm 0.041)\%$	1.192%	0.331%	0.100%	1.241%
Total	$(1.7607 \pm 0.0040 \pm 0.0512)\%$				
L0, $\pi^\pm$ TOS	$(8.122 \pm 0.017)\%$	0.039%	0.067%	0.100%	0.126%
HLT1, $\pi^\pm$ TOS	$(33.036 \pm 0.025)\%$	0.303%	0.028%	0.100%	0.320%
PID, $\pi^\pm$ TOS	$(73.350 \pm 0.031)\%$	0.274%	0.290%	0.100%	0.412%
Total	$(1.9681 \pm 0.0044 \pm 0.0377)\%$				
Sum over triggers	$(3.895 \pm 0.066)\%$				
Ghost cut	$(80.2410 \pm 0.0090)\%$	0.5382%	0.4081%	0.1000%	0.6828%
Total	$(80.2410 \pm 0.0090 \pm 0.6828)\%$				
Trigger pre-scale		0.00995819			
Simulated fraction saved		$(21.49 \pm 0.027)\%$			
Grand total (inc. syst.)		$(2.435 \pm 0.087) \times 10^{-6}$			

APPENDIX C

Simulated B^0, B_s^0 and B^+ datasets

The mixing analysis involved many simulation-based studies, where the 2010 Monte Carlo (MC10) was used. This was generated with 2010 LHC beam conditions, which means that $\sqrt{s}$=7 TeV and there were a mean of 2.5 interactions per bunch crossing (1.75 visible interactions). The simulations used a mixture of different decay modes, the branching fractions of which are not well known experimentally. Tables C.1 to C.3 show the decay modes input to the simulation from theoretical predictions. In these tables the branching fractions $D_s^{\pm} \to K^+K^-\pi^{\pm} = (5.49 \pm 0.27)\,\%$ and $D^{\pm} \to K^+K^-\pi^{\pm} = (9.54 \pm 0.26) \times 10^{-3}$ have been included. The sum of the branching fractions to each of these decay modes is the visible branching fraction ($\mathcal{B}_{\text{vis}}$) and

Table C.1: Proportion of events in B_s^0 cocktail samples (event-type code 13774002).

Decay mode	$\mathcal{B} \times 10^3$	Proportion
$B_s^0 \to D_s^- \mu^+ \nu_\mu$	1.07	0.234
$B_s^0 \to D_s^{*-} \mu^+ \nu_\mu \to D_s^- \mu^+ \nu_\mu X^0$	2.50	0.545
$B_s^0 \to D_{s0}^{*-} \mu^+ \nu_\mu \to D_s^- \mu^+ \nu_\mu X^0$	0.20	0.045
$B_s^0 \to D_{s1}^- \mu^+ \nu_\mu \to D_s^- \mu^+ \nu_\mu X^0$	0.20	0.045
$B_s^0 \to D_{s1}^{'-} \mu^+ \nu_\mu \to D_s^- \mu^+ \nu_\mu X^0$	0.35	0.078
$B_s^0 \to D_s^- \tau^+ \nu_\tau \to D_s^- \mu^+ \nu_\mu X^0$	0.070	0.015
$B_s^0 \to D_s^{*-} \tau^+ \nu_\tau \to D_s^- \mu^+ \nu_\mu X^0$	0.140	0.031
$B_s^0 \to D_{s1}^{'-} \tau^+ \nu_\tau \to D_s^- \mu^+ \nu_\mu X^0$	0.016	0.0035
$B_s^0 \to D_{s0}^{*-} \tau^+ \nu_\tau \to D_s^- \mu^+ \nu_\mu X^0$	0.016	0.0035
Total	4.58	1.00

these are:

$$\begin{aligned} \mathcal{B}_{\text{vis}}(B_s^0) &= (4.6 \pm 1.4) \times 10^{-3}, \\ \mathcal{B}_{\text{vis}}(B^0) &= (5.86 \pm 0.60) \times 10^{-4}, \\ \mathcal{B}_{\text{vis}}(B^+) &= (8.8 \pm 1.1) \times 10^{-5}. \end{aligned} \tag{C.1}$$

In addition, to study partially-reconstructed indistinguishable peaking backgrounds of the type $X_b \to (X_c \to \mu X) D_s^{(*)\pm} X$, a generator-level selection on inclusive-b Monte Carlo events is used to create a specific sample (event type code 10074000).

