Measurements of Indirect CP Violation in Charm at LHCb

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

2015

Mark Smith

School of Physics and Astronomy

University of Manchester

Contents

Word count: 35989

List of Figures

List of Tables

Abstract

Measurements of Indirect CP Violation in Charm at LHCb
Mark Smith
Doctor of Philosophy
The University of Manchester
September 2015

This thesis describes two pieces of work. The first is a study of the resolution of the LHCb vertex locator throughout Run 1. The second comprises analyses to measure the charm mixing and CP violation observables A_Γ and y_{CP}.

An estimate of the resolution of the LHCb vertex locator is required for use in the track fits. A method to measure the resolution with collision data has been developed and tested. The performance of the sub-detector throughout Run 1 of the LHC has been assessed. A significant degrading of the resolution has been seen. The effects of this on the track reconstruction has been examined with little change in the measured quantities being observed.

The measurement of indirect CP violation in neutral D meson transitions has been measured through the observables A_Γ and y_{CP}, using $1\,\text{fb}^{-1}$ of pp collisions with a centre of mass energy $7\,\text{TeV}$, collected by the LHCb detector in 2011.

A_Γ describes the CP asymmetry of the lifetime of the D^0 decaying to a CP eigenstate (K^+K^- or $\pi^+\pi^-$). The analysis documented here yields $A_\Gamma = (-0.17 \pm 0.54) \times 10^{-3}$ when the measurements are combined. This is the world's best result and represented a factor of four improvement over the previous best measurement.

The observable y_{CP} compares the effective lifetimes of the Cabibbo favoured decay $D^0 \to K^-\pi^+$ and the transition to a CP eigenstate (K^+K^- or $\pi^+\pi^-$). The unblinded result obtained in this document, averaged over both final states is $y_{CP} = (5.61 \pm 1.56) \times 10^{-3}$. This result is commensurate with the world average central value within 1.25 standard deviations and has significance of $3.6\,\sigma$ relative to zero.

Declaration

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

Mark Smith
School of Physics and Astronomy
University of Manchester
Oxford Road
Manchester
M13 9PL
September 2015

Copyright

*see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487

†see http://www.manchester.ac.uk/library/aboutus/regulations

Acknowledgements

Several individuals have provided much invaluable help in the preparation of this thesis. In particular I would like to express my gratitude to the following:

My supervisor Chris Parkes for all his patience, helpful advice, interesting discussions and careful proofreading of the thesis.

Marco Gersabeck for the enlightening charm discussions and advice.

Michael Alexander for his excellent efforts developing the analysis.

Vava Gligorov for his work on the swimming and very welcome encouragement.

Silvia Borghi for the support with both the physics analysis and the VELO study.

Tomasz Szumlak for his guidance with the VELO work.

Thomas Bird for all his vital computing aid.

All at the Manchester HEP group who have made the PhD a rewarding experience.

Chapter 1

The Standard Model and Flavour Physics

The Standard Model (SM) has proven itself to be a remarkably successful and predictive theory. With the discovery of the Higgs boson in 2012 [1, 2] the final SM particle has been found. Nevertheless theoretical questions remain that the SM is unable to answer. Alongside the contrary experimental observations, such as neutrino masses, dark matter, gravity and the magnitude of Nature's preference for matter over anti-matter there are several aesthetic motivations for searching for physics beyond the Standard Model. These include the fine tuning of particle masses and the 'strong *CP* problem' as well as the large number of arbitrary parameters in the theory. Although experiment has yet to directly observe new physics, precision tests such as those described in this document may characterise it.

In this chapter the Standard Model will be outlined, with the gauge symmetries and their resulting interactions listed. Some attention will be given to how flavour changing transitions arise. Finally charm physics will be examined in detail. An overview of the theoretical challenges faced in the charm sector will be followed by explanations of the principal techniques to measure the physics parameters.

1.1 The Standard Model

The Standard Model is a quantum field theory that is a member of the combination of $\mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ gauge groups. The requirement for invariance under each of these transformations introduces gauge fields that together describe the interactions of the Standard Model as summarised in table 1.1.

The eight fields of the SU(3) group give rise to the strong force. The

Group	Fields
SU(3)	$A^1 \dots A^8$
SU(2)	W^1, W^2, W^3
U(1)	B

Table 1.1: The Standard Model gauge groups and their associated fields.

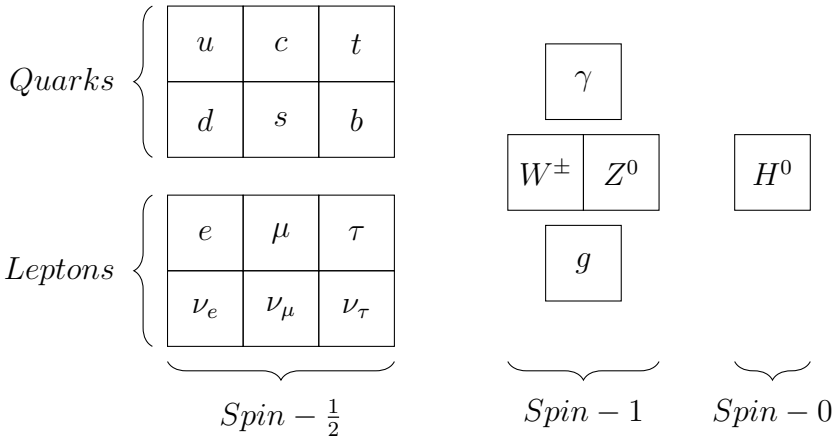


Figure 1.1: The Standard Model particles. On the left are the fundamental fermions, in the centre the gauge bosons and on the right is the Higgs boson.

SU(2) ⊗ U(1) symmetry of the electroweak force is spontaneously broken via the Higgs mechanism [3, 4, 5] to give the observed massive weak bosons and electromagnetic fields as well as the fermion masses. The particles of the Standard Model are summarised in Fig. 1.1. The fundamental particles are the spin-$\frac{1}{2}$ fermions; the quarks and leptons. In addition are the particles arising from the gauge fields, the bosons of spin-1 and spin-0.

1.1.1 Quantum chromodynamics

The Lagrangian for quantum chromodynamics (QCD) is

$$\mathcal{L}_{QCD} = \overline{\psi}(i\not{D} - m)\psi - \frac{1}{2}\mathrm{tr}\{G_{\mu\nu}G^{\mu\nu}\},$$

with the covariant derivative

$$D^\mu = \partial^\mu - igA^\mu(x),$$

and the gauge field tensor

$$G^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + g f_{abc} A^b_\mu(x) A^c_\nu(x). \tag{1.1}$$

Chirality	Left-handed			Right-handed		
Particle	Q	T_3	Y	Q	T_3	Y
u,c,t	$+\frac{2}{3}$	$+\frac{1}{2}$	$+\frac{1}{3}$	$+\frac{2}{3}$	0	$+\frac{4}{3}$
d,s,b	$-\frac{1}{3}$	$-\frac{1}{2}$	$+\frac{1}{3}$	$-\frac{1}{3}$	0	$-\frac{2}{3}$
ν_e,ν_μ,ν_τ	0	$+\frac{1}{2}$	-1	–	–	–
e,μ,τ	-1	$-\frac{1}{2}$	-1	-1	0	-2

Table 1.2: The assigned values of the third component of weak isospin, hypercharge and electric charge for the fermions. As right-handed neutrinos have not been experimentally verified their values have been left blank.

The Lagrangian is invariant under transformations of the SU(3) group.

Expanding out the Lagrangian, one finds that the gauge fields interact with the fermions and with themselves. That they self-interact leads to effects characteristic of QCD; confinement [6] and asymptotic freedom [7, 8, 9]. At large distances the coupling becomes strong and over short distances it is weak. This implies that for low energy interactions (which have large characteristic lengths) QCD is not perturbative and becomes difficult to calculate. Conversely for interactions involving a large momentum transfer a perturbative approach is a suitable approximation to make. Such considerations are particularly pertinent for calculating charm processes as discussed in section 1.3.3.

1.1.2 Electroweak interaction

The electroweak interactions are described by the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry. The conserved quantity Y is referred to as the hypercharge defined as $Y = 2(Q - T_3)$ where T_3 is the third component of weak isospin that is the generator of the SU(2) group. The values of the weak quantum numbers for the fundamental particles is listed in table 1.2. The gauge symmetry is spontaneously broken to leave the $\mathrm{U}(1)_Q$ symmetry of electromagnetism (Q being the conserved electric charge) and the three massive gauge bosons of the weak force. As shown in table 1.2 the electroweak force is a chiral theory; only left-handed particles participate in weak interactions.

Before symmetry breaking the free Lagrangian for some fermion field f is

$$\mathcal{L} = \bar{f} i \not{D} f - \frac{1}{4} W_a^{\mu\nu} W_{\mu\nu}^a - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} + \mathcal{L}_\phi + \mathcal{L}_{Yuk},$$

where the index a sums over the three W fields and the field strength tensors are defined analogously to equation 1.1. The fermion field is split into a left doublet

and right singlet such that

$$f = f_L + f_R,$$

and the covariant derivative is

$$D^\mu = \partial^\mu + ig\tau \cdot \mathbf{W}^\mu + \frac{i}{2}g'YB^\mu.$$

The Lagrangian for the scalar Higgs field ϕ is

$$\mathcal{L}_\phi = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi),$$

with the Higgs potential ϕ taking the Mexican hat shape

$$V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2.$$

Finally the Yukawa term introduces the fermion mass by the interactions of their fields with the Higgs scalar when it acquires a vacuum expectation

$$\mathcal{L}_{Yuk} = -g_f(\bar{f}_L\phi f_R + \bar{f}_R\phi f_L).$$

Inserting the vacuum expectation of the Higgs and considering the physical fluctuations $h(x)$ around it

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v+h) \end{pmatrix},$$

breaks the symmetry. The W^1 and W^2 together form the W^+ and W^- bosons with the same masses $(gv/2)$ whilst the W^3 and B fields mix to form the massive Z^0

$$Z^0 = \cos\theta_W W^3 - \sin\theta_W B,$$

and massless photon

$$A = \sin\theta_W W^3 + \cos\theta_W B.$$

The weak mixing angle θ_W is given by the couplings to the W and B fields, g and g'

$$\cos\theta_W = \frac{g}{(g^2+g'^2)^{\frac{1}{2}}}, \qquad \sin\theta_W = \frac{g'}{(g^2+g'^2)^{\frac{1}{2}}}.$$

1.2 Flavour physics

The precision study of flavour changing interactions has motivated much theoretical progress in the SM. The discovery of parity violation in weak decays [10] strongly hinted at the chiral structure of the weak interaction whilst the measurement of *CP* violation [11] suggested a rather extended quark sector compared to what had been discovered at the time [12]. The lack of flavour changing neutral currents in kaon decays demanded the existence of a fourth quark due to the GIM mechanism [13].

The charged current for lepton interactions is of the form

$$j^{\mu}_{CC} = \bar{u}_l(q_l)\frac{1}{2}\gamma^{\mu}(1-\gamma_5)u_{\nu_l}(p_{\nu_l}).$$

The current is of the V-A form that maximally violates parity; the left handed u_{ν_l} and right-handed $\bar{u}_l$ are picked out by the projection operator.

The quark charged currents are of a similar form but the strong quark states have been transformed by the CKM matrix

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

Imposing the constraint of unitarity and ignoring unobservable phases leaves four independent parameters for the 3×3 matrix; three rotations of the quark states and a complex phase. The matrix can be parametrised in any number of ways such as the Wolfenstein parametrisation [14] which indicates the hierarchy of the matrix. To fifth order it is

$$V_{CKM_{(5)}} = \begin{pmatrix} 1-\frac{\lambda^2}{2}-\frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda + A\lambda^5(\frac{1}{2}-\rho-i\eta) & 1-\frac{\lambda^2}{2}-\frac{\lambda^4}{8}(1+4A^2) & A\lambda^2 \\ A\lambda^3(1-\bar{\rho}-i\bar{\eta}) & -A\lambda^2 + A\lambda^4(\frac{1}{2}-\rho-i\eta) & 1-\frac{1}{2}A^2\lambda^4 \end{pmatrix}, \tag{1.2}$$

with

$$\bar{\rho} = \rho\left(1-\frac{\lambda^2}{2}\right), \qquad \bar{\eta} = \eta\left(1-\frac{\lambda^2}{2}\right).$$

$\lambda = \sin\theta_C \approx 0.22$ where θ_C is the Cabibbo angle of the 2×2 Cabibbo mixing matrix [15]. As is shown in equation 1.2 the Cabibbo matrix is very close to unitary (it is to third order in λ) and transitions between the third second generations are suppressed by the λ^2 factor. *CP* violation is introduced via the

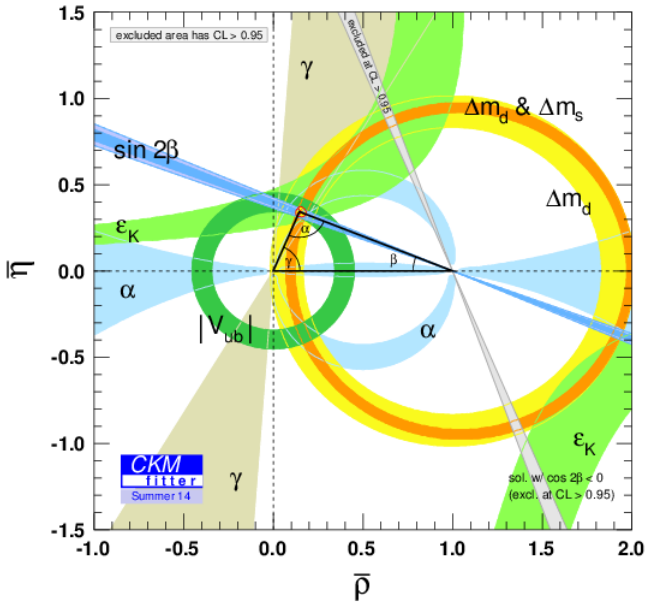


Figure 1.2: The current global fit of the unitary triangle for $\sum_{i=u,c,t} V_{id}V_{ib}^* = 0$. From [18].

complex element in the third generation.

The four parameters of the CKM matrix are fundamental quantities that must be determined by experiment. Doing so forms a significant part of the physics programme of the flavour factories* and is a prerequisite for making accurate SM predictions of flavour changing transitions. The unitarity condition places constraints on the CKM elements

$$\sum_k V_{ik}V_{jk}^* = 0,$$

which can be visualised in the complex plane as closed triangles. In Fig. 1.2 the current global fit of the triangle $\sum_{i=u,c,t} V_{id}V_{ib}^* = 0$ can be seen. This is experimentally an interesting triangle due to the similar lengths of the sides.

1.2.1 *CP* violation

The parity transformation $\hat{P}$ changes the sign of spatial coordinates

$$\vec{r} \to -\vec{r}.$$

*The flavour factories are experiments specifically designed to study flavour-changing transitions. In particular Belle [16] and BaBar [17] have measured large samples of B decays.

Experimental evidence that weak interactions are not invariant under parity transformations [10] motivated the $V - A$ structure of the weak theory [19]. The charge-parity transformation $\hat{C}$ changes particle to anti-particle and vice versa

$$\psi \to \bar{\psi},$$

and has also been shown not to be a symmetry of weak interactions [20]. Furthermore the combination of C and P, CP, is also not a perfect symmetry in weak processes.

The discovery that weak interactions are not symmetric under the combined operation of both transformations (CP) [11] led to the development of the theory of flavour changing interactions described previously. That experiment found CP violation in the mixing of neutral kaons, with asymmetries of the order of 0.2%. It was not until 1999 that CP violation in the decay amplitudes of kaons was conclusively discovered [21, 22]. In 2001 CP violation was found in B mesons [23, 24] and since then many measurements have been made in B decay modes (including B_s^0 and B^+), constraining the parameters of the CKM matrix. Currently CP asymmetry measurements are attractive for searches for new physics. In the Standard Model they are rare processes and so new physics should be clear to see through enhancements of CP asymmetries.

To observe CP violation in a physical transition two interfering amplitudes are required with different strong and weak phases. The difference in magnitude of a transition and its CP conjugate is proportional to $\sin(\delta)\sin(\phi)$ where δ and ϕ are the differences between the strong and weak phases of the contributing amplitudes. The strong phase does not change sign under the CP transformation, whereas the weak phase does.

Experimentally it has been seen that the hadronic weak interactions are the only known part of the Standard Model to exhibit CP violation; most results have shown consistency with the Standard Model, though some tension has become apparent recently (for example ΔA_{CP} in charm which will be discussed in section 1.3.4). However due to the degree of CP violation demanded by the Sakharov conditions [25] for the visible universe one knows there must be additional CP violation in nature, motivating the search for new sources of it.

The strong force also allows for significant CP violation. That it has not been seen, for example the neutron has no electric dipole moment [26], requires arbitrarily setting a Standard Model parameter to be very small. One noted solution to this 'strong CP problem' is that of the Peccei-Quinn mechanism [27], which introduces another symmetry into the Standard Model that is spontaneously broken. The resulting boson, the axion, is a rather attractive candidate for dark

matter. Experimental searches are ongoing but thus far no clear axion signal has been seen (see for example [28, 29]).

Neutrinos have been shown to mix [30] via the PMNS matrix [31], in a manner analogous to the mechanism for quarks, suggesting that *CP* violation could be present in leptonic weak interactions. Indeed neutrino mixing is already a sign of physics beyond the Standard Model; measuring the possible neutrino *CP* violating parameter δ is an important goal of neutrino experiments.

1.3 Charm Physics

Since its discovery in 1974 [32, 33] the charm quark has not garnered the same quantity of headlines that kaon and B physics have. However in this century charmed spectroscopy has generated interest due to the observation of several anomalous states [34, 35]. Indeed the charm quark forms the basis for the first unambiguously observed tetraquark [36] and pentaquark [37] states. It was not until 2007 that neutral meson oscillations were discovered in charm with the first single measurement of charm mixing with greater than $5\,\sigma$ significance being made in 2011 [38]. Thus far there have also been no unambiguous signs of *CP* violation in charm. Both mixing and *CP* violation are expected to be small, with potential enhancement from new physics. With the large data sets acquired by the flavour factories and the great statistical power afforded by the LHC the charm sector is now a burgeoning area of flavour physics research.

In this section the phenomenology of mixing and *CP* violation in neutral charm mesons will be outlined and the relevant physics parameters introduced; much of the formalism applies to other neutral meson systems. The *CP* violating observables y_{CP} and A_Γ, whose measurements are the subject of this document, will be derived. Finally overviews of the current theoretical and experimental status of charm flavour physics will be presented. The interested reader may find more information in the plethora of review articles, such as [39, 40].

1.3.1 Mixing and *CP* violation phenomenology

Considering the the two flavour eigenstates $|D^0\rangle$ and $|\overline{D}^0\rangle$ one can write the time evolution of the wave function $|\psi\rangle = \binom{D^0}{\bar{D}^0}$ from the Schrödinger equation

$$i\frac{\partial}{\partial t}|\psi\rangle = \mathbf{H}|\psi\rangle = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)|\psi\rangle, \tag{1.3}$$

where the Hamiltonian $\mathbf{H}$ is a 2×2 matrix. The elements are such that $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$. Non-zero values of the off-diagonal elements

of $\mathbf{H}$ implies oscillations between the D^0 and $\overline{D}^0$ whilst the diagonal part describes the decay.

Diagonalising the Hamiltonian allows one to arrive at the mass eigenstates of the neutral charm mesons, composed from linear combinations of the two flavour eigenstates

$$|D_1\rangle = p|D^0\rangle + q|\overline{D}^0\rangle \qquad\qquad |D_2\rangle = p|D^0\rangle - q|\overline{D}^0\rangle. \tag{1.4}$$

p and q may be complex and $|p|^2 + |q|^2 = 1$. If $p = q = \frac{1}{\sqrt{2}}$ these are CP eigenstates, $CP|D_{1,2}\rangle = \eta_{1,2}|D_{1,2}\rangle$, with the convention chosen that $\eta_1 = -1$ and $\eta_2 = +1$. The masses and widths of the two mass states are $m_{1,2}$ and $\Gamma_{1,2}$. The mass states are observables so the Hamiltonian in this basis is Hermitian.