Table C.4 shows the efficiency of the generator level cuts, along with the number of events passing both the stripping and full selection. Using the information in Eq. (C.1) and Table C.4, the ratio of expected B^+ to B^0 events

Table C.2: Proportion of events in $B_u^\pm$ cocktail samples (event-type code 12674011).

Decay mode	$\mathcal{B} \times 10^3$	Proportion
$B^+ \to D^- \pi^+ \mu^+ \nu_\mu$	0.0089	0.102
$B^+ \to D^{*-} \pi^+ \mu^+ \nu_\mu$	0.0070	0.080
$B^+ \to D_0^{*0} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X$	0.0231	0.264
$B^+ \to D_1^0 \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X$	0.0215	0.245
$B^+ \to D_1^{\prime 0} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X$	0.0074	0.084
$B^+ \to D_2^{*0} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X$	0.0197	0.224
Total	0.0876	1.00

in the final selected dataset can be calculated to be,

$$\frac{N_{B^+}}{N_{B_s^0}} = \frac{f_u \varepsilon_{B^+} \mathcal{B}_{\text{vis}}(B^0)}{f_d \varepsilon_{B^0} \mathcal{B}_{\text{vis}}(B^+)}, \tag{C.2}$$

where N_{B^+} (N_{B^0}) is the number of selected B^+ (B^0) events, f_u (f_d) is the fraction of b-quarks which hadronise to B^+ (B^0) and ε_{B^+} (ε_{B^0}) is the full selection efficiency of simulated B^+ (B^0) decays. The ratio of $N_{B^+}/N_{B_s^0}$ is then calculated to be 11 ± 2 %.

Table C.3: Proportion of events in B_d^0 cocktail samples (event-type code 11874021).

Decay mode	$\mathcal{B} \times 10^3$	Proportion
$B_d^0 \to D^- \mu^+ \nu_\mu$	0.273	0.466
$B_d^0 \to D^{*-} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X^0$	0.204	0.348
$B_d^0 \to D_1^- \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X^0$	0.00233	0.040
$B_d^0 \to D_0^{*-} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X^0$	0.0164	0.028
$B_d^0 \to D_1^{'-} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X^0$	0.011	0.017
$B_d^0 \to D_2^{*-} \mu^+ \nu_\mu \to D^- \mu^+ \nu_\mu X^0$	0.0022	0.037
$B_d^0 \to D^- \mu^+ \pi^0 \nu_\mu$	0.011	0.019
$B_d^0 \to D^{*-} \mu^+ \pi^0 \nu_\mu \to D^- \mu^+ \nu_\mu X^0$	0.0085	0.015
$B_d^0 \to D^- \tau^+ \nu_\tau \to D^- \mu^+ \nu_\mu X^0$	0.0022	0.0037
$B_d^0 \to D^{*-} \tau^+ \nu_\tau \to D^- \mu^+ \nu_\mu X^0$	0.0011	0.019
$B_d^0 \to D_1^- \tau^+ \nu_\tau \to D^- \mu^+ \nu_\mu X^0$	0.00071	0.0013
$B_d^0 \to D_0^{*-} \tau^+ \nu_\tau \to D^- \mu^+ \nu_\mu X^0$	0.00078	0.0013
$B_d^0 \to D_1^{'-} \tau^+ \nu_\tau \to D^- \mu^+ \nu_\mu X^0$	0.0013	0.0023
$B_d^0 \to D_2^{*-} \tau^+ \nu_\tau \to D^- \mu^+ \nu_\mu X^0$	0.0014	0.0022
Total	0.586	1.00

Table C.4: Details about the selection efficiency for the simulated events used in this study.