The differences in mass and width of $|D_1\rangle$ and $D_2\rangle$ are

$$\Delta M = 2\Re\sqrt{\mathbf{H}_{12}\mathbf{H}_{21}} \qquad\qquad \Delta\Gamma = -4\Im\sqrt{\mathbf{H}_{12}\mathbf{H}_{21}},$$

with $\mathbf{H}_{12}$ and $\mathbf{H}_{21}$ referring to the elements of the Hamiltonian of the flavour states.

$$\mathbf{H}_{12} = \mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12} \qquad\qquad \mathbf{H}_{21} = \mathbf{M}^*_{12} - \frac{i}{2}\mathbf{\Gamma}^*_{12},$$

and

$$\left(\frac{q}{p}\right)^2 = \frac{\mathbf{H}_{12}}{\mathbf{H}_{21}}.$$

In the event of CP conservation $\mathbf{H}_{12} = \mathbf{H}_{21}$. Calculating the elements of the Hamiltonian is generally rather tricky, with the D system bringing its own issues which are discussed further in 1.3.3.

CP violation arises in mixing when p and q are different ($|p/q| \neq 1$). An asymmetry due to mixing, A_m, is defined as

$$\left|\frac{q}{p}\right|^{\pm 2} \approx 1 \pm A_m. \tag{1.5}$$

By inserting equations 1.4 into the time dependent Schrödinger equation one can establish the amplitude of finding the flavour eigenstates at time t after the formation of an initial flavour state. In the case of an initial D^0 the wavefunction is:

$$|\Psi(t)\rangle = \frac{1}{2}\left[\left(e^{-im_2 t - \frac{1}{2}\Gamma_2 t} + e^{-im_1 t - \frac{1}{2}\Gamma_1 t}\right)|D^0\rangle - \frac{q}{p}\left(e^{-im_2 t - \frac{1}{2}\Gamma_2 t} - e^{-im_1 t - \frac{1}{2}\Gamma_1 t}\right)|\overline{D}^0\rangle\right]. \tag{1.6}$$

With $\Delta m = m_2 - m_1$ and $\Delta\Gamma = \Gamma_2 - \Gamma_1$,$m_{1,2}$ and $\Gamma_{1,2}$ being the masses and widths of the mass states, the following quantities that define the mixing of the flavour states are defined:

$$\begin{aligned} m &= \frac{m_1 + m_2}{2} & \Gamma &= \frac{\Gamma_1 + \Gamma_2}{2} \\ x &= \frac{\Delta m}{\Gamma} & y &= \frac{\Delta\Gamma}{2\Gamma}. \end{aligned} \tag{1.7}$$

The decay amplitudes from a D to a final state f or its charge conjugate $\overline{f}$ via the Hamiltonian $\mathscr{H}$ are

$$\begin{aligned} A_f &= \langle f|\mathscr{H}|D\rangle & A_{\overline{f}} &= \langle \overline{f}|\mathscr{H}|D^0\rangle \\ \overline{A}_f &= \langle f|\mathscr{H}|\overline{D}^0\rangle & \overline{A}_{\overline{f}} &= \langle \overline{f}|\mathscr{H}|\overline{D}^0\rangle. \end{aligned} \tag{1.8}$$

Direct *CP* violation occurs for those amplitudes where $\left|A_f\right| \neq \left|\overline{A}_{\overline{f}}\right|$. The direct *CP* asymmetry, A_d, for decays to self conjugate final states is

$$\left|\frac{\overline{A}_f}{A_f}\right|^{\pm 2} \approx 1 \pm A_d, \tag{1.9}$$

and the parameter λ_f is defined as

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f} = \left|\frac{q}{p}\right|\left|\frac{\overline{A}_f}{A_f}\right| \exp^{i(\phi+\delta)}. \tag{1.10}$$

CP violation is apparent if the magnitude of λ_f deviates from one or the phase between mixing and decay is not zero. The strong phase δ is apparent in those decays that are not to a *CP* eigenstate, for example $D^0 \to K^-\pi^+$; there is a phase difference between the doubly Cabibbo suppressed amplitude and that for mixing followed by decay, shown in Fig. 1.3. The strong phase does not change sign under the *CP* transformation, the weak phase ϕ does. In the decays to *CP* eigenstates (K^+K^- or $\pi^+\pi^-$) there is no strong phase difference ($\delta = 0$).

The time dependent decay rate to a final state f from an initial D^0 state is given by $|\langle f|\mathscr{H}|\psi(t)\rangle|^2$. Inserting the parameters in 1.7 where appropriate one finds

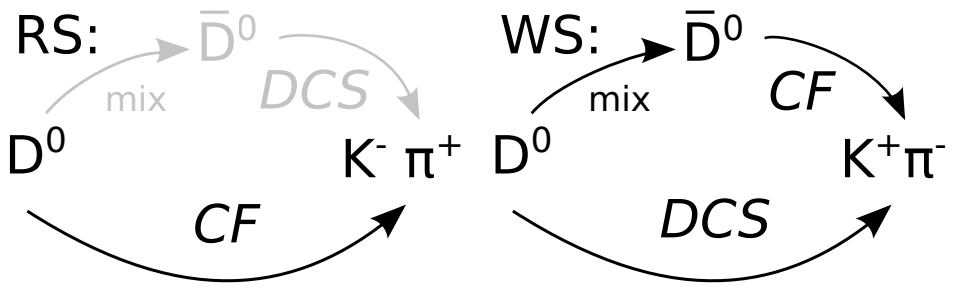


Figure 1.3: The paths of the right-sign $D^0 \to K^-\pi^+$ (left) and wrong-sign $D^0 \to K^+\pi^-$ (right) decays. The right sign decay is completely dominated by the Cabibbo favoured amplitude. The wrong-sign can proceed via the doubly Cabibbo supressed decay or by mixing and then a Cabibbo favoured decay. There is a strong phase δ between the two possible paths.

$$\begin{aligned}\Gamma\left(D^0(t) \to f\right) = \frac{1}{2}e^{-\Gamma t}|A_f|^2[\left(1+|\lambda_f|^2\right)\cosh(y\Gamma t) + \left(1-|\lambda_f|^2\right)\cos(x\Gamma t) \\ + Re(\lambda_f)\sinh(y\Gamma t) - Im(\lambda_f)\sin(x\Gamma t)] \\ \Gamma\left(\overline{D}^0(t) \to f\right) = \frac{1}{2}e^{-\Gamma t}|\overline{A}_f|^2[\left(1+|\lambda_f^{-1}|^2\right)\cosh(y\Gamma t) + \left(1-|\lambda_f^{-1}|^2\right)\cos(x\Gamma t) \\ + Re(\lambda_f^{-1})\sinh(y\Gamma t) - Im(\lambda_f^{-1})\sin(x\Gamma t)]. \quad (1.11)\end{aligned}$$

1.3.2 y_{CP} and A_Γ

For decays to *CP* eigenstates (K^+K^- or $\pi^+\pi^-$) one inserts the appropriate expression for λ_f with its phase. Experiments have shown that for D^0 mesons x and y are small; the Heavy Flavour Averaging group lists $x = (0.41^{+0.14}_{-0.15})\,\%$ and $y = (0.63^{+0.07}_{-0.08})\,\%$ [41] (allowing for both direct and indirect *CP* violation). Thus equation 1.11 can be expanded to first order to reach

$$\begin{aligned}\Gamma\left(D^0(t) \to K^+K^-\right) = e^{-\Gamma t}|A_{KK}|^2[1 - \frac{q}{p}(y\cos\phi - x\sin\phi)\Gamma t] \\ \Gamma\left(\overline{D}^0(t) \to K^+K^-\right) = e^{-\Gamma t}|\bar{A}_{KK}|^2[1 - \frac{p}{q}(y\cos\phi + x\sin\phi)\Gamma t]. \quad (1.12)\end{aligned}$$

Experimentally one seeks to fit the lifetime to an exponential decay $|A_f|^2 e^{-\hat{\Gamma}t}$. The Taylor expansion to first order of the exponential is

$$|A_f|^2 e^{-\hat{\Gamma}t} \approx |A_f|^2\left(1-\hat{\Gamma}t\right). \quad (1.13)$$

Therefore approximating the terms in parentheses in equation 1.12 as an

exponential and collecting the terms appropriately one finds

$$\hat{\Gamma}(D^0 \to K^+K^-) = \Gamma[1 + |\frac{q}{p}|(y\cos\phi - x\sin\phi)] \tag{1.14}$$

$$\hat{\Gamma}(\overline{D}^0 \to K^+K^-) = \Gamma[1 + |\frac{p}{q}|(y\cos\phi + x\sin\phi)]. \tag{1.15}$$

Here $\hat{\Gamma}$ is the effective width of the decay and Γ is the total width.

Finally one arrives at two measurable time dependent quantities defined as

$$\begin{aligned} y_{CP} &= \frac{\hat{\Gamma}(D^0 \to K^+K^-) + \hat{\Gamma}(\overline{D}^0 \to K^+K^-)}{2\Gamma} - 1 \\ &= \frac{|\frac{q}{p}| + |\frac{p}{q}|}{2} y\cos\phi - \frac{|\frac{q}{p}| - |\frac{p}{q}|}{2} x\sin\phi \\ &\approx \left(1 - \frac{1}{8}A_m^2\right) y\cos\phi - \frac{1}{2}A_m x\sin\phi, \end{aligned} \tag{1.16}$$

and

$$\begin{aligned} A_\Gamma &= \frac{\hat{\Gamma}(D^0 \to K^+K^-) - \hat{\Gamma}(\overline{D}^0 \to K^+K^-)}{2\Gamma} \\ &= \frac{|\frac{q}{p}| - |\frac{p}{q}|}{2} y\cos\phi - \frac{|\frac{q}{p}| + |\frac{p}{q}|}{2} x\sin\phi \\ &\approx \frac{1}{2}(A_m + A_d) y\cos\phi - x\sin\phi. \end{aligned} \tag{1.17}$$

The parameter Γ (the total width) in the denominator can be measured from the decay to the non *CP* eigenstates $K^-\pi^+$. This decay is Cabibbo favoured and can very closely be approximated to an exponential leading to the assertion that $\hat{\Gamma}(D^0 \to K^-\pi^+) = \Gamma$. Assuming *CP* violating effects are small one can define y_{CP} and A_Γ as

$$y_{CP} = \frac{\hat{\Gamma}(D^0 \to K^+K^-)}{\hat{\Gamma}(D^0 \to K^-\pi^+)} - 1 \tag{1.18}$$

$$A_\Gamma = \frac{\hat{\Gamma}(D^0 \to K^+K^-) - \hat{\Gamma}(\overline{D}^0 \to K^+K^-)}{\hat{\Gamma}(D^0 \to K^+K^-) + \hat{\Gamma}(\overline{D}^0 \to K^+K^-)}. \tag{1.19}$$

These can both be defined with the $\pi^+\pi^-$ final state instead of K^+K^-, yielding the same physics parameters.

In the limit of no *CP* violation $|q/p| = 1$ and $\phi = 0$ leading to $A_\Gamma = 0$ and y_{CP} becomes the mixing parameter y. To obtain a determination of *CP* violation in mixing the y_{CP} result has to be compared with an independent measurement of

y (for example [42, 43]). Equation 1.18 shows that by measuring the lifetime τ (where $\tau = 1/\hat{\Gamma}$) for D^0 decays to the various final states one can calculate A_Γ and y_{CP}. Such lifetime measurements are the subject of this document.

In the SM the phase ϕ is independent of the final state as it only describes the *CP* phase

$$\phi = \arg\left(\frac{M_{12}}{\Gamma_{12}}\right).$$

Assuming negligible direct *CP* violation this implies that A_Γ is final state independent. New physics could well introduce some final state dependency so measuring A_Γ for two different decays separately gives another test of the SM. If ϕ varies with decay the difference in A_Γ between K^+K^- and $\pi^+\pi^-$ is

$$\begin{aligned}\Delta A_\Gamma &= A_\Gamma(K^+K^-) - A_\Gamma(\pi^+\pi^-) \\ &\approx \Delta A_d y \cos\phi + (A_m + A_d) y\Delta \cos\phi - x\Delta \sin\phi. \end{aligned} \tag{1.20}$$

1.3.3 Theoretical overview

At the quark level D^0 oscillations occur via box diagrams such as that shown on the left in Fig. 1.4. As well as these short range interactions there is a considerable contribution from longer range processes with the possibility of the intermediate quarks forming hadronic states. This is in contrast to the B system where the mixing is dominated by short range processes as the quarks involved in the transition are heavy and there is a large momentum transfer (implying a short range). In the D system however the charm quark is comparatively light so there is a smaller momentum transfer, implying a longer range of the interaction with the possibility that intermediate hadronic states are an influence. That the D^0 is the only up type meson to mix makes it unique compared to the B and K systems; it is the only one where the mixing occurs through intermediate down type quarks. The diagram in Fig. 1.4 suggests that the short range contributions to mixing should be small in the SM as there is large GIM suppression due to the near degeneracy of the d and s quark masses and the CKM suppression of the contribution of the b [44]. That same suppression implies that *CP* violation is small in mixing in the SM (the amplitude will be suppressed by $V_{cb}V_{ub} \sim \lambda^5$).

Similarly in the decay, shown at tree level on the right of Fig. 1.4, the weak phase of the third generation would enter in second order diagrams such as in Fig. 1.5, again implying that *CP* violation is a small effect. These loop diagrams make charm an attractive arena for the search for new physics as new particles can enter into them, potentially giving significant enhancements to the *CP* violating

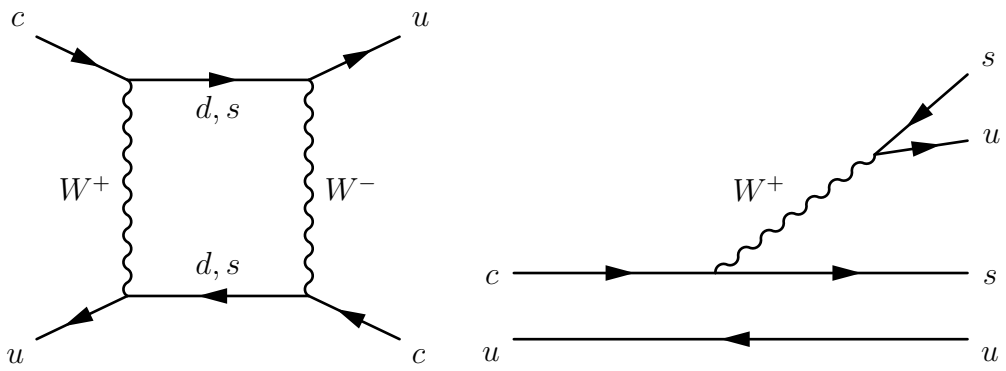


Figure 1.4: The leading order Feynman diagrams for D^0 mixing (left) and a singly Cabibbo suppressed decay, $D^0 \to K^+K^-$ (right).

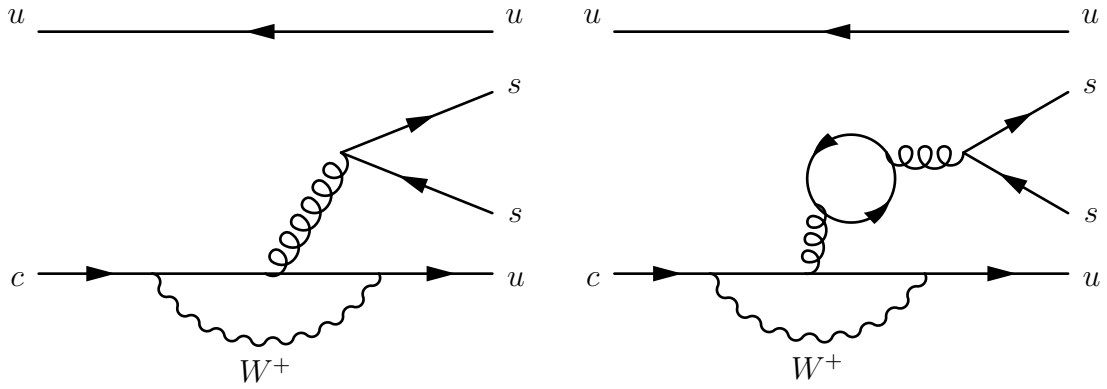


Figure 1.5: Two penguin diagrams that may contribute to the *CP* violation in the singly Cabibbo suppressed $D^0 \to K^+K^-$ decay.

amplitudes.

Making theoretical predictions for the charm sector has proved a tricky endeavour. For example consider Fig 1.6, an assembly of SM predictions for the charm mixing parameters x and y from a few years before the discovery of charm mixing. The numbers plotted show a spread of over three orders of magnitude. Whilst theoretical calculations have improved since then there are still no quantitative predictions for *CP* violation and calculations for the mixing parameters and D lifetimes show some discrepancy with experiment.

One could take either of two approaches to calculating amplitudes of a charm transition. An exclusive approach relies on explicitly accounting for every possible intermediate state (see for example [46, 47]). Alternatively an inclusive method may be utilised where the Hamiltonian describing the transition is replaced by an effective Hamiltonian, parametrised as a series expansion of operators of increasing dimensions in powers of $\frac{1}{m_q}$, m_q being the mass of the heavy (in this case c) quark.

In the B system the Heavy Quark Expansion (HQE) has been used to successfully make predictions of lifetimes (see for example the review [48]). The b quark is heavy so the expansion converges readily and long range QCD effects are small. In the charmed system however the intermediate mass of the quark gives

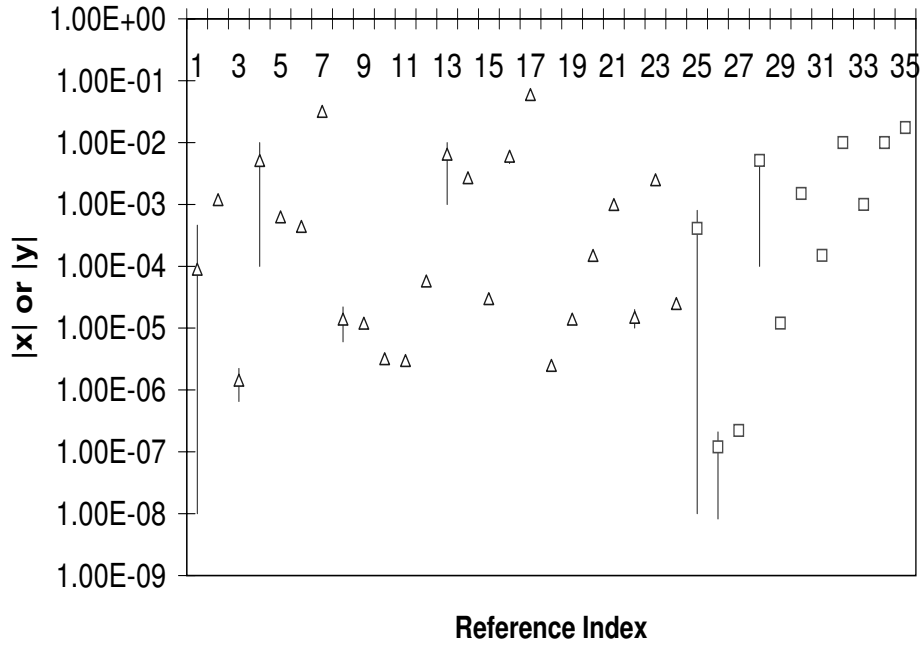


Figure 1.6: A collection of SM predictions for the charm mixing parameters x (triangles) and y (squares). Taken from [45], see therein for the numbered references.

larger contributions to the higher order operators. Additionally the long range, non-perturbative, QCD contributions are potentially much larger; unfortunately they are also difficult to calculate.

The lack of a clear theoretical picture means that one cannot make a statement about what to expect from *CP* violation in charm with any kind of authority. As noted in [49] "*CP violation of the order of one per mille is an unambiguous signal for new physics*" seems to be the consensus for many. However the author notes that such statements should be met with caution and by including large $\mathrm{SU}(3)_F$ symmetry breaking effects *CP* violation of order several 10^{-3} could be accommodated within the Standard Model.

Such thoughts were rather apparent after the LHCb ΔA_{CP} result in 2011. A direct *CP* asymmetry had been measured of order a few per mille with significance of more than three standard deviations [50], a result corroborated by other experiments [51, 52]. Such a large deviation from 0 had not been expected and led to a reconsidering of the SM expectations. Some have claimed that the SM could be stretched to accommodate such large *CP* violation [53, 54] although numerical predictions are still unavailable. More recent results have reduced the significance of the central value for ΔA_{CP} [55, 56] and the theoretical picture remains unclear.

As far as the measurements presented in this document are concerned the figure usually quoted by the analysts for the SM expectation is $A_\Gamma \sim \mathcal{O}(10^{-4})$. That number comes from the averaged mixing parameter values and the above

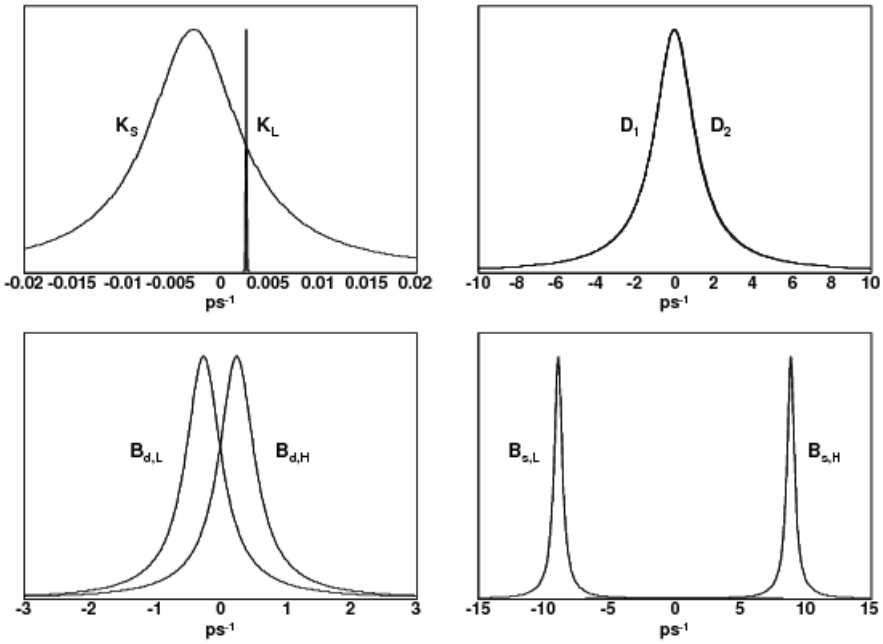


Figure 1.7: The mass splits and widths of the neutral meson mass states. The widths are the inverse of the lifetimes of the mesons. The mass differences determine the oscillation frequencies. The two D mass states have almost the same masses and widths which presents an experimental challenge to discern them. Reproduced from [39].

statement about the expected size of A_m being less than % level, assuming a very small phase. As already noted A_Γ should be very close to final state independent in the SM (assuming negligible direct *CP* violation). Any difference between the K^+K^- and $\pi^+\pi^-$ final states at the experimental precision currently achievable would be an unambiguous indication of new physics.

The picture for y_{CP} is no clearer; predictions currently do not reflect the experimental value of y. Again any kind of deviation from y is likely to be small due to the A_m^2 term and the assumption that ϕ is small. Assuming no *CP* violation the measurement of y_{CP} does provide a statistically powerful tool for measuring y and will constrain the averages of the mixing parameters.

1.3.4 Experimental overview

The charm sector has its own unique challenges for flavour physics measurements. As shown in Fig. 1.7 the D mass states are very similar in mass and lifetime. In contrast the B^0 and B_s^0 mesons have sizeable mass differences leading to rapid oscillations of period much smaller than their mean lifetimes. Kaons have a small mass splitting but very different widths allowing for the two mass states to be discerned readily.

This similarity between the masses and widths of the mass states of the neutral D mesons makes measuring the mixing difficult (of course the B and K mesons present experimental challenges of their own). As the oscillation period is much longer than the mean lifetime most D mesons decay before the oscillation of the mixed states has developed. The contrast with the B^0 mesons can be seen in Fig. 1.8. The current HFAG averages of the charm mixing parameters are [41]

$$x = (0.49^{+0.14}_{-0.15})\%$$
$$y = (0.62 \pm 0.08)\%.$$

Identifying a D decay from the particles produced in a proton-proton collision presents its own challenges. The momentum distribution of the decay products of the D is lower than for a B so the requisite looser cuts inevitably leads to a greater amount of background appearing in the data sample. If the selection is based upon finding a displaced vertex (as for example at LHCb) the short lifetime of the D again demands a loose selection (as well as good detector precision), reducing the purity of the sample. Displaced vertex selections also cut out a large quantity of the signal with low lifetimes. For a D^0 produced at the interaction point, referred to as 'prompt', the initial flavour of the D^0 is determined by searching for the strong decay $D^{*+} \to D^0\pi^+$, with the charge of the pion indicating the D^0 flavour. Such a tagged analysis also has the advantage of greatly reducing combinatorial backgrounds.

An alternative is to select charm from B decays. This has the advantage of much more efficient selections and in particular includes D^0 at very low decay times. At LHCb such analyses are carried out using semi-leptonic B decays, giving the results the moniker 'semi-leptonic', where the charge of the lepton indicates the initial D^0 flavour. The downside of such measurements is the lower statistics due to the reduced production cross-section of B particles with respect to charm and the subsequent factor due to the branching ratio of the decay of interest. Nevertheless they do provide independent data samples for checking the results for charm produced at the interaction point.

The measurements of the charm mixing and CP violation parameters is achieved through a veritable smörgåsbord of complementary analyses. A summary of the LHCb efforts (by no means exhaustive) and how they contribute to the determination of the physics variables is shown in Fig. 1.9.

Mixing of neutral charm mesons was first established in 2007 through a combination of several measurements [58, 59, 60]. It was not until 2012 that the first single measurement of charm mixing was made with statistical significance greater than $5\,\sigma$ using $1\,\text{fb}^{-1}$ of data taken in 2011 by LHCb [38].

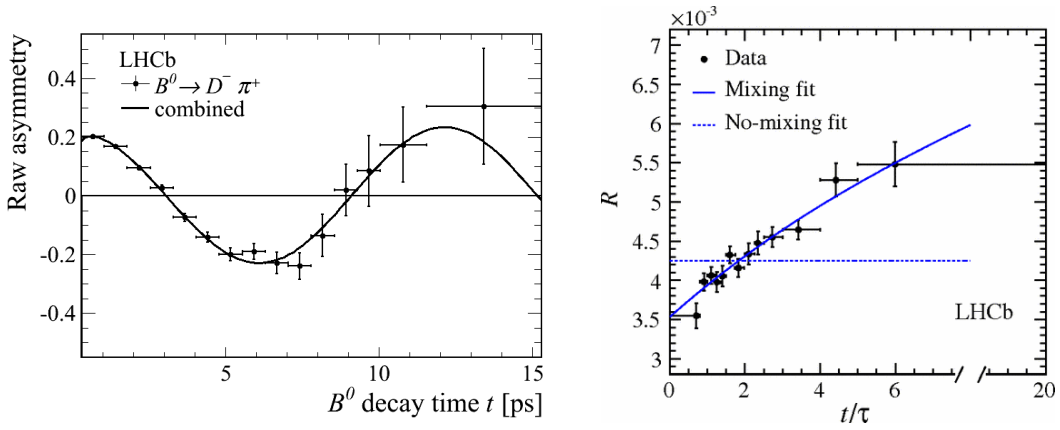


Figure 1.8: The time evolution of the B^0 (left) [57] and D^0 (right) mesons [38]. The left diagram shows the time evolution of the difference between the number of B^0 candidates that have changed flavour (mixed) before decaying and those that have not. The right hand plot displays the ratio of 'right-sign' to 'wrong-sign' decays as a function of time. In both cases the effect of mixing can be clearly seen.

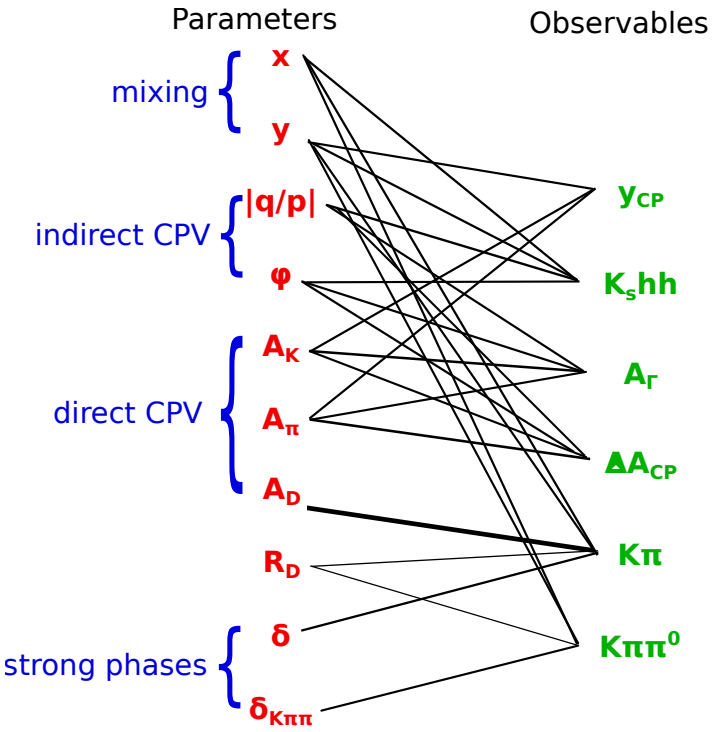


Figure 1.9: The charm flavour physics parameters (left) and the observables measured by LHCb to ascertain them (right). This figure is not exhaustive and neglects the great contributions of other flavour experiments. From Marco Gersabeck.

LHCb studied the time evolution of the yields of $D^0 \to K^+\pi^-$ and $D^0 \to K^-\pi^+$ (with the charge conjugates), the wrong-sign (WS) and right-sign (RS) decays respectively. Wrong-sign decays are heavily supressed as they can only proceed via a doubly Cabibbo supressed decay or via mixing followed by the Cabibbo favoured decay. The right-sign decay is dominated by the Cabibbo favoured transition. As shown in the LHCb result in Fig. 1.8 mixing causes the ratio of the yields, R, to develop over proper time, the fitted curve giving access to y' and x'^2, the mixing parameters (equation 1.7) rotated by the strong phase δ, as

$$R(t) = \frac{N_{WS}(t)}{N_{RS}(t)} = R_D + \sqrt{R_D}y't + \frac{x'^2 + y'^2}{4}t. \tag{1.21}$$

R_D is the ratio of the amplitudes of the Cabibbo favoured and and doubly Cabibbo suppressed decays and $N_{WS}(t)$ and $N_{RS}(t)$ are the yields of the WS and RS decays at proper time t. The no-mixing hypothesis was excluded by LHCb with significance of $9.1\,\sigma$ in 2011 and the subsequent update with the full Run 1 data set significantly improved on that [61]. Belle, Babar and CDF have made similar measurements, albeit with less statistical reach. Of course other combinations of RS and WS final state can be used, such as $D^0 \to K^-\pi^+\pi^+\pi^-$ and $D^0 \to K^+\pi^-\pi^+\pi^-$, with different strong phases and amplitude ratios.

Hadronic decays are also used to measure y_{CP}, which as previously noted becomes y in the limit of no CP violation. Both Belle [62] and Babar [63] have completed analyses on their full data sets, measuring $y_{CP} \neq 0$ with significance of $4.5\,\sigma$ and $3.3\,\sigma$ respectively. It is hoped that the LHCb analysis presented in this document makes the first individual measurement of y_{CP} with greater than $5\,\sigma$ significance.

For getting direct access to the mixing parameters one considers semi-leptonic decays, again comparing WS ($D^0 \to l^-\overline{\nu}_l X$) and RS ($D^0 \to l^+\nu_l \bar{X}$) modes. The WS decay can only be seen due to mixing (as opposed to the doubly Cabibbo suppressed channel in the hadronic modes) so that one has direct access to $\frac{1}{2}(x^2 + y^2)$. Such analyses have been performed at the flavour factories [64, 65] but statistically they are not competitive due to the very small expected mixing rate. At LHCb such studies would be difficult due to the missing energy of the neutrino.

In addition to the semi-leptonic measurements, lepton colliders benefit from being able to produce quantum correlated D^0–$\overline{D}^0$ pairs for use in charm studies via the channel $e^+e^- \to \psi(3770) \to D^0\overline{D}^0$. As they are the only results of the resonance decay they must be a quantum correlated C-odd state meaning that ascertaining the CP of one D decay implies the other is in the opposite CP state. For mixing this allows rather direct access to the strong phase in $K^-\pi^+$ decays [66] through the asymmetry of the CP state decays [67, 68].

Although mixing is now well established, *CP* violation in either its direct or indirect manifestations has yet to make an unambiguous appearance. As noted previously the SM expectations are that it should be small, with $\mathcal{O}(10^{-4})$ usually quoted for asymmetries such as A_Γ though up to $\mathcal{O}(10^{-3})$ could be reasonable. Experimental progress is such that precisions of a few 10^{-3} are achievable.

Significant excitement was generated in the charm community by the measurement of ΔA_{CP} by LHCb in 2011 [50]. ΔA_{CP} is a time-integrated *CP* asymmetry defined as

$$\begin{aligned}\Delta A_{CP} &= A_{meas}(K^+K^-) - A_{meas}(\pi^+\pi^-) \\ &\approx \Delta A_{CP}^{dir}\left(1 + y_{CP}\frac{\langle \bar{t}\rangle}{\tau}\right) + a_{CP}^{ind}\frac{\Delta\langle t\rangle}{\tau},\end{aligned} \tag{1.22}$$

where the individual asymmetries for final state f are

$$A_{meas}(f) = \frac{N(D^0 \to f) - N(\overline{D}^0 \to f)}{N(D^0 \to f) + N(\overline{D}^0 \to f)},$$

which can be written as the sum of the physics A_{CP}, production A_{prod} and detection A_{det} asymmetries

$$A_{meas}(f) = A_{CP} + A_{prod} + A_{det}.$$

Taking the difference between the final states should to a good approximation cancel the last two leaving only the difference in the physics asymmetries which are expected to be of opposite sign [69]. As the last term of eq. 1.22 is very small ΔA_{CP} is mostly an exploration of direct *CP* violation; it can be thought of as similar to a time integrated version of A_Γ.

In 2011 a preliminary result of $\Delta A_{CP} = (-0.82 \pm 0.21_{stat} \pm 0.11_{syst})\%$ was released by LHCb, a deviation from 0 of $3.5\,\sigma$ significance, using $0.6\,\text{fb}^{-1}$ of data taken in 2011 [50]. That was confirmed by the CDF analysis on their full $9.7\,\text{fb}^{-1}$ data set which showed good agreement with LHCb and significance of $2.7\,\sigma$ [51] as well as Belle with $2.1\,\sigma$ [52]. Such a value of ΔA_{CP} was larger than expected in the SM and lead to a reconsideration by the theoretical community of what to expect as noted in section 1.3.3.

However, the resulting update with the full $1\,\text{fb}^{-1}$ 2011 data set [55] (there was a significant change to the method so this result supersedes the previous one), along with a complementary analysis using a statistically independent data set from LHCb [56], referred to as the semi-leptonic sample (using charm from B decays instead of that produced in the proton interaction), showed significantly better agreement with 0. The current world average for Δa_{CP}^{dir} is

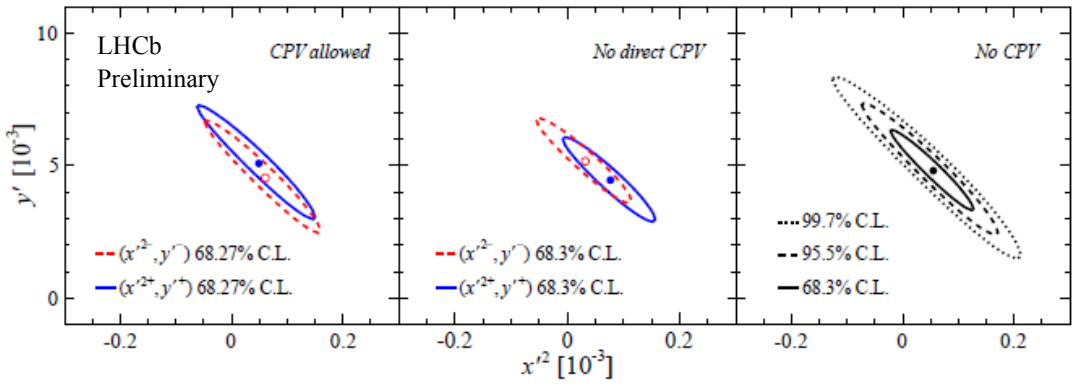


Figure 1.10: The fit results for y' and x'^2 from the LHCb 2012 WS analysis. In blue is the D^0 fit, the $\overline{D}^0$ being in red; the two agree within $1\,\sigma$ uncertainty so no *CP* violation in decay or mixing is seen. The right hand plot shows the total fit for x'^2 and y' without *CP* violation. The no mixing point at (0,0) is very well excluded. Reproduced from [61].

$\Delta a_{CP}^{dir} = (-0.253 \pm 0.104)\%$; new measurements are needed to clarify the situation.

In indirect *CP* violation searches LHCb results again give the highest precision. The WS analysis that measured mixing has been updated to look for *CP* violation with the full Run 1 data set [61] and with the data fitted under the hypothesis of both direct and indirect *CP* violation. As can be see in Fig 1.10 none can be distinguished in the fits but stringent limits on the physics parameters can be set. The measurement of A_Γ, a time dependent *CP* asymmetry, has similarly helped to constrain the physics parameters and is the subject of this thesis.

In addition to these modes multi-body hadronic decays can be studied. As these proceed via intermediate resonances the phase space of the pairs of final state particles must be considered to be sensitive to the contributions of each one. Considering only the total final state (a global asymmetry) ignores the potential for *CP* violation in the Dalitz plane that overall cancels. Such 'Dalitz' analyses can be carried it in a model independent way such as in [70, 71]. Implementing a model of the resonant structure of the decay does allow one to disentangle the contributions to an observed asymmetry but then one is limited by veracity of the model employed.

Experimental status of y_{CP} and A_Γ

Before the LHCb Run 1 measurements, y_{CP} and A_Γ have both been measured extensively, the former using an array of possible Cabibbo favoured final states. Determining y_{CP} has been one of the principle avenues in the search for mixing. These measurements have centred around measuring the effective lifetimes in the

manner of the analyses that make up the bulk of this document. Doing so at the flavour factories is somewhat more straightforward than at a hadron collider. The low event rates and multiplicities at these facilities rather negate the need for a complex trigger or any kind of detailed pre-selection before the event is reconstructed; the datasets are therefore unbiased in proper time. The trigger requirements of hadron colliders and the acceptance effects they introduce are a great challenge for the analysts and novel methods have been developed (see 'swimming' in section 4.2.3).