Decay descriptor	Event type code	Generation cut efficiency			Events simulated	Events stripped	Candidates stripped	Candidates fully-selected	Luminosity[a] equiv. / pb^{-1}
$B_s^0 \to D_s^{(*)-}\mu^+\nu_\mu X$	13774002	0.1693	±	0.0014	5049943	214103	222685	112660	~ 110
$B_d^0 \to D_d^{(*)-}\mu^+\nu_\mu X$	11874021	0.1708	±	0.0022	5034246	244380	254794	128864	~ 220
$B^+ \to D_d^{(*)-}\mu^+\pi^+\nu_\mu X$	12674011	0.1671	±	0.0021	5039967	237834	257744	98831	~ 1500
$X_b \to D_s^{(*)\pm}(X_c \to \mu X)\ldots$	10074000	0.00144578	±	0.00000062	9979397	50052	58334	4366	~ 12

[a] Calculated with a $b\bar{b}$ cross-section of $(284 \pm 20 \pm 49)\,\mu$b from Ref. [136], and the branching ratios used as input to the simulation.

List of Abbreviations

ALICE A Large Ion Collider Experiment

ASIC application-specific integrated circuit

ATLAS A Toroidal LHC Apparatus

BDT boosted decision tree

CERN European Organization for Nuclear Research

CKM Cabbibo-Kobayashi-Maskawa matrix

CALO calorimeter

CB crystal ball function

CL confidence level

CLFV charged lepton-flavour violation

CM centre-of-mass

CMS Compact Muon Solenoid experiment

CPV $\mathcal{CP}$-violation (Charge conjugation and Parity symmetry Violation)

DAQ data acquisition

ΔLL difference in logarithms of likelihoods

DOCA distance of closest approach

DTF decay tree fit

ECAL electromagnetic calorimeter

FCNC flavour-changing neutral currents

GBT gigabit transceiver

GEM gas electron multiplier

GIM Glashow-Iliopoulos-Maiani mechanism

GPD general purpose detector

HCAL hadronic calorimeter

HEP high energy physics

HLT high-level trigger

HPD hybrid photon detector

IP impact parameter

IT Inner Tracker

L0 hardware level-0 trigger

LEP Large Electron–Positron Collider

LFV lepton-flavour violation

LHC Large Hadron Collider

LHCb large hadron collider beauty experiment

LHCf large hadron collider forward experiment

LLT low-level trigger

LS1 Long Shutdown 1 (2013-2015)

LS2 Long Shutdown 2 (2018-2020)

LS3 Long Shutdown 3 (2023-2025)

LS4 Long Shutdown 4 (2028-2030)

MAPMT multianode photomultiplier tube

MC Monte Carlo simulations

MEG Mu to E Gamma experiment

MOEDAL Monopole and Exotics Detector at the LHC

MSSM minimal supersymmetric standard model

MWPC multi wire proportional chamber

NLL negative logarithm of the likelihood

NN neural network

OS opposite-side b-quark, this is the non-signal b-quark

OT outer tracker

PDF probability density function

PID particle identification

PIDCalib particle identification calibration software package

PMNS Pontecorvo-Maki-Nakagawa-Sakata matrix

PMT photomultiplier tube

PS pre-shower detector

PS Proton Synchrotron

PSB Proton Synchrotron Booster

PSI Paul Scherrer Institute

PV primary *pp* interaction vertex

QCD quantum chromodynamics

QED quantum electrodynamics

QGP quark gluon plasma

RF radio frequency

RICH ring imaging cherenkov detector

RMS root mean squared deviation

RPV R-parity violation

RPV-MSSM R-parity violating MSSM

SciFi scintillating fibre tracker

SiPM silicon photomultiplier

SM standard model of particle physics

νSM standard model with non-degenerate massive neutrinos

SPD scintillating pad detector

SPS Super Proton Synchrotron

SS same-side b-quark, this is the signal b-quark

ST silicon tracker

SUSY supersymmetric standard model

SV secondary vertex

TIS trigger independent of signal

TOS triggered on signal

TOTEM total, elastic and diffractive cross-section measurement experiment

TPG thermal pyrolytic graphite

TrigCalib trigger calibration software package

TT tracker turicensis

UT upgrade tracker

VELO vertex locator detector

VL VeloLite detector

VP VeloPix detector

WLS wavelength-shifting fibre

References

[1] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

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

[3] B. Pontecorvo, *Mesonium and anti-mesonium*, Sov. Phys. JETP **6** (1957) 429, [Zh. Eksp. Teor. Fiz. **33** (1957) 549].

[4] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Sov. Phys. JETP **7** (1958) 172, [Zh. Eksp. Teor. Fiz. **34** (1957) 247].

[5] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870.

[6] E. Majorana, *Theory of the Symmetry of Electrons and Positrons*, Nuovo Cim. **14** (1937) 171.

[7] S. Stone, *New physics from flavour*, PoS **ICHEP2012** (2013) 033, arXiv:1212.6374.

[8] C. Elsasser, *Feynman Diagram Library*, http://www.physik.uzh.ch/~che/FeynDiag/index.php, Mar, 2014.

[9] T. Bird, *$\mathcal{CP}$ violation in semi-leptonic B decays at LHCb*, J. Phys. Conf. Ser. **447** (2013) 012021.

[10] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the flavour-specific CP-violating asymmetry a^s_{sl} in B^0_s decays*, Phys. Lett. **B728** (2014) 607, arXiv:1308.1048.

[11] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in $B^0 - \overline{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, arXiv:1409.8586.

[12] P. J. Mohr and D. B. Newell, *Resource letter fc-1: The physics of fundamental constants*, Am. J. Phys. **78** (2010), no. 4 338.

[13] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K^0_L \to \mu\bar{\mu}$, $K^+ \to \pi^+ \nu\overline{\nu}$ and $K^0 \leftrightarrow \overline{K}^0$*, Prog. Theor. Phys. **65** (1981) 297, [Erratum: Prog. Theor. Phys. **65** (1981) 1772].

[14] A. J. Buras, M. Jamin, and P. H. Weisz, *Leading and Next-to-leading QCD Corrections to ϵ Parameter and $B^0 - \overline{B}^0$ Mixing in the Presence of a Heavy Top Quark*, Nucl. Phys. **B347** (1990) 491.

[15] A. Lenz *et al.*, *Anatomy of New Physics in $B - \overline{B}$ mixing*, Phys. Rev. **D83** (2011) 036004, arXiv:1008.1593.

[16] A. J. Buras and R. Fleischer, *Quark mixing, CP violation and rare decays after the top quark discovery*, Adv. Ser. Direct. High Energy Phys. **15** (1998) 65, arXiv:hep-ph/9704376.

[17] A. Lenz and U. Nierste, *Theoretical update of $B^0_s - \overline{B}^0_s$ mixing*, JHEP **06** (2007) 072, arXiv:hep-ph/0612167.

[18] A. Lenz and U. Nierste, *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons*, in *CKM unitarity triangle. Proceedings, 6th International Workshop, CKM 2010, Warwick, UK, September 6-10, 2010*. arXiv:1102.4274.

[19] A. Lenz, *Theoretical update of B-Mixing and Lifetimes*, in *Electroweak Interactions and Unified Theories: Proceedings of the 47th Rencontres de Moriond on Electroweak Interactions and Unified Theories, La Thuile, March 3-10, 2012*. arXiv:1205.1444.

[20] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, arXiv:hep-ph/0406184.