An alternative method for determining A_Γ is to use a binned ratio method, similar to that of the WS analysis. As the acceptance of the trigger should treat D^0 and $\overline{D}^0$ equally for the final state it will cancel in ratios. This method was used as a complement to the lifetime measurement presented in this document, although its larger systematic uncertainty meant it served as cross-check [72]. The ratio method is an attractive prospect due to the reduced complexity of the fits and the lower computing resources required. However this method is not suitable for y_{CP} in Run 1 due to the different final states having different acceptances that do not cancel in the ratio.

The final states K^+K^- and $\pi^+\pi^-$ as used for measuring A_Γ and y_{CP} in this thesis are two-body *CP* eigenstates. Alternatively one may use multi-body final states that are self-conjugate (for example $D \to \pi^+\pi^-\pi^0$) if the *CP* content of the decay is known [73]. Generally A_Γ can be considered as the asymmetry of D^0 and $\overline{D}^0$ decaying to a *CP*-even final state, with y_{CP} being the ratio of the effective width of the *CP* even decay to the mean width of the D^0. Therefore using the known *CP*-even fraction, F_+, of the multi-body decay one can define an effective A_Γ and y_{CP} [73] and add further input to the fits for the parameters ϕ, x, y and $\left|\frac{q}{p}\right|$. The *CP* content of the multi-body decays maybe ascertained using quantum-correlated D^0 pairs as studied with CLEO [74]. Although the sensitivity to the mixing and *CP* violation parameters would be reduced due to the fractional *CP* content, such an analysis would benefit from the comparatively large branching fraction of the multi-body mode and would help to constrain the parameters.

The state of y_{CP} and A_Γ prior to the two analyses presented in this thesis is shown in Fig. 1.11 with the Heavy Flavor Averaging Group (HFAG) average. The 'LHCb 2012' result was achieved with $29\,\text{pb}^{-1}$ of data collected in 2010 [75]; this analysis proved the validity of the method applied here although due to the low statistics a number of simplifications were made to the background treatment. Nevertheless it already was competitive in y_{CP} and showed promise for A_Γ.

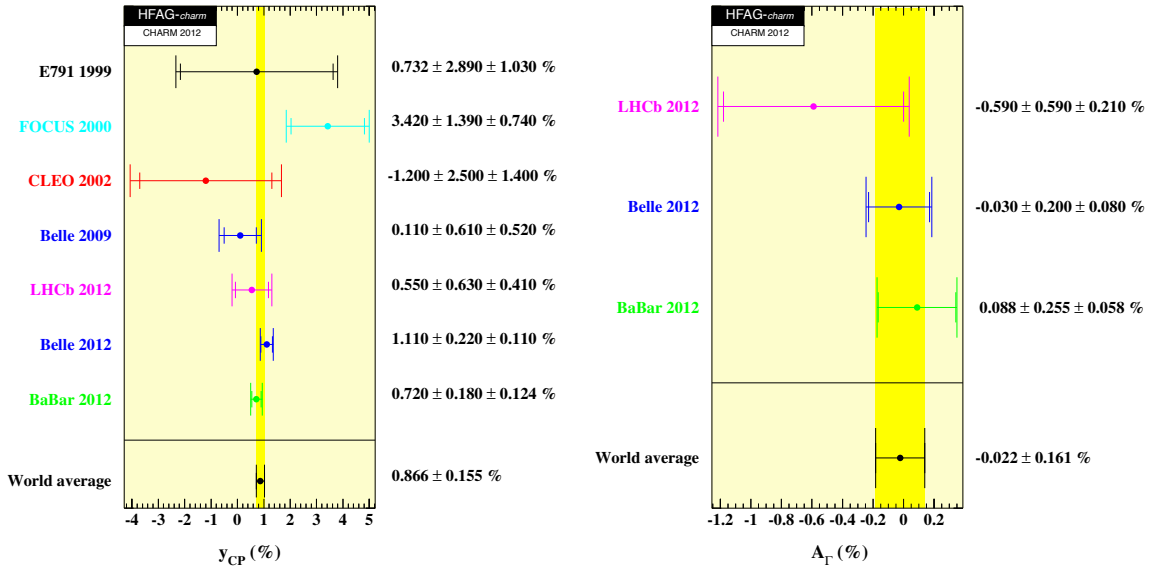


Figure 1.11: The HFAG averages for y_{CP} (left) and A_Γ (right) from May 2012. From [41].

1.3.5 Heavy Flavour Averaging Group fit

The fit of direct and indirect *CP* violation, using ΔA_{CP} and A_Γ measurements prior to the release of the results presented in this thesis is shown in Fig. 1.12.

The fit of the parameters of indirect *CP* asymmetry, the magnitudes of p and q and the phase between them is shown in Fig. 1.13.

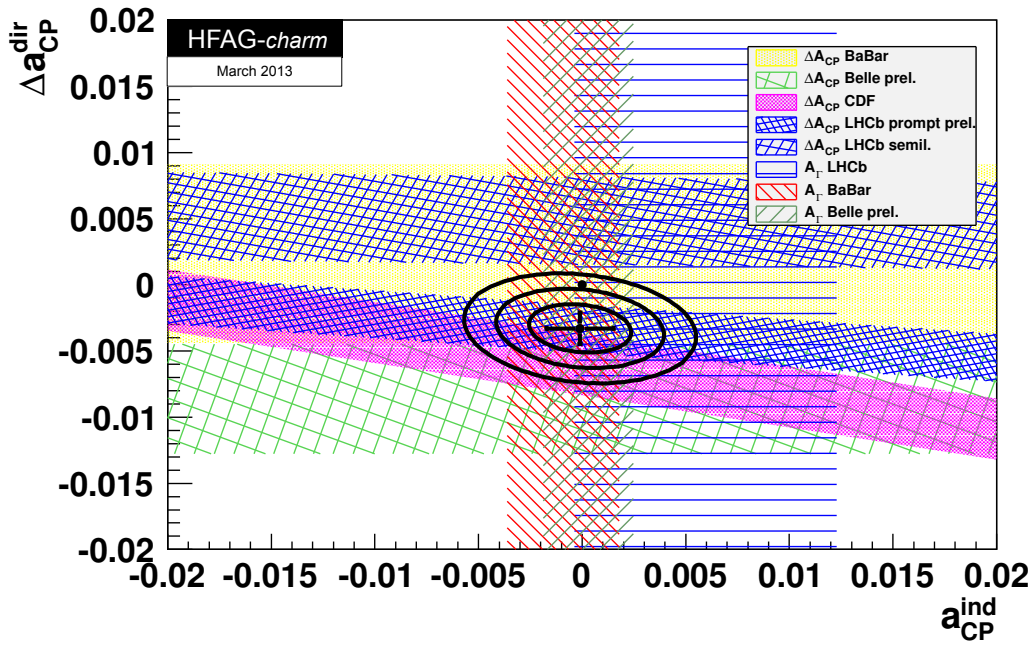


Figure 1.12: The HFAG fit for direct and indirect *CP* violation in charm from March 2013. Shown is the direct *CP* violation difference between K^+K^- and $\pi^+\pi^-$ final states as the LHCb prompt analysis does not unfold the individual asymmetries. From [41].

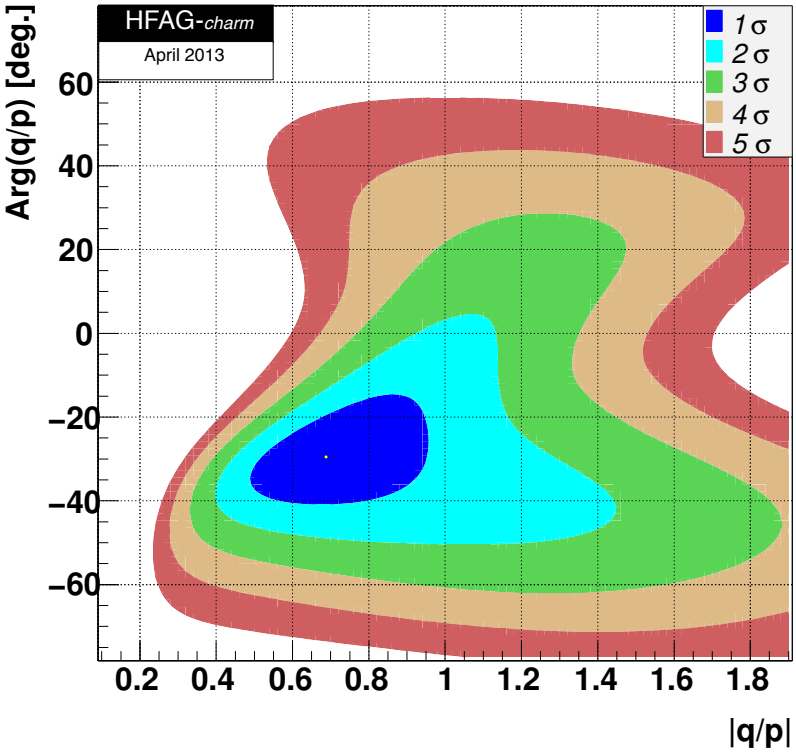


Figure 1.13: The HFAG fit for the magnitudes and phase of the mixing parameters p and q. From [41].

Chapter 2

LHC and LHCb

The measurement of the charm CP violation observables y_{CP} and A_Γ introduced in the previous chapter are in this instance measured by studying the decay of charmed hadrons produced in proton-proton collisions. In this chapter the experimental apparatus used will be introduced.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a hadron synchrotron at the European Organisation for Nuclear Research (CERN) on the Franco-Swiss border to the north-west of the city of Geneva. Its 27 km circumference occupies the tunnel that was formerly used by the LEP collider until 2000. Construction of the LHC was approved in 1994 and proton beams first circulated in 2008. In 2010 data from collisions was first taken with the subsequent running in 2011 and 2012 providing the bulk of the data for physics analyses. In 2013 the machine was shut down for repairs and upgrades, bringing an end to the Run 1 programme.

The LHC is designed to accelerate and collide beams of protons or lead ions; data has been taken with proton-proton, lead-lead and proton-lead (in both directions) interactions. The design energy of proton collisions is 14 TeV, an order of magnitude greater than that managed by the TEVATRON [76] which was previously the most powerful collider to have been built and operated. Thus far the maximum proton collision energy was that for the 2012 data taking of 8 TeV.

The accelerator contains eight straight sections; in four are situated the large experiments (ALICE, ATLAS, CMS and LHCb). The RF acceleration, beam dump and beam collimation are contained in the others.

2.1.1 Running conditions for 2011

In 2011 the proton collision energy was 7 TeV with 50 ns bunch spacing. The instantaneous luminosity at LHCb varied over the year but had reached $4 \times 10^{32}\,\mathrm{cm}^2\ \mathrm{s}^{-1}$ with about 1.8 visible proton interactions per bunch crossing. In total $1.1\,\mathrm{fb}^{-1}$ of integrated luminosity was recorded by the detector.

The instantaneous luminosity at LHCb is somewhat lower than for the general purpose detectors. The performance of the vertex detector is somewhat degraded by additional reconstructed primary vertices; it is estimated that the vertex resolution decreases by 5–10% for each additional primary vertex [77]. Therefore a lower luminosity is favourable for the LHCb physics programme with its focus on high precision. Because of this the luminosity is stable over the course of a fill at LHCb in contrast to the exponential decreases seen at the other experiments due to the 'luminosity levelling' [78] implemented at the LHCb interaction point.

2.2 LHCb

A full description of the detector and its components can be found in [79]. Here a brief introduction to the LHCb hardware and some relevant software will be presented. Those parts that are particularly pertinent to the subjects of this thesis will be expounded upon in a little more detail. These are the vertex locator and the offline reconstruction. Some attention will be given to their performance over Run 1. The interested reader may find more complete descriptions in the relevant design reports, performance papers and conference presentations; where appropriate the citations for these are given.

LHCb is situated at point 8 of the LHC ring adjacent to Geneva airport and at a depth of 100 m. The experimental cavern was previously occupied by the LEP experiment DELPHI; that piece of history remains underground having been moved to one side. Measuring $21 \times 13 \times 10$ m at its extremities, weighing 5600 tonnes and with a capital cost of 75 MCHF, LHCb is the smallest of the four large LHC experiments [80]. The collaboration comprises roughly 700 members.

The detector itself is usually described by some amalgamation of the words forward, spectrometer (occasionally with a bonus 'single arm'), detector, precision, (dedicated) heavy flavour and experiment. Such descriptions sum up the nature of the experimental equipment in a succinct manner.

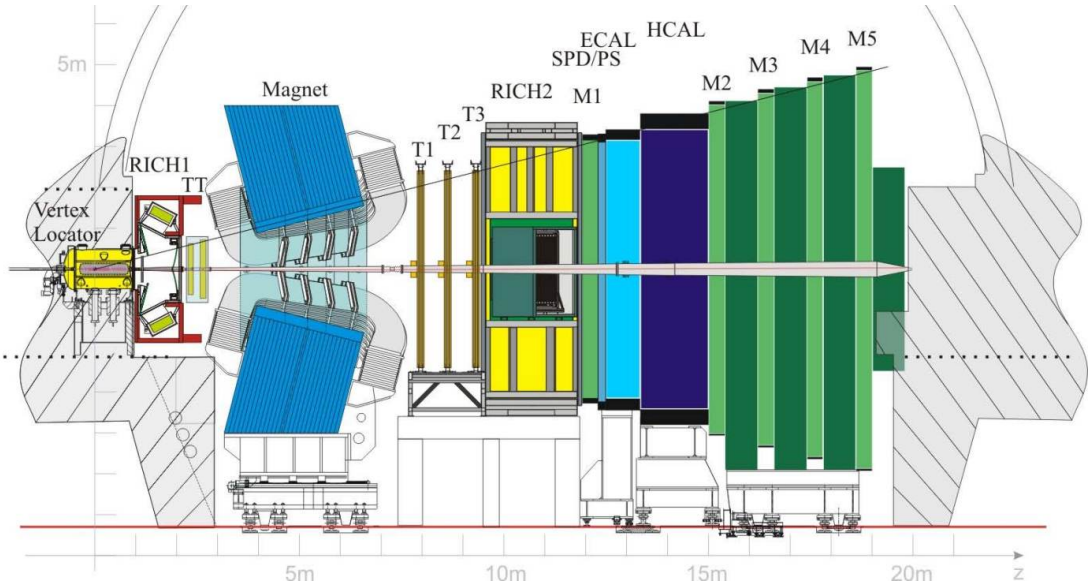


Figure 2.1: A cross-section diagram of the LHCb detector. Reproduced from [79].

2.2.1 Detector overview

The primary physics goal of LHCb is to study flavour physics, particularly with respect to the b quark. The roadmap document of 2009 [81] details six measurements that were expected to be made and have been; they include measurements of the CKM angle γ [82], measurements of B_s^0 mixing [83] and the discovery of the rare $B_s^0 \to \mu^+\mu^-$ decay [84] amongst others. The focus for all of these is very high precision from the detector and great statistical power due to the large event rates expected at a hadron collider. These are needed to study rare decays or processes with small amplitudes such as CP violation. Such requirements for accuracy were the motivation for the resulting design of the detector, which can be seen in Fig. 2.1.

The detector design is optimised for coverage of the region in which most b quarks are produced in p-p collisions. As shown in Fig. 2.2 this is mostly in cones around the beam pipe in the forward or backward directions ($\theta = 0$ or $\theta = \pi$). As these are the particles of interest only the forward region is instrumented, shown by the red region on the left of Fig. 2.2 or the red box on the right, unlike the hermetic general purpose detectors. Doing so allows for the equipment to be optimised for high precision. Of course a large amount of b quarks produced cannot be measured; the simulation of Fig 2.2 estimates about 25% of b (or $\overline{b}$) quarks produced are within the LHCb acceptance, compared with about 50% in ATLAS or CMS giving the general purpose detectors an advantage in statistical power (due to production only) of ~ 1.4. However the b quarks produced in the forward region are greatly boosted making them much easier to trigger on due to

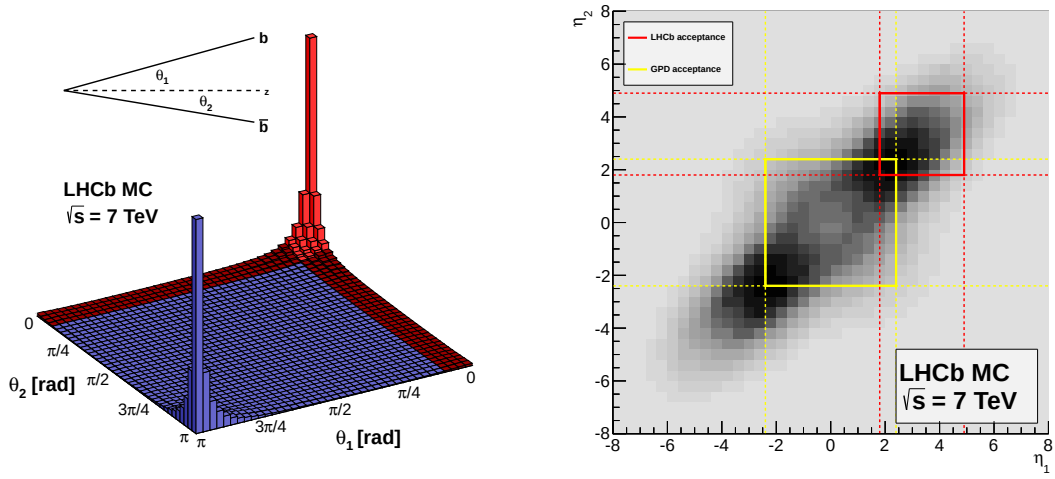


Figure 2.2: A simulation of the production of b pairs in 7 TeV proton-proton collisions. On the left the production rate as a function of the quark angles with respect to the beam is shown, with the LHCb acceptance highlighted in red. On the right the production is plotted as a function of the pseudo-rapidity of the two quarks with the acceptances of LHCb (the red box) and the general purpose detectors (GPD, ATLAS and CMS, in yellow) overlaid. Reproduced from [87].

the large decay vertex displacement. The acceptance of LHCb is defined by the pseudo-rapidity η * as $2<\eta<5$, compared to roughly $\eta<2.5$ for the general purpose detectors. Therefore LHCb covers a unique pseudo-rapidity range that can complement ATLAS and CMS measurements (for example in hadron [85] and electroweak boson [86] production) in addition to its precision flavour physics programme (see Fig. 2.3).

The protons are brought to collision on the left of Fig. 2.1 in the vertex locator at the edge of the LHC tunnel. The direction to the right of the interaction point (towards the rest of the detector) is henceforth referred to as forward or downstream; to the left and into the cavern wall is backward or upstream. The tracking system comprises of the Tracking Turicensis (TT) and the stations T1, T2 and T3. The two Cherenkov detectors, RICH1 and RICH2, together provide good hadronic particle identification capabilities. There are two calorimeters, ECAL and HCAL, which provide energy estimates and with the pre-shower detectors (PS) and scintillating pad detector (SPD) can be used to identify electromagnetic and neutral particles. The ensemble of sub-detectors is completed by the five muon stations, M1–M5, at the rear of the detector. There is also a conventional dipole magnet (as opposed to the superconducting magnets of the general purpose detectors) to allow for momentum estimation .

*The pseudo-rapidity is defined as $\eta \equiv -\log\left(\frac{\theta}{2}\right)$, where θ is the angle between the particle trajectory and the beamline.