[21] S. Aoki *et al.*, *Review of lattice results concerning low-energy particle physics*, Eur. Phys. J. **C74** (2014) 2890, arXiv:1310.8555.

[22] J. Urban, F. Krauss, C. Hofmann, and G. Soff, $B^0\bar{B}^0$ *mixing involving supersymmetry*, Mod. Phys. Lett. **A12** (1997) 419.

[23] J. Urban, F. Krauss, U. Jentschura, and G. Soff, *Next-to-leading order QCD corrections for the* $B^0\bar{B}^0$*-mixing with an extended Higgs sector*, Nucl. Phys. **B523** (1998) 40, arXiv:hep-ph/9710245.

[24] DELPHI Collaboration, J. Abdallah *et al.*, *Search for* $B_s^0 - \bar{B}_s^0$ *oscillations and a measurement of* $B_d^0 - \bar{B}_d^0$ *oscillations using events with an inclusively reconstructed vertex*, Eur. Phys. J. **C28** (2003) 155, arXiv:hep-ex/0303032.

[25] T. Bird, *Towards Measuring B Mixing in Semileptonic Decays at LHCb*, Master's thesis, Southampton U., 2011, CERN-THESIS-2011-184.

[26] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) 235, arXiv:physics/0503066, [Transp. Theory Statist. Phys.1,186(1971)].

[27] R. H. Bernstein and P. S. Cooper, *Charged Lepton Flavor Violation: An Experimenter's Guide*, Phys. Rept. **532** (2013) 27, arXiv:1307.5787.

[28] MEG Collaboration, J. Adam *et al.*, *New constraint on the existence of the $\mu^+ \to e^+\gamma$ decay*, Phys. Rev. Lett. **110** (2013) 201801, arXiv:1303.0754.

[29] V. A. Baranov *et al.*, *Search for $\mu^+ \to e^+e^+e^-$ decay*, Sov. J. Nucl. Phys. **53** (1991) 802, [Yad. Fiz. **53** (1991) 1302].

[30] SINDRUM II Collaboration, W. H. Bertl *et al.*, *A Search for $\mu - e$ conversion in muonic gold*, Eur. Phys. J. **C47** (2006) 337.

[31] MEG Collaboration, J. Adam *et al.*, *New limit on the lepton-flavour violating decay $\mu^+ \to e^+\gamma$*, Phys. Rev. Lett. **107** (2011) 171801, arXiv:1107.5547.

[32] A. de Gouvea and P. Vogel, *Lepton Flavor and Number Conservation, and Physics Beyond the Standard Model*, Prog. Part. Nucl. Phys. **71** (2013) 75, arXiv:1303.4097.

[33] 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.

[34] BNL E871 Collaboration, D. Ambrose *et al.*, *New limit on muon and electron lepton number violation from $K^0_L \to \mu^\pm e^\mp$ decay*, Phys. Rev. Lett. **81** (1998) 5734, arXiv:hep-ex/9811038.

[35] LHCb Collaboration, R. Aaij *et al.*, *Search for the lepton-flavor violating decays $B_s^0 \to e^\pm\mu^\mp$ and $B^0 \to e^\pm\mu^\mp$*, Phys. Rev. Lett. **111** (2013) 141801, arXiv:1307.4889.

[36] Belle collaboration, M. Petric *et al.*, *Search for leptonic decays of D^0 mesons*, Phys. Rev. **D81** (2010) 091102, arXiv:1003.2345.

[37] L.-B. Jia, M.-G. Zhao, H.-W. Ke, and X.-Q. Li, *Searching for New Physics in $D^0 \to \mu^+\mu^-$, e^+e^-, $\mu^\pm e^\mp$ at BES and/or Super Charm-Tau Factory*, Chin. Phys. **C38** (2014), no. 10 103101, arXiv:1312.7649.

[38] H. Georgi, *Unparticle physics*, Phys. Rev. Lett. **98** (2007) 221601, arXiv:hep-ph/0703260.