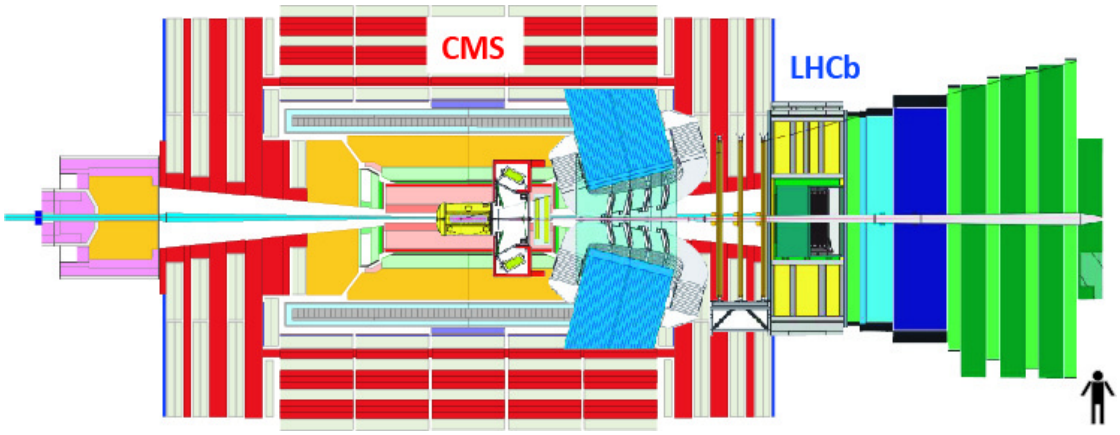


Figure 2.3: The CMS and LHCb experiments, one on top of the other, showing their respective geometries. Reproduced from [88].

LHCb uses a right-handed coordinate system with the z axis aligned with the beam pipe and pointing from the interaction point to the rest of the detector. The z–x plane is horizontal with the orthogonal y axis pointing vertically upwards.

2.2.2 VErtex LOcator

The proton collision point is surrounded by the VErtex LOcator (VELO), a silicon strip detector for reconstructing production and decay vertices of short lived particles. It is unique at the LHC for its exceptional accuracy which is due to some novel design features.

A schematic of the VELO sensor layout can be seen in Fig. 2.4. It is composed of two halves, each half containing 21 sensing modules arranged orthogonally to the beam direction and slightly offset from each other longitudinally. The detecting surface is 8 mm from the beam at the innermost part and it extends out to 42 mm. Such proximity gives excellent vertex resolution. During LHC filling, the two halves of the VELO can be retracted by ∼3 cm to avoid potential damage due to the increased dynamic aperture of the machine; once stable beams have been achieved they are brought together to overlap slightly, giving complete azimuthal coverage.

Each module contains two silicon strip sensors, one measuring the radial distance of the particle hit from the beam, the r coordinate, the other ascertains the azimuthal angle of the hit, ϕ. The r strips are concentric circles whilst those on the ϕ sensors are at an angle to the radial direction. The pitch of the strips varies between 40–100 µm from centre to edge. The sensors themselves are of a radiation hard n-type bulk, with n^+-type strips and p back (there is one station with n-on-p sensors). The sensors are aligned behind a thin foil to prevent picking

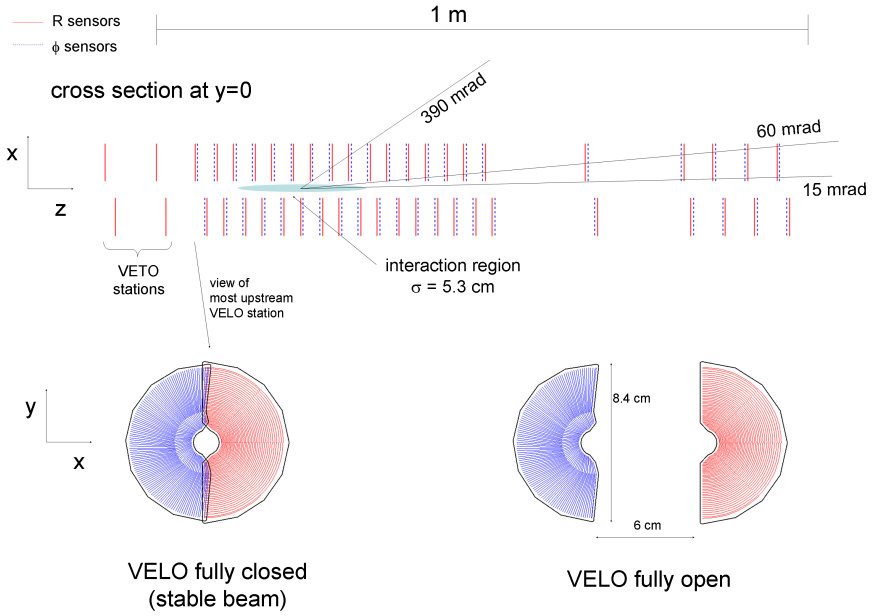


Figure 2.4: A diagram of the VELO layout. The top shows the arrangement of the 42 modules into two halves with r and ϕ sensors on each one. On the bottom the open and closed positions of the VELO halves can be seen. Reproduced from [79].

up the RF signal of the proton beam and separate the LHC vacuum from the detector as shown in Fig 2.5.

The reconstruction of hits and the measurement of the hit resolution of the sensors is described in detail in chapter 3 where it is measured and parametrised. For the moment it is sufficient to note that such resolution depends on the pitch of the strips at which the particle has traversed the sensor and its angle with respect to the sensor strips. Fig. 2.6 shows the hit resolution for r strips for two ranges of angles as a function of the strip pitch. The best resolution at the smallest strip pitch is 4 µm [77].

The primary vertex (PV) resolution is displayed in Fig. 2.7 and that of the impact parameter (IP) [†] in Fig. 2.8. A good measurement of the IP is important as LHCb primarily seeks displaced vertices in order to identify long lived hadrons (B and D). To that end impact parameter constraints are typically used as it is a quantity that can be quickly measured and cut on. The impact parameter resolution is less than 30 µm for high momenta tracks. Furthermore the VELO achieves excellent proper-time resolution of 50 fs [91].

The VELO has been extensively studied in Run 1 to characterise the effects of radiation damage on the silicon sensors [92]. Type inversion of the irradiated

[†]The IP is defined as the shortest perpendicular distance between a track and the reconstructed collision point.

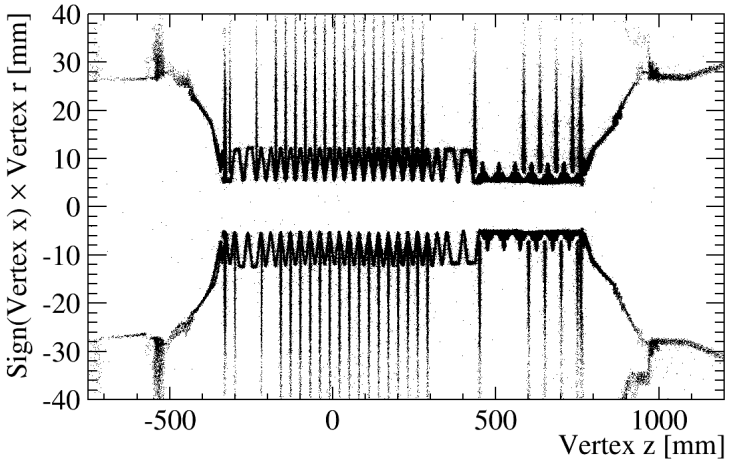


Figure 2.5: The VELO layout, imaged by reconstructing the primary vertices of particles produced by interactions with detector material from beam-gas collisions. The exceptional accuracy of the VELO makes such a visualisation possible. Taken from [89], see [90] for details.

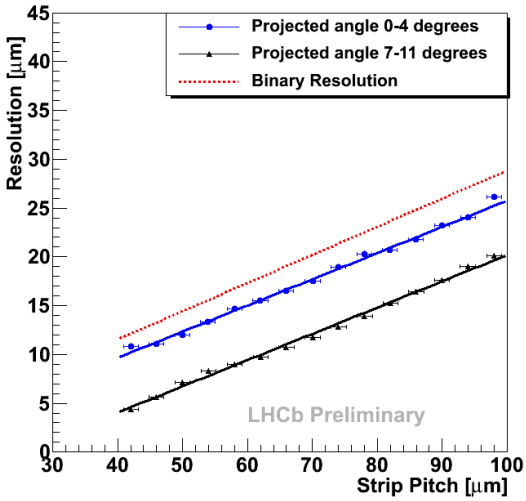


Figure 2.6: The projected hit resolution of the VELO r sensors as a function of strip pitch for two bins of track angle with respect to the sensor. The binary resolution is also shown. Reproduced from [77].

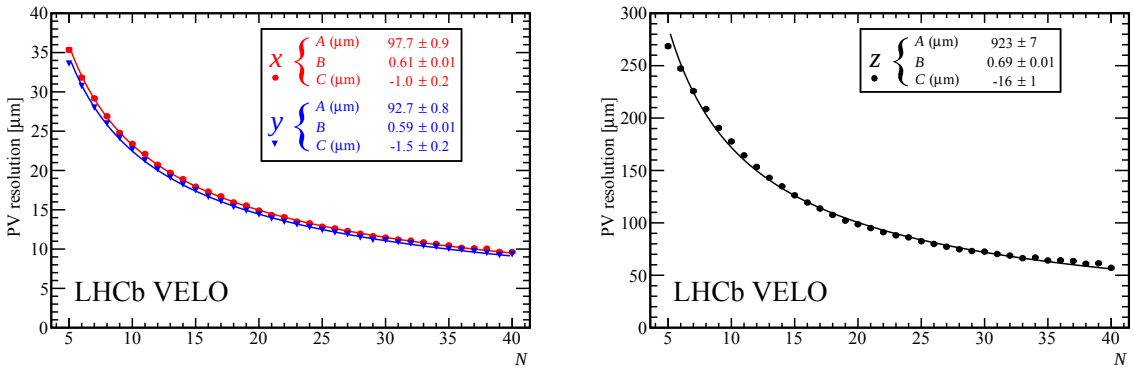


Figure 2.7: The primary vertex resolution of the VELO as a function of track multiplicity N. On the left is the resolution in the horizontal, x (red), and vertical, y (blue), coordinates. The right plot shows the resolution in z along the beam direction. The distributions have been fitted with a curve of the form $\sigma_{PV} = \frac{A}{N^B} + C$. Reproduced from [77].

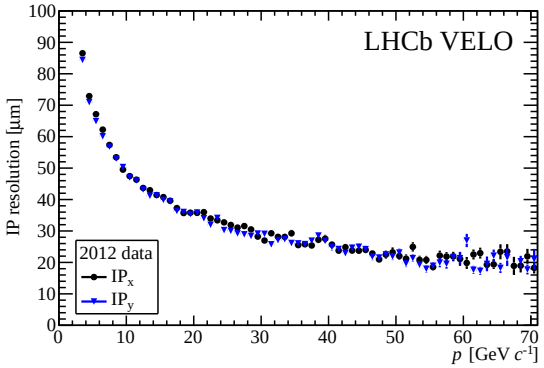


Figure 2.8: The impact parameter resolution in x (black) and y (blue) as a function of the track momentum p. Reproduced from [77].

silicon has been observed as well as a coupling between the sensor and the metal lines reading out the strips. A more complete discussion is given in chapter 3.

2.2.3 Tracking

Aside from the VELO, charged particle tracking is performed by the four tracking stations: TT, T1, T2 and T3. The first is upstream of the magnet, the others just after it. The three large downstream stations each contain an inner and outer tracker (IT and OT respectively). The inner tracker is a silicon detector, the outer gas filled straw tubes. The highest densities of tracks in events at LHCb are to be found adjacent to the beam pipe so in this region the higher resolution of a silicon detector is required. Further away from the beam pipe the larger area can be more cost effectively covered by straw tubes.

Each OT station contains four module layers; two are arranged vertically whilst the other two are arranged at a ± 5 deg angle with respect to the vertical in order to provide stereo resolution. The modules themselves are comprised of two layers of 4.9 mm diameter straw tubes. The single hit resolution of the outer tracker has been measured to be 205 μm [93].

The TT and IT are silicon strip detectors of 200 μm pitch. The TT covers the full LHCb acceptance, whilst the IT covers the central regions of T1, T2 and T3. As before, the silicon detectors stations are each comprised of four layers, arranged in the same manner as the OT with orientation of the strips of two of them at an angle with respect to the vertical. The hit resolutions of the TT and IT detectors have been measured to be about 52.6 μm and 50.3 μm respectively [93].

2.2.4 Cherenkov detectors

The majority of the differentiation between the particle identification hypotheses of hadrons (pions, kaons and protons) is achieved with the two Ring Imaging Cherenkov detectors, RICH1 and RICH2. The former is situated adjacent to the VELO, the latter is between the tracking stations and calorimetry. Together they cover the momentum range 2–100 GeV/c.

RICH1 is optimised for lower momenta particles and so covers the entire LHCb acceptance. Two radiators are used in this detector. Gaseous C_4F_{10} provides pion-kaon separation down to about 9 GeV/c (the threshold momentum for kaons to emit Cherenkov light). The momentum range below that is catered for by an additional layer of silica aerogel.

The second detector uses CF_4 as the radiation medium. As high momentum tracks tend to have smaller angles with respect to the beam pipe RICH2 covers a

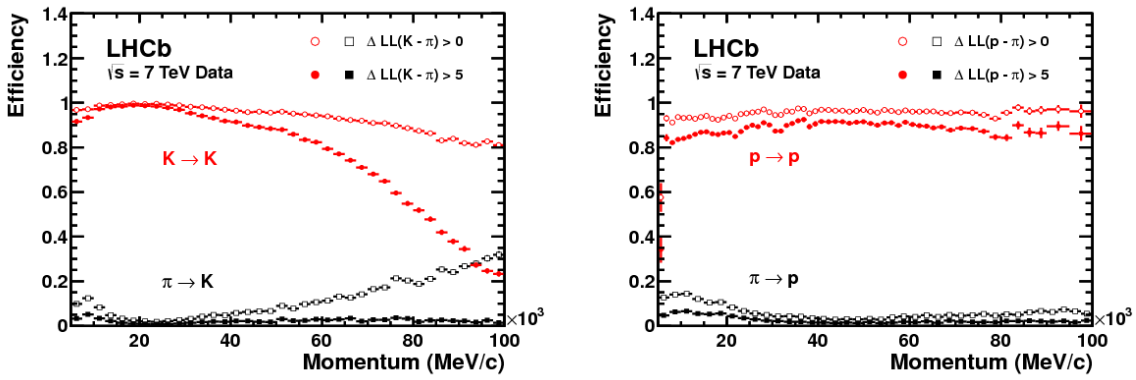


Figure 2.9: The PID performance for kaons (left) and protons (right) as a function of momentum. Two different $\Delta \log \mathcal{L}$ conditions are shown along with the pion mis-ID rates. Reproduced from [95].

smaller area than the total acceptance of the experiment (up to 120 mrad angle compared to 300 mrad of LHCb acceptance).

In both instances the emitted Cherenkov light is focused onto sensing planes of hybrid pixel detectors (HPD). These are situated outside the LHCb acceptance in order to reduce the material budget.

The extraction of particle identification is done via a global likelihood method [94]. For a single track the angle of the emitted Cherenkov light is proportional the particle speed. Combining the speed with the measured momentum leads to an estimate of the rest mass of the particle. For the large multiplicity of tracks in a typical event a likelihood is constructed for the event with HPD hits being associated with each track assuming all particles are pions. This likelihood is then maximised by varying the particle identification hypothesis (pion, kaon, proton, electron, muon) of each track in turn. The results is that for each track one can calculate the relative change in the logarithm of the overall event likelihood through changing the particle hypothesis from pion to some other possibility, the $\Delta \log \mathcal{L}$.

The PID performance can be measured on single tracks where kinematic constraints on particle combinations are sufficient for PID. Fig. 2.9 shows the measured kaon and proton PID performances as measured with data.

2.2.5 Calorimetry

The calorimetry is split into three parts, a pre-shower detector (SPD/PS), electromagnetic (ECAL) and hadronic (HCAL) calorimeters. Together they allow for energy estimates for use in the trigger, identification of electromagnetic and neutral particles and some reconstruction capabilities for neutrals. This

information is particularly used in the hardware trigger. All of these components contain regions of differing granularities, the highest being close to the beam pipe, reflecting the higher occupancy in those areas.

The first part is the SPD/PS, consisting of two layers of scintillating detectors (the SPD and PS in order of increasing z) on either side of a lead plate of 2.5 radiation lengths thickness. Scintillating light from each pad is collected and transferred to a photo-multiplier tube that reads out the signal.

The ECAL is of a sampling type with alternating planes of 2 mm thick lead and 4 mm thick plastic scintillating tiles. Again the scintillated light is collected by scintillating fibres and then read out at the back of the ECAL with photo tubes. In total there are 66 layers of scintillator and lead giving a depth of 42 cm, or 25 radiation lengths. Such a thickness should mean that high energy photons are completely contained within the ECAL, allowing for a good energy measurement, a prerequisite for the reconstruction of a π^0. The mass resolution for neutral pions has been measured to be $\sim 30\,\mathrm{MeV}/c^2$ at high momentum[93].

Finally the HCAL is another sampling device of alternating layers of steel absorber and scintillator. In the direction of the beam the length of the layers are one interaction length. In total the HCAL is 5.6 interaction lengths in thickness. Hadronic showers do not need to be completely reconstructed and the information gained is sufficient for the first level of the trigger to select high p_{T} candidates.

Seeing where in the calorimeter the shower develops allows for particle identification. Electrons and photons both shower in the lead between the SPD and PS, the first detector allowing for charged particle discrimination. Hadrons will only shower significantly in the HCAL. The PID performance using only the calorimeter information is shown in Fig. 2.10. The identification efficiency is $\sim 92\%$ with a mis-ID rate of $\sim 4.5\%$ for $\Delta \log \mathcal{L}^{\mathrm{CALO}}(e-h)>2$; with additional information from the RICH detectors these figures can be improved to 97% and 2% respectively [93].

2.2.6 Muon stations

At the back of the detector can be found the muon detectors which provide muon identification for the trigger and offline analysis. The first station, M1, is situated before the calorimeters with the remaining stations afterwards. M1 is used only in the first level muon trigger where a fast estimate of p_T is required. Tracks are formed in M2–M5 and then matched to a hit in M1. The direction of the track provides a fast estimate of the p_T and the extra hit in M1 improves the resolution over using just M2–M5 by $\sim 25\,\%$ [96]. In the high level trigger and offline a full track reconstruction is performed whereby the muon tracks are matched to the rest

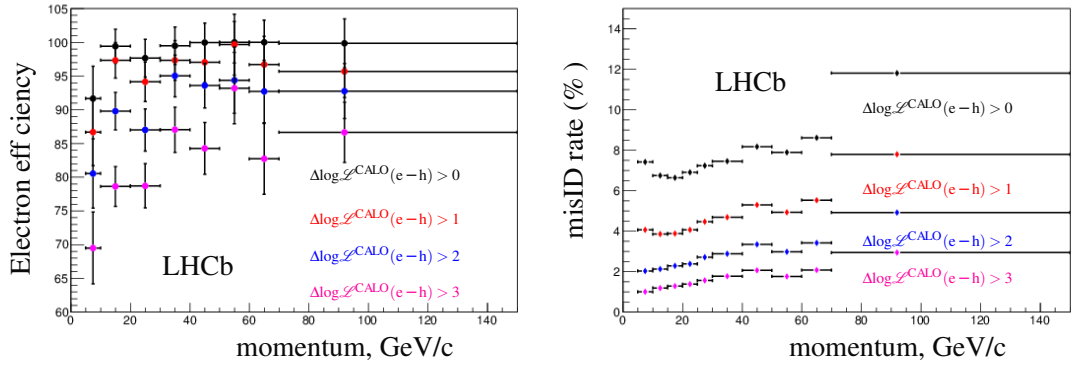


Figure 2.10: The PID efficiency (left) and mis-ID rate (right) of the calorimeters. Different $\Delta \log \mathcal{L}$ cuts using only the calorimeter information are shown. Reproduced from [93].

of the tracking systems, so the hit in M1 is not utilised.

The detectors cover a total area of 435 m ^{2} using multi wire proportional chambers (MWPC) for the most part; the innermost region of M1 is instrumented with gas electron multipliers (GEM). In between the stations M2 to M5 are 80 cm thick steel absorbers such that a particle passing through the whole of the calorimeter and muon systems would have traversed about 20 interaction lengths.

The first three stations (M1–M3) have high granularity and can contribute to the tracking and momentum measurement of the muon candidate. All the muon stations contribute to muon PID via a simple binary ('IsMuon') selection and a likelihood estimate. The first accounts for how many stations a track traverses. Those with momenta below 6 GeV/c only need to leave hits in M2 and M3 to be considered a muon; higher momenta tracks need to be seen in M4 and M5.

A muon likelihood can then be assigned to a track depending on how well the hits in the muon stations can be linearly extrapolated to the track in the forward tracking. The closer they are to a linear track the higher the likelihood (see [97] for details). Fig 2.11 shows the efficacy of the muon identification using information only from the muon systems. Combining these likelihoods with those from the two RICH detectors and the calorimeters further improves performance.

2.2.7 Magnet

The LHCb magnet is a normal conducting dipole magnet of integrated field 4 Tm. The field is orientated vertically so the bending plane is horizontal. The polarity can be reversed, with the field direction up or down. The field map can be seen in Fig. 2.12.

A normal conducting magnet benefits from reduced cost and increased safety

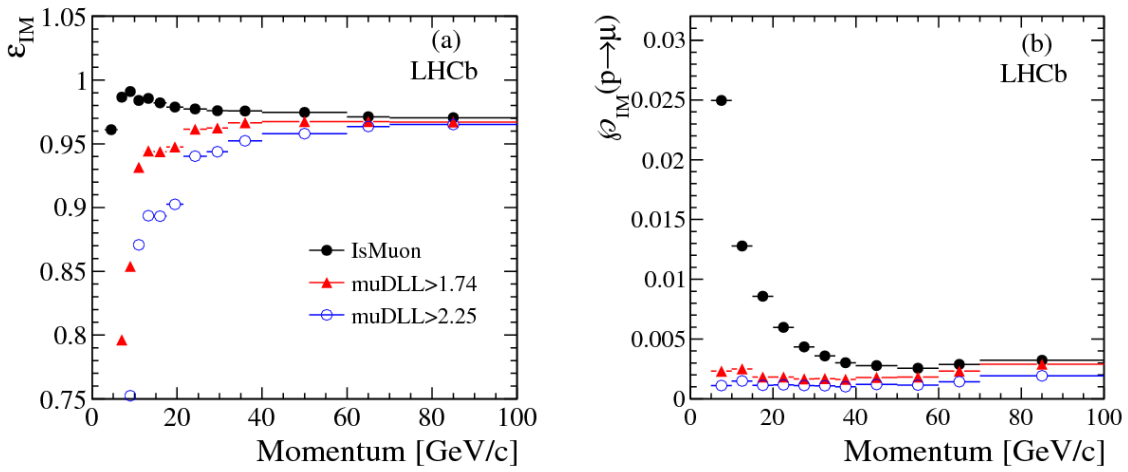


Figure 2.11: The identification efficiency (left) and mis-ID rate (right) of muons using the muon chambers. The black circles represent the binary selection, whilst the two colours show the improvements achieved with the muon likelihoods. Reproduced from [97].

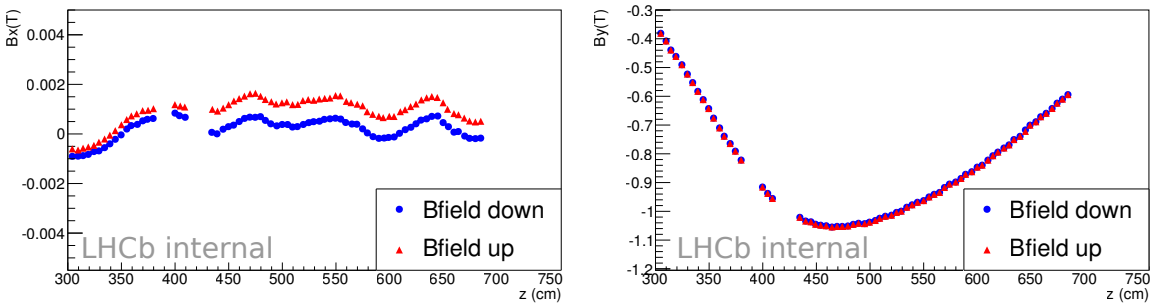


Figure 2.12: The measured field of the LHCb dipole magnet for the 2011 data taking period. Shown is the strength of the field in the x (left) and y (right) directions, both perpendicular to the beam pipe, as a function of the z position. The sign of the magnet up polarity measurements have been inverted. The measurements are made at fixed x and y, slightly offset from the beam axis. Reproduced from [98].

compared to a superconducting apparatus. Additionally the magnetic field can be ramped up and its polarity reversed quickly, which is an important feature. As positive and negative particles are deflected in opposite directions any differences in performance between the left and right sides of the detector could lead to extra charge detection asymmetries. Performing analyses averaged over data taken with both magnet polarities should cancel such effects.

A momentum resolution of $\frac{\delta p}{p} \sim 0.5\%$ is achievable for high momentum tracks (100 GeV/c) [93]. It is expected that the effect of the particle scattering in the detector material is the limiting factor in the resolution. Both the detector material and the accuracy of the measured field map are included in the reconstruction algorithms.

2.2.8 Trigger

The LHCb detector is read out at a rate of ∼2–5 kHz, a scaling of the bunch crossing frequency (20 MHz for Run 1) of order 10^{-4}. Such a reduction is achieved by the LHCb trigger. It is worth noting that LHCb has a significantly greater readout rate compared to the general purpose experiments (reading out at ∼500 Hz) as befits their respective physics goals.

The trigger is arranged in three parts: L0, HLT1 and HLT2. The first is a hardware trigger due to the strict limitations on the processing time available to it. L0 reduces the rate to 1 MHz. The high level trigger (the collective term for HLT1 and HLT2) is implemented in software and further reduces the output rate to 43 kHz and then successively to the 3 kHz that gets read out and stored for further offline analysis.

L0 seeks to make decisions based on information that is immediately available and does so within 1 μs. It looks for a hard scatter with high momentum transfer by seeking electrons, hadrons or muons with large p_{T}. These manifest themselves as either large deposits in the calorimeter or hits on the muon chambers. The calorimeter trigger fires on the single largest energy deposit, whilst the muon trigger takes either a single high p_{T} track or the sum of two above a threshold. The reduced rate of 1 MHz allows the rest of the detector to be read out; their rates being limited by their front end electronics.

Those events that fire the L0 trigger are passed onto the first software stage. It is here that a fast pattern recognition is carried out and tracks are first formed and fitted, to which selections may be applied. For the most part one seeks long-lived particles whose decay products will have a large impact parameter (IP) with respect to the collision point. Therefore in HLT1 tracks can be quickly constructed and their IP cut on (in practice the χ^2_{IP} is used) to provide a fast

method for preferentially selecting long-lived hadrons (e.g. B or D) without having to reconstruct the particle in question and its decay vertex, a process that takes a significant amount of time.

From HLT1 events pass to HLT2 which can perform a full reconstruction, although information from the Cherenkov detectors is generally not used due to time constraints. Using this full reconstruction decays of heavy particles can be reconstructed from their decay products and more stringent selections applied, for example on the invariant mass. As well as the exclusive triggers operating purely on a cut based method, topological triggers accounting for missing transverse momentum and multivariate selections may be utilised.

Each layer of the HLT contains several selection algorithms (lines) optimised for the analysis for which they have been created. Together they form the trigger configuration described by a Trigger Configuration Key (TCK). Being implemented in software gives the HLT great flexibility and allows it to be optimised throughout the data taking. It also allows for the trigger to be re-run offline, either on data or simulation, in exactly the same configuration as for data taking. This is important for assessing trigger efficiencies or potential biases due to the trigger (as will be outlined in section 4.2.3).

The trigger decisions at each level are split into three groups depending upon the desired signal candidate. If the signal candidate is what caused the event to pass a particular trigger line it is referred to as TOS (triggered on signal) with respect to that line. On the occasions that some part of the event other than the signal candidate fired the trigger it is referred to as TIS (triggered independently of the signal). If the trigger is fired by both the signal candidate and another part of the event it is referred to as TISTOS. Knowing what caused an event to pass the trigger is essential for understanding potential trigger induced biases or asymmetries.

In 2012 the triggering capabilities of LHCb were further improved by the introduction of a 'deferred trigger' [99]. From the L0 output 25% of events were stored in a buffer. At the end of an LHC fill and before the start of the next the buffered events were processed by the HLT at a time when the computing farm would otherwise be idle. Thus the HLT could process a greater amount of information with the same accuracy in the reconstruction.

2.2.9 Offline selection and analysis

The data that passes the trigger is stored and then reconstructed offline. To the resulting candidates further selection can be applied after the trigger for each analysis. More stringent PID information becomes available offline as well as a

slightly improved reconstruction due to the greater time available.

Each analysis can select decays using the standard LHCb software framework which defines all the detector objects, the decay products and reconstructed combined particles in a standard set of objects. Each analysis creates their selections which are combined and then the raw data from the detector is processed in one go with the results being written out to separate streams and made available to the analysts. This is performed centrally for greater efficiency and to negate the need for individuals to deal with enormous data sets. This processing is periodically repeated with updated reconstruction and selection algorithms. Subsequently individuals can apply further selections for their particular needs.

2.3 Charm at LHCb

The LHCb experiment has proved a fruitful environment for charm physics analyses. Those qualities of the detector that were optimised for b physics serve the charm sector just as well. In particular the high resolution of the VELO is important for charm due to the relatively short flight distances of charmed mesons; of the order of mm, instead of cm for B particles. The excellent impact parameter resolution allows for efficient triggering on these comparatively short lived particles whilst the proper time resolution of 50 fs [77], of the order of a tenth that of the D^0 mean lifetime, facilitates precise time dependent measurements (a prerequisite for the analyses presented in this thesis).

As well as the qualities of the detector, LHCb benefits from great statistical power. The charm production cross section in the region of the LHCb acceptance is roughly 20 times that for b particles [100, 101],

$$\sigma_{b\bar{b},acc}(7\,\text{TeV}) = (75.3 \pm 14.1)\,\mu\text{b}$$
$$\sigma_{c\bar{c},acc}(7\,\text{TeV}) = (1419 \pm 134)\,\mu\text{b},$$

providing LHCb with a very large sample of charm decays for analysis. Additionally charm originating from B decays (usually via a semi-leptonic decay with a muon, the lepton providing a flavour tag) provides a completely independent sample with which to perform analyses. Such measurements have less statistical power but provide cleaner data samples which are easier to trigger on. In turn the trigger has a high output rate of $\sim$2 kHz for charm.

Many of the LHCb achievements with regards charm flavour physics have been highlighted already in chapter 1. In addition to these LHCb data is making new contributions to charmed hadron spectroscopy [37, 36] and setting stringent limits for searches of very rare decays[102, 103].

Chapter 3

VELO cluster resolution

The VErtex LOcator (VELO), situated at the interaction point, is an exceptionally precise detector, capable of accurate measurements of production and decay vertices and track impact parameters [77]. That excellent performance is due in part to the ability to find and fit tracks accurately. For the purposes of track fitting the uncertainty on the measurement of the position of each hit in the detector needs to be known. If the relative uncertainty on the hit position is incorrect then the optimum tracking performance will not be obtained. If the absolute error is false then the total χ^2 of the fit, along with the χ^2 of associated quantities, will also be incorrect. The error is ascertained by means of a parametrisation that reflects the resolution of the detector. The estimation used throughout Run 1 was based on simulation studies where the resolution was tuned to test-beam data. The aim of the work presented here is to measure the hit resolution of the VELO using collision data and update that parametrisation.

In this chapter the track fitting strategy of the LHCb VELO will be described, highlighting the importance of knowing the true error of the hit measurements and how to extract the unbiased residuals from the track fit. The VELO sensors will be described in some detail with the physical parameters pertinent to the resolution introduced. A method to measure the resolution will then be outlined followed by a check with simulated data. The resolution for a selection of short data-taking periods throughout Run 1 have been measured, with unexpected results. They will be presented with some investigation as to the possible causes of the trends seen. Finally the effect on the track fit and the implications for physics analyses will be commented on.

3.1 Track fitting

One can never know precisely the true hit positions in the sensors from a particle track; the individual measurements have an inherent uncertainty. Instead one can

estimate the true particle path using the measurements from all the sensors. The Kalman filter [104][105] seeks to do so via a recursive method where successive hits are taken into account. This strategy is an alternative to a global track fitting method (for example a global minimum χ^2 fit). Such an approach has the advantage of being comparatively quick and computationally straightforward to carry out; there are no large global covariance matrices to calculate. It is also excellent for dealing with 'noisy' data as one finds with the hit measurements in a track fit. Assuming the uncertainties on the hit positions follow a normal distribution then the Kalman filter is well-suited to extracting the true hit positions from the noise of the hit uncertainties.

3.1.1 Kalman filters

The Kalman filter formalism is rather standard fare and can be found in most textbooks on signal processing, an application for which it was initially developed. Indeed the method can be applied to any processes that can be modelled in an iterative manner. Here the most pertinent expressions are outlined in the language of a track fitting algorithm. The process is carried out in three steps: *prediction*, *filtering* and *smoothing*.

The true description of the track at the sensor labelled z is given by the state vector $\mathbf{x}_z$. The state vector contains all the relevant information about the track; its hit coordinates, slope and momentum. The estimate of the state vector that is obtained by the Kalman filter is denoted $\hat{\mathbf{x}}_z$, with a covariance matrix describing the uncertainties, $\mathbf{P}_z$.

The first step, *prediction*, extrapolates the estimated measurement at the sensor at $z-1$ to the sensor at z via the transport matrix $\mathbf{F}_z$ (there should be nothing affecting the particle path between the sensors, such as residual magnetic field), producing the estimate at z given the estimate at the previous station

$$\hat{\mathbf{x}}_{z|z-1} = \mathbf{F}_z \hat{\mathbf{x}}_{z-1}. \tag{3.1}$$

The covariance matrix, is likewise transported in the prediction

$$\mathbf{P}_{z|z-1} = \mathbf{F}_z \mathbf{P}_{z-1} \mathbf{F}_{z-1}^{\mathrm{T}} + \mathbf{Q}_z, \tag{3.2}$$

where $\mathbf{Q}_z$ is the covariance of the noise associated with the transport. This noise is caused by scattering in the detector material (the success of the fit also relies on accurate descriptions of the detector material); it has a mean of 0. Having made the prediction the estimate is updated with the *measurement* at z, denoted $\mathbf{m}_z$. It is assumed that the uncertainties on both the prediction and the measurement are

Gaussian in nature, such that the resulting error is also a Gaussian. The mapping of the true state to the measurement is achieved via

$$\mathbf{m}_z = \mathbf{H}_z \mathbf{x}_z + \mathbf{v}_z, \tag{3.3}$$

with $\mathbf{v}_z$ representing the error on the measurement, again with mean 0 and covariance $\mathbf{V}_z$. The update of the prediction with the measurement at z to give the *filtered* estimate is described by

$$\hat{\mathbf{x}}_{z|z} = \hat{\mathbf{x}}_{z|z-1} + \mathbf{K}_z(\mathbf{m}_z - \mathbf{H}_z \hat{\mathbf{x}}_{z|z-1}), \tag{3.4}$$

with the updated covariance matrix

$$\mathbf{P}_{z|z} = \mathbf{P}_{z|z-1} - \mathbf{K}_z \mathbf{H}_z \mathbf{P}_{z|z-1}. \tag{3.5}$$

The Kalman gain matrix $\mathbf{K}_z$ gives a relative weighting to the prediction and measurement

$$\mathbf{K}_z = \mathbf{P}_{z|z-1} \mathbf{H}_z^{\mathbf{T}} (\mathbf{H}_z \mathbf{P}_{z|z-1} \mathbf{H}_z^{\mathbf{T}} + \mathbf{V}_z)^{-1}. \tag{3.6}$$

The filtered estimate contains information from all measurements prior to that z station. The filtering stage is shown in Fig. 3.1.

The *smoothing* stage includes all measurements in the track, from $z = 1$ to $z = N$, where N is the total number of hits on the track. In essence this is achieved by running the filter in reverse (extrapolating from $z + 1$ back to z) with the filtered estimate used instead of the measurement and taking something akin to the average of the two results

$$\hat{\mathbf{x}}_{z|N} = \hat{\mathbf{x}}_{z|z} + \mathbf{A}_z(\hat{\mathbf{x}}_{z+1|N} - \hat{\mathbf{x}}_{z+1|z}). \tag{3.7}$$

The matrix $\mathbf{A}_z$ is the smoother gain matrix

$$\mathbf{A}_z = \mathbf{P}_z \mathbf{F}_z^{\mathrm{T}} (\mathbf{P}_{z+1|z})^{-1}, \tag{3.8}$$

and the subsequent smoothed covariance matrix is

$$\mathbf{P}_{z|N} = \mathbf{P}_z + \mathbf{A}_z(\mathbf{P}_{z+1|N} - \mathbf{P}_{z+1|z})\mathbf{A}_z^{\mathrm{T}}. \tag{3.9}$$

The residual of the estimate before filtering is

$$\mathbf{r}_{z|z-1} = \mathbf{m}_z - \mathbf{H}_z \hat{\mathbf{x}}_{z|z-1}, \tag{3.10}$$

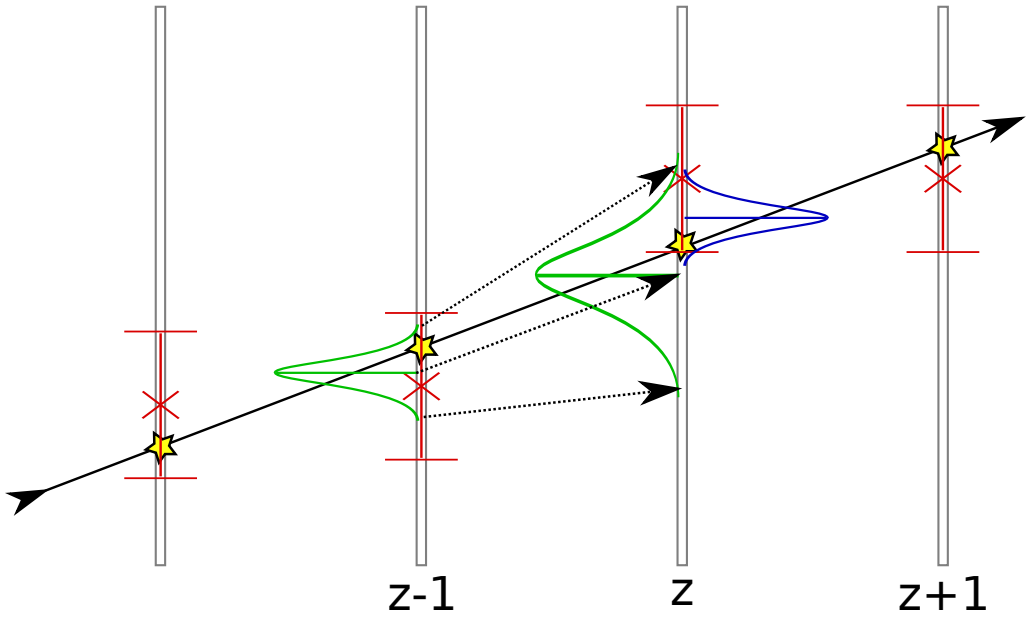


Figure 3.1: The Kalman filter. The true hit positions are denoted by the stars, with the measurements $\mathbf{m_z}$ and their associated uncertainties in red. The green curve at $z-1$ represents the filtered estimate $\hat{\mathbf{x}}_{z-1|z-1}$. The extrapolation of the hit and its covariance is shown up to the sensor at z, with the subsequent filtered result indicated by the blue curve, which will then be transported up to the $z+1$ station for the next stage of the recursive process.

with covariance

$$\mathbf{R}_{z|z-1} = \mathbf{V}_z + \mathbf{H}_z \mathbf{P}_{z|z-1} \mathbf{H}_z^{\mathrm{T}}. \tag{3.11}$$

After filtering, the residuals and their covariances are

$$\mathbf{r}_{z|z} = (\mathbf{I} - \mathbf{H}_z \mathbf{K}_z)\mathbf{r}_{z|z-1} \mathbf{R}_{z|z} = (\mathbf{I} - \mathbf{H}_z \mathbf{K}_z)\mathbf{V}_z. \tag{3.12}$$

Hence the inputs required in the procedure are a description of the detector material, the measured hit positions and the estimate of the hit errors. For measuring the hit measurement errors the distribution of the residuals is considered.