[39] V. Barger, C.-W. Chiang, P. Langacker, and H.-S. Lee, *Z′ mediated flavor changing neutral currents in B meson decays*, Phys. Lett. **B580** (2004) 186, arXiv:hep-ph/0310073.

[40] SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257, arXiv:hep-ex/0509008.

[41] MEGA Collaboration, M. L. Brooks *et al.*, *New limit for the family number nonconserving decay $\mu^+ \to e^+\gamma$* , Phys. Rev. Lett. **83** (1999) 1521, arXiv:hep-ex/9905013.

[42] R.-M. Wang *et al.*, *Studying the lepton number and lepton flavor violating $D^0 \to e^\pm\mu^\mp$ and $D^+_{d/s} \to \pi(K)^+e^\pm\mu^\mp$ decays*, Int. J. Mod. Phys. **A29** (2014), no. 29 1450169.

[43] F. Tahir, A. Mir, and S. Mahmood, *Study of pure and semileptonic decays of D_s meson within R-parity violating supersymmetric model*, Chin. Phys. **C38** (2014) 123101.

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

[45] J. Haffner, *The CERN accelerator complex*, OPEN-PHO-ACCEL-2013-056 (2013).

[46] M. Brice, *Aerial View of the CERN taken in 2008*, CERN-MI-0807031 (2008).

[47] *Some LHC milestones*, BUL-NA-2008-132 (2008).

[48] ATLAS Collaboration, *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*, LHC Tech. Proposal, CERN, Geneva, 1994, CERN-LHCC-94-43.

[49] CMS Collaboration, G. L. Bayatian *et al.*, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, Technical Design Report CMS, CERN, Geneva, 2006, CERN-LHCC-2006-001.

[50] ATLAS Collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[51] CMS Collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[52] ALICE Collaboration, *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC*, LHC Tech. Proposal, CERN, Geneva, 1995, CERN-LHCC-95-71.

[53] LHCb Collaboration, *LHCb: letter of intent*, tech. rep., CERN, Geneva, 1995, CERN-LHCC-1995-005.

[54] MoEDAL Collaboration, J. Pinfold *et al.*, *Technical Design Report of the MoEDAL Experiment*, tech. rep., CERN, Geneva, Jun, 2009, CERN-LHCC-2009-006.

[55] LHCf Collaboration, O. Adriani *et al.*, *LHCf experiment: Technical Design Report*, Technical Design Report LHCf, CERN, Geneva, 2006, CERN-LHCC-2006-004.

[56] TOTEM Collaboration, V. Berardi *et al.*, *Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report*, Technical Design Report TOTEM, CERN, Geneva, 2004, CERN-LHCC-2004-002.

[57] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, arXiv:1412.6352.

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

[59] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

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

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

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

[63] M. Kraan, *Mechanical Design Vertex Locator*, `http://www.nikhef.nl/pub/departments/mt/projects/lhcb-vertex/`, Jun, 2012.

[64] B. Verlaat, M. Van Beuzekom, and A. Van Lysebetten, *CO_2 cooling for HEP experiments*, in *proceedings of the Topical Workshop on Electronics for Particle Physics, TWEPP2008, Naxos, Greece*, 2008. doi: 10.5170/CERN-2008-008.328.

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

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

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

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

[69] A. D. Webber, *Radiation damage studies in the LHCb VELO detector and measurement of the flavour-specific asymmetry in semileptonic B-decays*, PhD thesis, Manchester U., 2013, CERN-THESIS-2013-117, 21 Feb 2013.

[70] R. W. Lambert, *LHCb Hybrid Photon Detectors and Sensitivity to Flavour Specific Asymmetry in Neutral B-Meson Mixing*, PhD thesis, Edinburgh U., Edinburgh, 2009, CERN-THESIS-2009-001, 07 Jan 2009.