3.2 VELO clusters

The measurements of particle hits in the VELO sensors is displayed in Fig. 3.2. The particle path indicated by the line has released free charge carriers in the sensor bulk that are collected by the strips on the surface. The charge detected at each strip is indicated by the green bars. Clusters are initialised by strips that

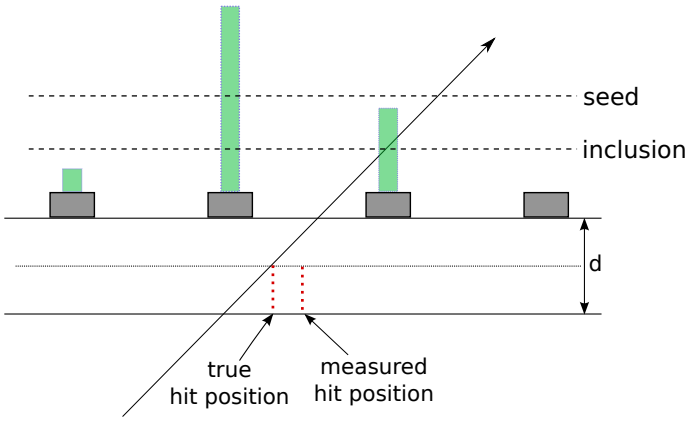


Figure 3.2: The construction of clusters in the VELO sensors. The charge collected on each sensing strip is indicated by the green bars. The resulting measured and true positions are indicated.

have charge greater than the *seeding* cut [77]. Having seeded the cluster the adjacent strips are examined and if their collected charge is greater than the inclusion cut they are added to the cluster. The resulting measured hit position is simply the mean of the strip positions x_i, weighted by their collected charge w_i

$$x_{meas} = \frac{\sum^{strips} w_i x_i}{\sum^{strips} w_i}. \tag{3.13}$$

As a consequence the resolution of clusters with two-strips is somewhat better than that for single-strip clusters due to the extra information about the hit position. A more complex treatment to calculate the measured hit position could be used, such as the η function described in [106]. This more accurately models the charge sharing between the strips. Such a function is not currently used in the reconstruction although it is under investigation. Preliminary results suggest that it is not expected to yield large improvements to the resolution.

3.2.1 Error parametrisation

To each measurement must be assigned an uncertainty for use in the Kalman filter. This depends on the pitch of the strips as the greater the distance between them, the larger the error. It also depends on the component of the track passing through the sensor that is perpendicular to the strips, described by the projected angle, α_p. This quantity is defined as the angle between the component of the track perpendicular to the strip direction in the plane of the sensor and the plane of the sensor as shown in Fig. 3.3.

For the sensor thickness d and distance travelled through the sensor

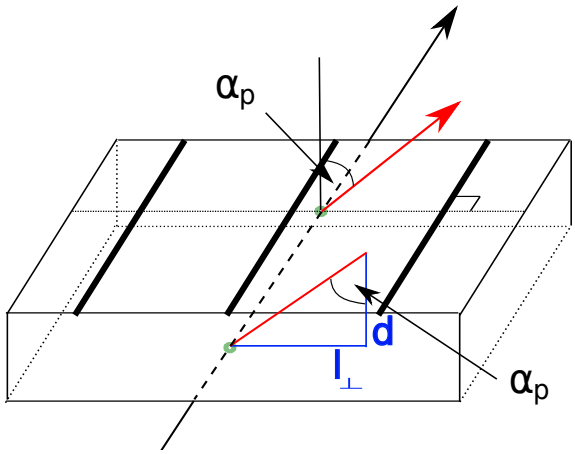


Figure 3.3: The projected angle α_p. d is the thickness of the sensor, $l_\perp$ is the distance the particle travels through the sensor in the direction perpendicular to the strip orientation.

perpendicular to the strip orientation $l_\perp$ the projected angle can be calculated as

$$\tan \alpha_p = \frac{l_\perp}{d}. \tag{3.14}$$

As strips on the ϕ sensors have a pseudo-radial geometry their clusters mostly have very small projected angles unless the track has originated far from the beam axis. The r clusters on the other hand have few hits at low or very large angles, instead covering a range that mostly covers the detector geometry for long tracks that traverse the tracking downstream of the magnet (they lie in the nominal acceptance $2 < \eta < 5$).

The dependence of the resolution on the projected angle is due to the charge sharing between the sensing strips. For small angles (more perpendicular to the sensor face) the charge is likely to be collected by a single strip giving a fairly poor resolution. Multi-strip clusters can be formed at small angles but they are mostly for hits equidistant from two strips. At larger angles there is a larger probability of forming multi-strip clusters so the resolution is improved. Single-strip clusters must be hits from much closer to a strip to prevent the charge sharing creating sufficient charge for a second strip to be included in the cluster. The resolution therefore improves with projected angle up to a minimum. At large angles the resolution worsens again as the charge deposited by the track is spread out over a large distance. The pitch and projected angle dependences of the resolution of the VELO clusters is shown in Fig. 3.4, where they have been studied with simulated data.

To facilitate the assigning of a measurement uncertainty to the hit the resolution is parametrised as a function of pitch and α_p. Therefore when the track

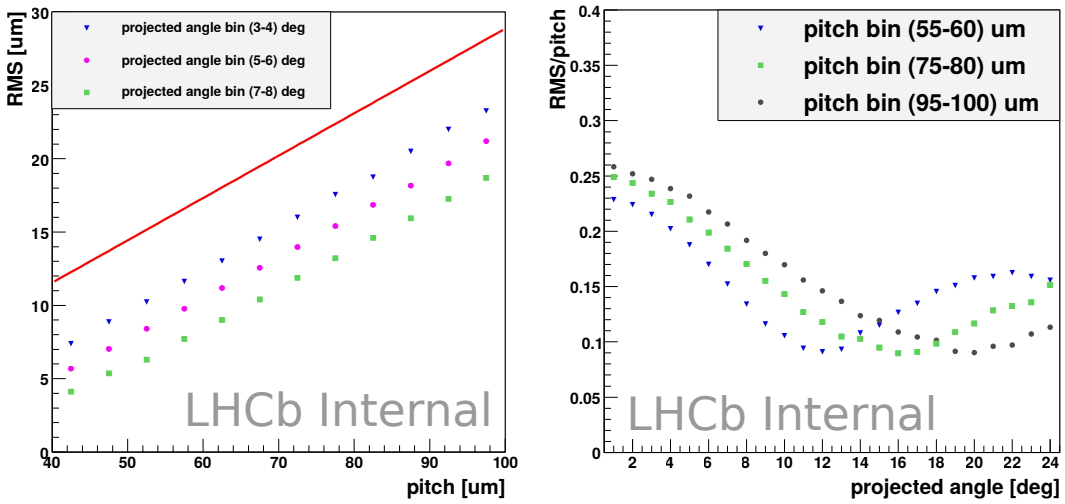


Figure 3.4: The dependence of the cluster resolution on pitch (left) and projected angle (right). The red line in the left plot indicates the binary resolution. These measurements were made on simulated data before the start of LHC Run 1. From [107].

is constructed an estimate of α_p and pitch is used to calculate the measurement error which is used in the track fit. The track may subsequently be re-fitted with a better estimate of the requisite parameters.

The parametrisation forms a two dimensional surface as a function of pitch and α_p. This is achieved by measuring the resolution in one degree bins of α_p as a function of the pitch. As shown in the left plot of Fig. 3.4 the resolution follows a linear dependence with pitch so for each α_p bin the resolution at the centre is parametrised by a first order polynomial

$$error(\alpha_p, pitch) = p_0(\alpha_p) + p_1(\alpha_p) \times pitch \tag{3.15}$$

The resolution between the bin centres is then interpolated by implementing a spline between the values of p_0 and p_1. Thus for a cluster of given pitch and α_p, p_0 and p_1 are determined and the resolution returned for the track fit.

The error parametrisation used thus far is taken from studies on simulated data performed before the start of data taking in 2010. As stated previously the aim of this work is to update that parametrisation with resolution measurements made using real collision data as well as to investigate how the resolution changed during Run 1.

3.3 Measuring the cluster resolution

3.3.1 Method

The resolution of the detector can be ascertained by considering the distribution of the residuals. For data the residual is the distance between the fitted track position at the sensor in question and the measured track hit. The inclusion of the sensor under consideration in the track fit biases this residual, so it is corrected by a factor $\sqrt{\frac{V}{R}}$ [108, 106], where V and R are the variance of the measurement and residual respectively. An example of such a distribution is shown on the left of Fig. 3.5. Fitting it with a single Gaussian function and extracting its width indicates the average error of the measurements.

These errors are measured in bins of α_p, which is further binned by pitch. The resolution as a function of pitch is then plotted for each bin of α_p and fitted to extract the parameters p_0 and p_1. Thus a new error parametrisation has been found which can then be inserted back into the track reconstruction. As the error parametrisation affects both the track fit and the un-biasing of the residuals the process is repeated until successive iterations return the same measured residual distributions. The veracity of the resulting parametrisation can be checked by considering the *pull* distributions. The pull is defined as

$$pull = \frac{biased\,residual}{residual\,error}.$$

If the errors are properly estimated the distribution of the pulls should be Gaussian in shape with width one. When measuring the pull distribution width the standard deviation is used such that any biases, which would most likely be indicative of misalignment and not of immediate interest for this study, do not affect the result. A Gaussian fit to the distribution should give an identical result, although it would not deal well with a shape that is not entirely as expected.

The reconstruction that was used for the study was exactly the same as that used for analyses, with the exception of the modified error parametrisation. Certain selections were made on the tracks to ensure that they were good quality. They were required to have at least 10 hits in the VELO, a momentum of more than 5 GeV and χ^2 of between 0.1 and 5. Only tracks that had traversed the VELO and tracking systems and left energy in the calorimeters were selected (called *long* tracks). These tracks have momentum measurements so the uncertainty from multiple scattering can be accounted for, although the cut on momentum excludes those tracks that suffer most from multiple scattering. Finally the first and last VELO hits for each track were not included in the measurements; these are extrapolated only from one side so the resolution returned by the track

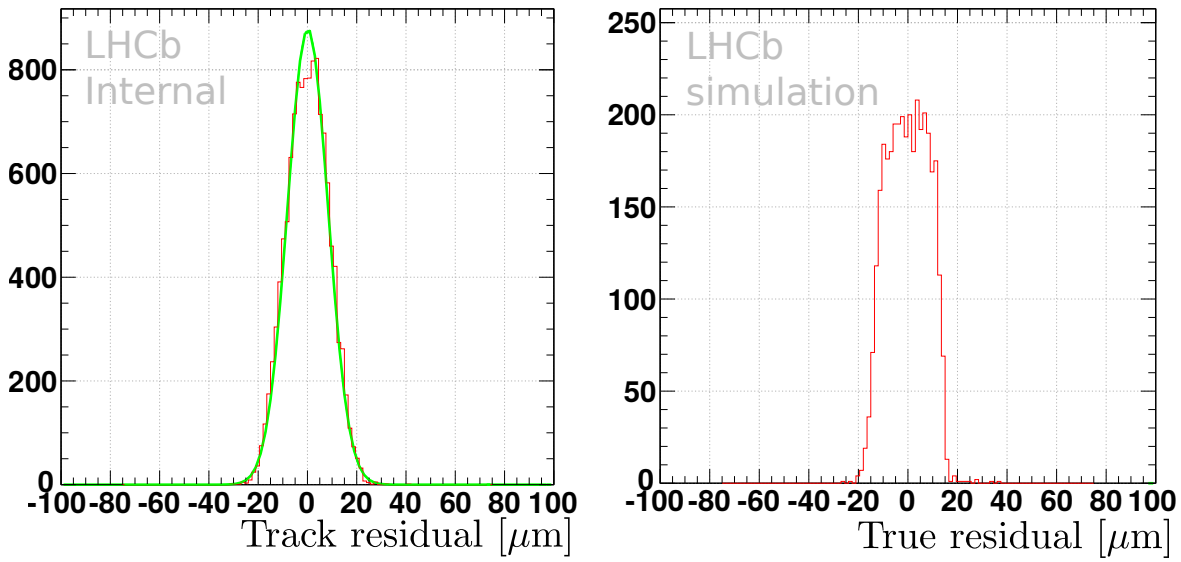


Figure 3.5: Example residual distributions for collision data (left) and simulation (right). The green curve on the left plot represents a Gaussian fit. The plots show different bins of pitch and projected angle and should not be compared.

fit may be worse than the true hit resolution. The r and ϕ clusters were considered separately, with r being the subject of the simulation tests. They cover different regions of α_p due to the strip geometry, ϕ having mostly very small α_p and r slightly larger but should have the same resolution. Nevertheless they could be affected differently by any misalignments in the detector.

3.3.2 Validation of the method

To test whether the proposed method reproduces the resolution of the detector a study was carried out using simulated data. The true resolution of the simulated detector can be ascertained as the exact paths of the particles are known. To do so the distribution of the difference between the true hit (using the simulation generator information) position and that of the measurement is considered, henceforth referred to as the *true residual* distribution. This is entirely independent of any information of the Kalman filter such as the description of multiple scattering or the error parametrisation.

Initially the true resolution was measured using the simulation information. An example of the residual distributions of the true residuals is shown on the right of Fig. 3.5; as this is not close to a Gaussian shape the RMS was taken for the width.

Using these widths the true resolution of the simulated detector was measured and parametrised. It was noted that the resolution of one-strip and two-strip clusters is rather different, a result that is to be expected; henceforth

they were considered separately. Three-strip clusters were not accounted for in either the study with simulated data, or that on collision data. They make up a small fraction of clusters (see Fig. 3.15, the simulated data sample size was rather limited so a large number of three-strip clusters was not present) and have a poor resolution.

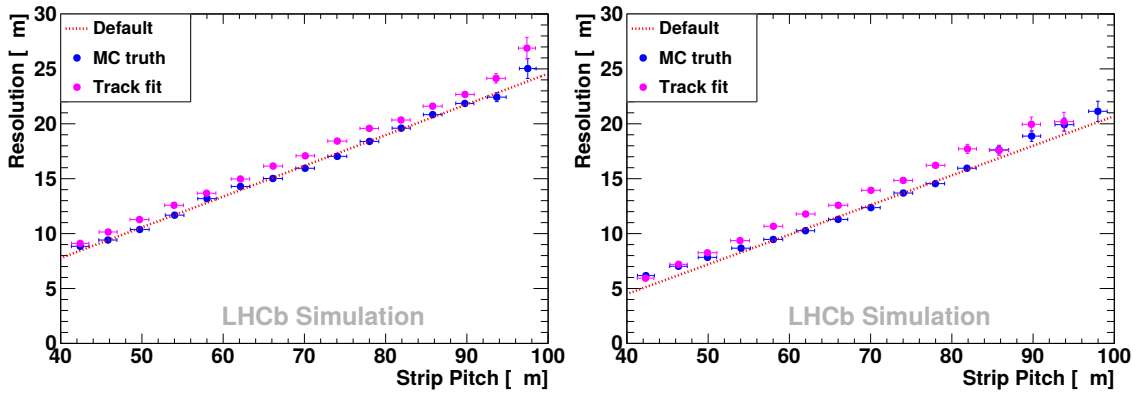
The parametrisation of the true resolution was put into the reconstruction which was then run again on the same data. The residuals of the track fit were then compared to the true residuals to check that the correct errors were being returned. The results are shown in Fig. 3.6. For the two-strip r clusters the resolution measured using the track residuals is the same as that determined using the truth information. For 1-strip clusters the track errors are slightly larger. This is due to the distributions of the truth residuals being rather square for 1-strip clusters, whereas the track fit produces a more Gaussian like distribution; the truth measurement has probably underestimated the true resolution. Taking the resulting measurement and iterating once more returns the same resolution.

The hypothesis is born out by the plots in Fig. 3.7 comparing the pulls of the reconstruction using the error parametrisation gleaned from the truth information (left) and the subsequent iteration (right). For both the width of the pulls for the two-strip clusters are close to one. For one-strip clusters the truth parametrisation has produced pulls a little larger than one, implying that the error has been underestimated. The subsequent iteration brings them closer to one. Henceforth this measured resolution (the pink line in Fig. 3.6) is considered as the true resolution of the simulated sample.

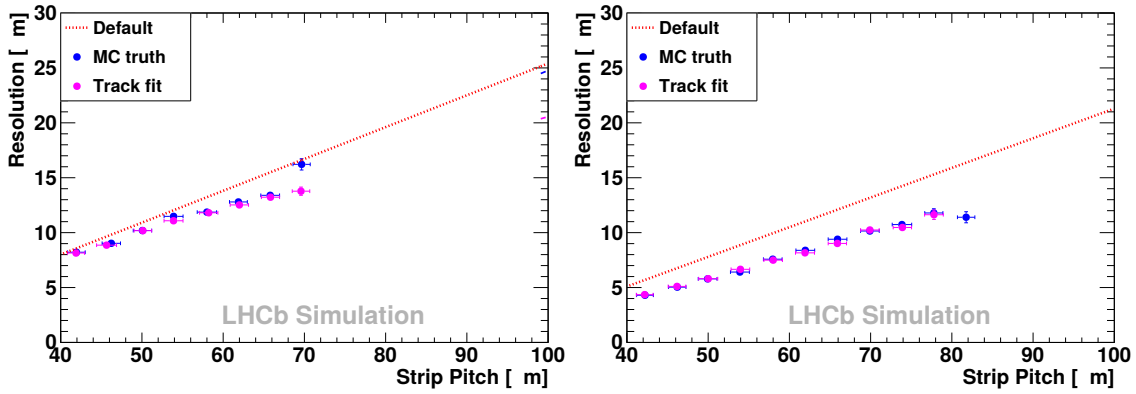
Next the method for updating the parametrisation was trialled. Starting from the default error parametrisation the reconstruction was run and the resolution measured from the track fit results, parametrised and put back into the reconstruction which was run again. The resolution using this parametrisation was measured and compared to that used to produce the parametrisation; the two were found to be identical suggesting that the method had converged quickly. As the true resolution and the default parametrisation are not too dissimilar such a quick convergence may be expected.

Thus the study with simulated data has shown that by measuring the cluster resolution and iterating the parametrisation of the hit error a measurement of the true resolution can be obtained. The resolution as a function of projected angle of the simulated data sample is shown in Fig. 3.8. The best resolution is about $2\,\mu$m for two strip clusters with projected angle of about 8° in simulation.