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

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

[73] V. Gligorov, *Trigger Schemes*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html, Mar, 2015.

[74] LHCb collaboration, *LHCb online system, data acquisition and experiment control: Technical Design Report*, CERN-LHCC-2001-040. LHCb-TDR-007.

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

[76] A. Puig, *The LHCb trigger in 2011 and 2012*, LHCb-PUB-2014-046.

[77] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[78] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

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

[80] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.

[81] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[82] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[83] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019. LHCb-TDR-011.

[84] A. Pearce *et al.*, *LHCb Starterkit*, https://lhcb.github.io/first-analysis-steps/, Nov, 2015.

[85] LHCb Collaboration, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001. LHCC-I-018.

[86] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001. LHCb-TDR-015.

[87] M. Ferro-Luzzi, *Upgrade VELO aperture*, http://indico.cern.ch/event/189371/, May, 2012.

[88] R. B. Appleby *et al.*, *VELO aperture considerations for the LHCb Upgrade*, Tech. Rep. LHCb-PUB-2012-018, CERN-ATS-Note-2012-101, CERN, Geneva, Dec, 2012.

[89] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021. LHCb-TDR-013.

[90] T. Bird *et al.*, *Simulated Performance of a Strip-Based Upgraded VELO*, Tech. Rep. LHCb-PUB-2014-033. CERN-LHCb-PUB-2014-033, CERN, Geneva, May, 2014, LHCb-PUB-2014-033.

[91] X. Llopart *et al.*, *Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements*, Nucl. Instr. Meth. Phys. Res. A **581** (2007), no. 1-2 485, Proceedings of the 11th International Vienna Conference on Instrumentation.

[92] Poco Graphite, *Poco Graphite: Thermal Management Material*, http://www.poco.com/portals/0/literature/semiconductor/78962v2pocofoamflyer.pdf.

[93] The LHCb VELO Upgrade Group, *LHCb VELO Upgrade Technology Review - Support Document*, Tech. Rep. LHCb-INT-2013-025, CERN-LHCb-INT-2013-025, CERN, Geneva, May, 2013.

[94] CERN PH-DT, VELO Group, O. A. de Aguiar Francisco *et al.*, *Evaporative CO_2 microchannel cooling for the LHCb VELO pixel upgrade*, JINST **10** (2015), no. 05 C05014.

[95] T. Bird, *The upgrade of the LHCb Vertex Locator*, JINST **9** (2014), no. 12 C12041, arXiv:1410.0812.

[96] BaBar Collaboration, J. P. Lees *et al.*, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. **D88** (2013), no. 7 072012, arXiv:1303.0571.

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

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

[99] R. Brun and F. Rademakers, Root- *an object oriented data analysis framework*, vol. 389 of *AIHENP'96 Workshop, Lausanne*, pp. 81–86, Sep, 1996. doi: 10.1016/S0168-9002(97)00048-X.

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

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

[102] A. Martín Sánchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, LHCb-PUB-2011-026. CERN-LHCb-PUB-2011-026.

[103] Mark II Collaboration, K. Riles *et al.*, *Limit on the Decay D0 $\to e^{\pm}\mu^{\mp}$*, Phys. Rev. **D35** (1987) 2914.

[104] E791 Collaboration, E. M. Aitala *et al.*, *Search for rare and forbidden dilepton decays of the D^+, $D^+_{(s)}$, and D^0 charmed mesons*, Phys. Lett. **B462** (1999) 401, arXiv:hep-ex/9906045.

[105] BaBar Collaboration, J. P. Lees *et al.*, *Search for the decay modes $D^0 \to e^+e^-$, $D^0 \to \mu^+\mu^-$, and $D^0 \to e^{\pm}\mu^{\mp}$*, Phys. Rev. **D86** (2012) 032001, arXiv:1206.5419.

[106] T. Bird, *The search for $D^0 \to e^{\pm}\mu^{\mp}$*, in *Proceedings, 6th International Workshop on Charm Physics (Charm 2013)*, 2013. arXiv:1312.1525.

[107] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0 \to K^{*0} e^+ e^-$ branching fraction at low dilepton mass*, JHEP **05** (2013) 159, arXiv:1304.3035.

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

[109] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[110] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, Ann. Statist. **29** (2001) 1189.

[111] M. Perrin-Terrin and G. Mancinelli, *Optimisation of the binning of the discriminating variables used in the computation of $\mathcal{B}\left(B_s^0 \to \mu^\pm \mu^\mp\right)$ upper limits with the modified frequentist approach*, LHCb-INT-2012-003.

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

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

[114] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727, [Erratum: Eur. Phys. J. **C73** (2013) 2501].

[115] L. Moneta *et al.*, *The RooStats Project*, PoS **ACAT2010** (2010) 57, arXiv:1009.1003.

[116] LHCb, R. Aaij *et al.*, *Search for the lepton-flavour violating decay* $D^0 \to e^{\pm}\mu^{\mp}$, arXiv:1512.0032 (2015).

[117] ALEPH Collaboration, D. Buskulic *et al.*, *Observation of the time dependence of* B^0_d-$\bar{B}^0_d$ *mixing*, Phys. Lett. **B313** (1993) 498.

[118] CDF Collaboration, A. Abulencia *et al.*, *Observation of* B^0_s - $\bar{B}^0_s$ *Oscillations*, Phys. Rev. Lett. **97** (2006) 242003, arXiv:hep-ex/0609040.

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

[120] M. Grabalosa and M. Musy, *Flavour Tagging developments within the LHCb experiment*, PhD thesis, Barcelona U., Mar, 2012, CERN-THESIS-2012-075, 15 May 2012.

[121] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.

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

[123] M. Calvi, O. Leroy, and M. Musy, *Flavour Tagging Algorithms and Performances in LHCb*, tech. rep., CERN, Geneva, May, 2007, CERN-LHCB-2007-058.

[124] N. T. Leonardo, *Analysis of* B_s *flavor oscillations at CDF*, FERMILAB-THESIS-2006-18 (2006).

[125] M. S. Anzelc, *Study of B_s Mixing at the DZero Detector at Fermilab Using the Semi-leptonic Decay $B_s \to D_s \mu \nu X$*, FERMILAB-THESIS-2008-07 (2008).

[126] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi \pi^+\pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600.

[127] M. Williams, *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*, JINST **5** (2010) P09004, arXiv:1006.3019.

[128] J. Amoraal *et al.*, *Application of vertex and mass constraints in track-based alignment*, Nucl. Instrum. Meth. **A712** (2013) 48, arXiv:1207.4756.

[129] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001, and 2013 partial update for the 2014 edition.

[130] J.-B. J. Fourier, *Théorie analytique de la chaleur*, Chez Firmin Didot, père et fils (1822).

[131] S. D. Conte and C. de Boor, *Elementary numerical analysis*, McGraw Hill Inc., 1980.

[132] H. G. Moser and A. Roussarie, *Mathematical methods for $B^0\bar{B}^0$ oscillation analyses*, Nucl. Instrum. Meth. **A384** (1997) 491.

[133] LHCb collaboration, R. Aaij *et al.*, *Observation of B_s^0–$\bar{B}_s^0$ mixing and measurement of mixing frequencies using semileptonic B decays*, Eur. Phys. J. **C73** (2013) 2655, arXiv:1308.1302.

[134] S. de Boer and G. Hiller, *Flavor & new physics opportunities with rare charm decays into leptons*, arXiv:1510.0031 (2015).

[135] LHCb, R. Aaij *et al.*, *Measurement of the B^0–$\bar{B}^0$ oscillation frequency Δm_d with the decays $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi K^{*0}$*, Phys. Lett. **B719** (2013) 318, arXiv:1210.6750.

[136] 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.