As previously stated, the residual distributions are not entirely Gaussian like in their shape. Three strategies were used to evaluate the spread of the distribution to extract a resolution measurement: the standard deviation, a fit of a

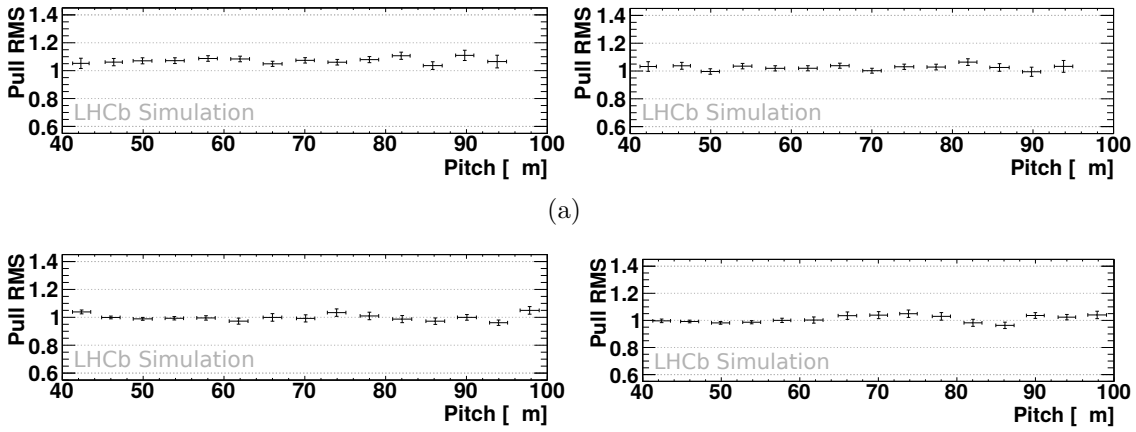


(a)



(b)

Figure 3.6: Comparison of the true resolution with that measured by the track fit for one-strip (a) and two-strip (b) r clusters for two angle bins. The MC truth measurement is shown in blue, and the resulting resolution measured from the track fit is in pink. The dashed red line shows the default error parametrisation used in the reconstruction.



(a)

(b)

Figure 3.7: Examples of the RMS of the pull distributions for the simulated data. The reconstruction using the error parametrisation of the MC truth is on the left, the subsequent iteration on the right. The top plots are for one-strip clusters with projected angles between six and seven degrees. The lower are two-strip clusters with angles between three and four degrees.

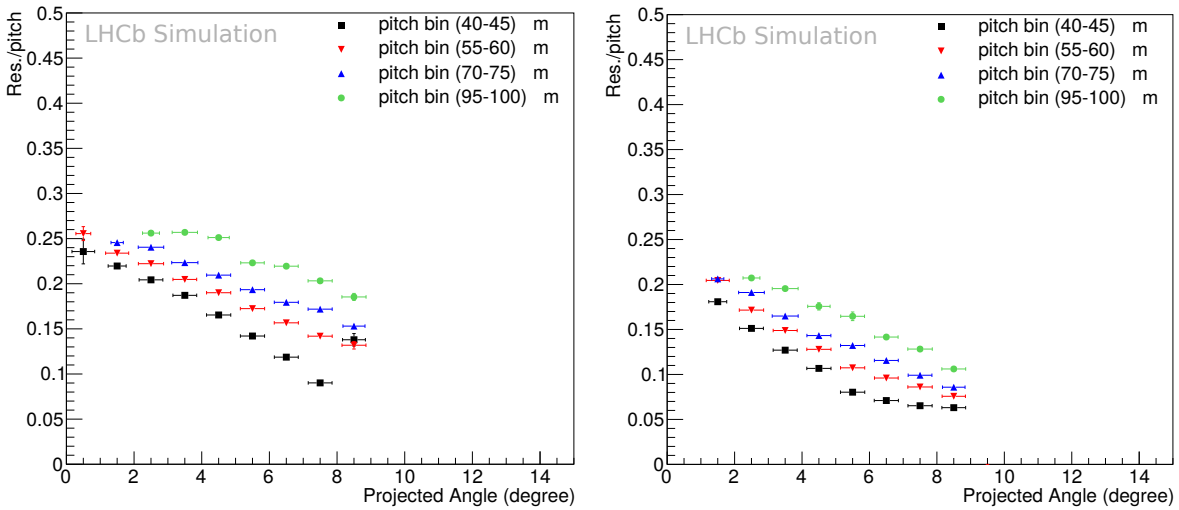


Figure 3.8: The measured resolution divided by pitch for the simulated data sample as a function of the projected angle. A range of pitch bins are plotted for one (left) and two (right) strip clusters on r type sensors with the resolution evaluated using the truth information.

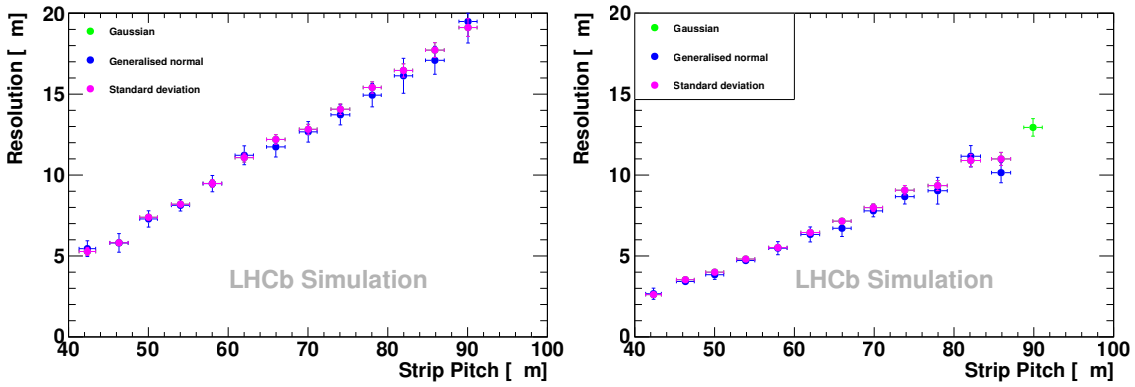


Figure 3.9: A comparison of three methods to fit the residual distributions and extract their widths. Shown are the standard deviations of Gaussian (green) and generalised normal distributions (blue) fitted to the distributions as well as the calculated standard deviation (pink). The three methods are shown to be in good agreement.

Gaussian function and a fit of a generalised normal function. The standard deviation has the benefit over the RMS that it is not be affected by a bias (neither is the width of a fitted function). Fitting a function to the distributions should negate the effects of outliers. Fig. 3.9 shows the measured residuals for two angle bins of the simulated data. As can be seen the three methods yield essentially identical results if the bins contain a large number of entries.

3.3.3 2010 parametrisation

The hit resolution has been established using a sample of data recorded in 2010, early in the Run 1 data taking period, using the method described above. A selection of the results is shown in Figs. 3.10 and 3.11. The measured resolution broadly follows expectations. In particular as a function of projected angle it is shown to reach a minimum at around eight degrees. The best resolution is measured to be $\sim 4\,\mu\text{m}$ (10 % of the pitch), slightly worse than that of the simulated data.

A selection of the pulls for the reconstruction run using the error parametrisation shown in Fig. 3.10 is shown in Fig. 3.12. Generally the RMS of the pull distributions are around one; there is some deviation at the lowest and largest pitches suggesting that the linear parametrisation of resolution with strip pitch, although a good approximation, is not exact.

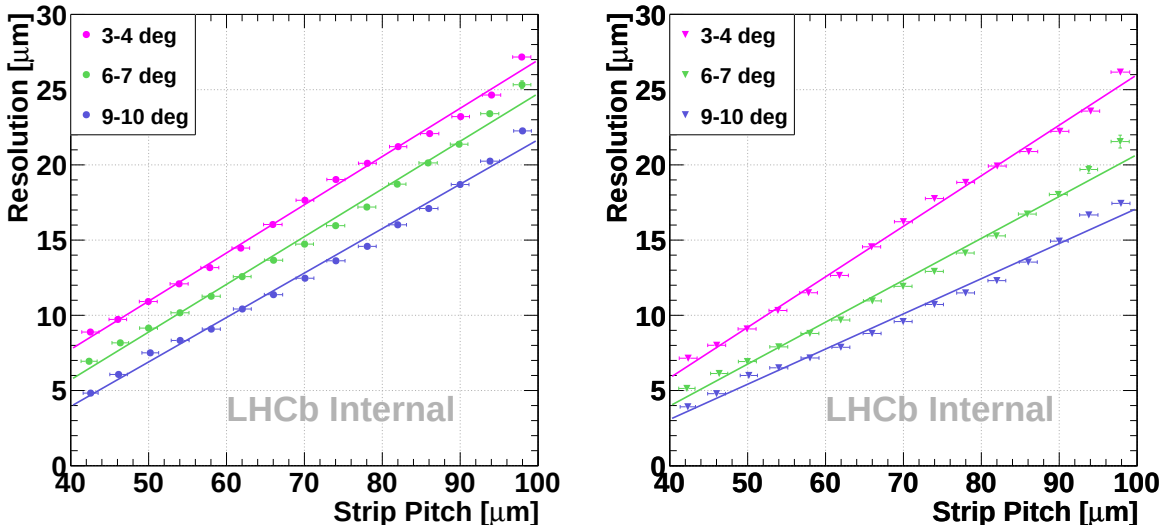


Figure 3.10: The measured resolution for the 2010 data sample as a function of the strip pitch. A range of projected angle bins are plotted for one (left) and two (right) strip clusters on r type sensors.

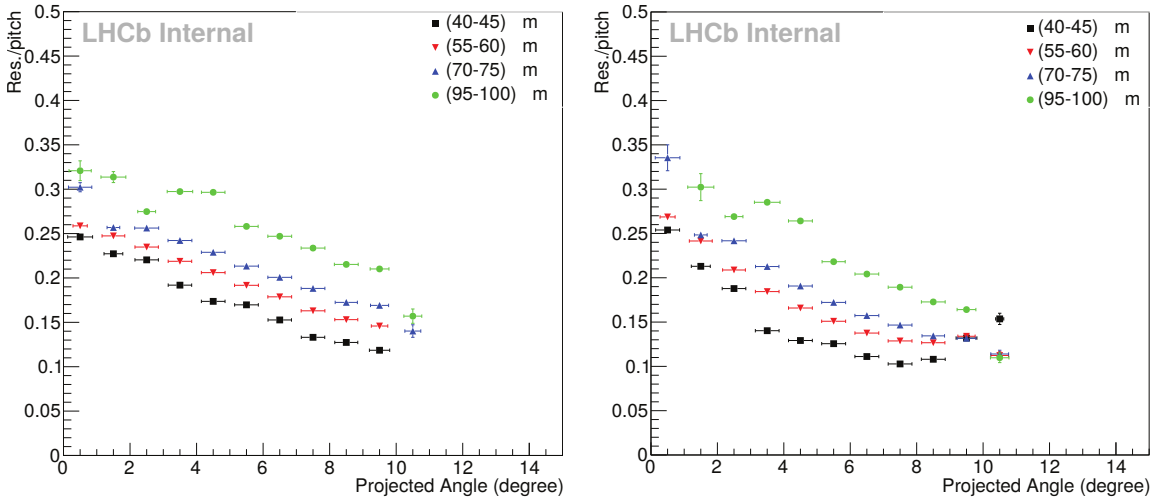


Figure 3.11: The measured resolution divided by pitch for the 2010 data sample as a function of the projected angle. A range of pitch bins are plotted for one (left) and two (right) strip clusters on r type sensors.

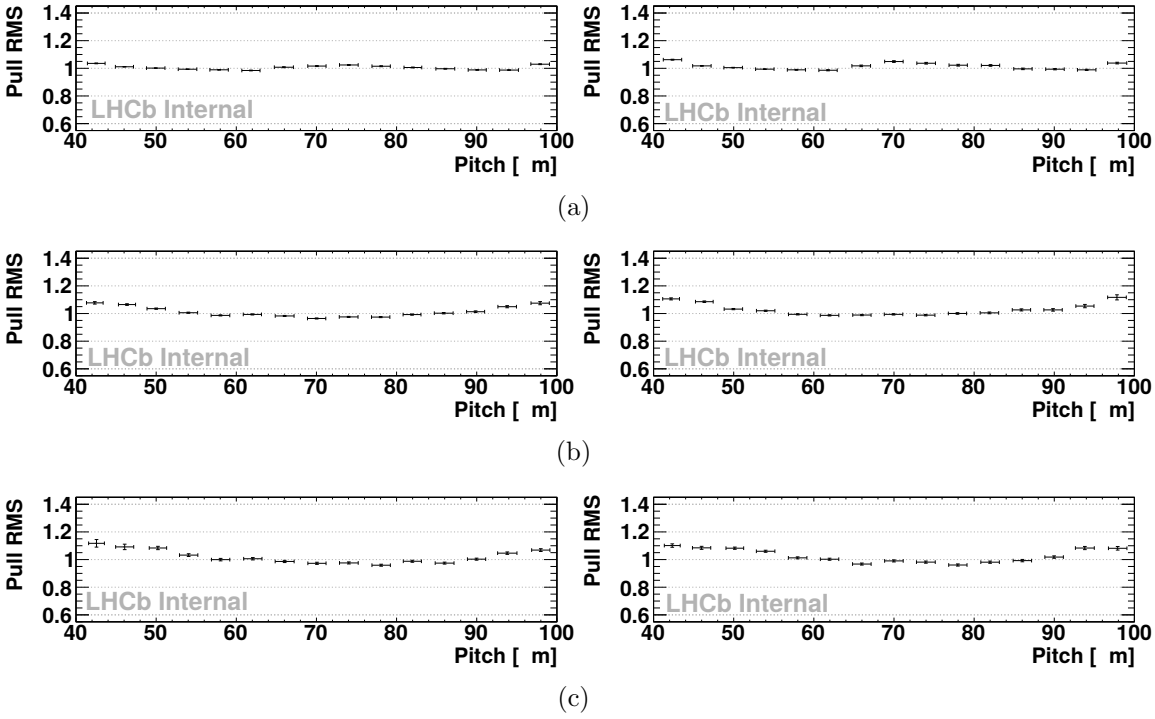


Figure 3.12: The RMS of the pull distributions for clusters reconstructed using the new error parametrisation in Fig. 3.10. Shown are one (left) and two (right) strip clusters for the angle bins 3–4 (a), 6–7 (b) and 9–10 (c) degrees.

3.4 Resolution throughout Run 1

The measurement of the resolution of both simulated data and the sample from the start of Run 1 confirmed that the proposed method of updating the error parametrisation would lead to correct results. Subsequently the evolution of the VELO resolution over the data taking period from 2010 to 2012 was examined. It was expected that the resolution would not change to a large degree. However initial investigation suggested otherwise so the resolution was measured with data taken at four different periods of Run 1: 2010, April 2011, March 2012 and November 2012. All data was collected with an applied bias voltage of 150 V.

3.4.1 Effects of radiation damage

Radiation damage is known to have caused some significant changes in the sensors [77, 92]. For example consider Fig. 3.13 showing the effective depletion voltage (EDV) with fluence. Initially the EDV decreases with radiation, until the n bulk undergoes type-inversion and the effective depletion voltage rises. Furthermore charge loss to the readout metal layer has been observed (the lines reading out the signals from the strips), leading to a reduced signal charge being collected by the strips.

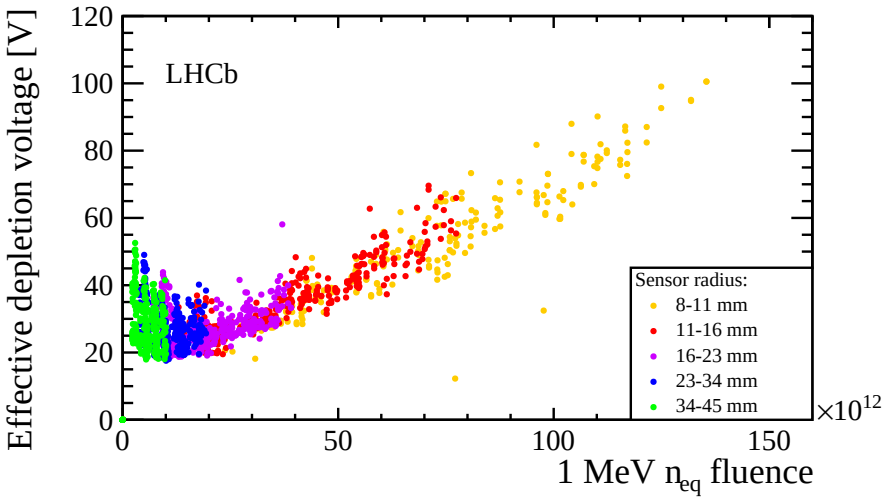


Figure 3.13: The change in the effective depletion voltage of the VELO sensors with fluence. From [77].

The effect of radiation damage on the charge collected per strip over Run 1 can be seen in Fig. 3.14 which shows the variation of the charge collected per strip from the start to end of the data taking period. By November 2012 less charge is collected.

Additionally one can examine the fraction of VELO hits that are one, two and three strip clusters, presented in Fig. 3.15. By the end of 2012 the fraction of one strip clusters has considerably decreased at small projected angles. Correspondingly the percentage of two and three strip clusters has increased, strikingly for three strip clusters. This is initially surprising because for small projected angles the tracks are almost perpendicular to the face of the sensor so one would expect that charge sharing between strips would be minimal. The evolution of the percentage of one two and three strip clusters with projected angle is also much reduced in the 2012 sample; the distributions are much flatter. Whilst a change in these distributions may not necessarily indicate that the hit resolutions for the individual cluster types (one, two or three strips) have changed they do suggest that the average hit resolution has. From the previous section it was seen that one and three strip clusters have poorer resolution than two strip clusters.

Two physical processes were mooted to be behind the changes. The first would be that the electric field inside the detector bulk has changed due to radiation damage. The EDV first decreases with fluence before type inversion occurs followed by increasing EDV (see Fig. 3.13). Because of this the charge deposited by a particle track spreads out further before reaching the strips, increasing the likelihood of creating multi-strip clusters.

Alternatively the capacitive coupling between adjacent strips may have

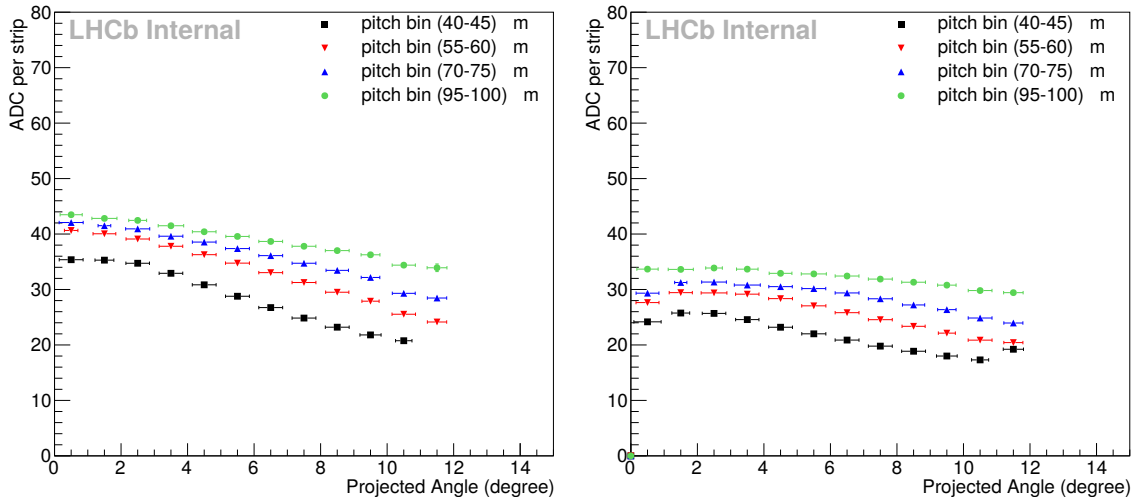


Figure 3.14: The average charge collected per strip (ADC counts) for all clusters on r sensors for the 2010 (left) and November 2012 (right) samples. The results are plotted against projected angle in bins of pitch.

changed. An increase in coupling again suggests larger clusters as it is more likely that strips adjacent to a seed will have collected enough charge to be above the inclusion threshold. Both effects have been crudely investigated using simulated data.

The simulated data is the same as that used in section 3.3.2; simulated beam-gas collisions with a minimum bias trigger. As the simulation files have been retained the subsequent detector simulation can be re-run with different configurations. The detector digitisation software gives a rather simplified description of LHCb and the results must be interpreted with that in mind.

3.4.2 Bias voltage

The bias voltage in the simulation is more a description of how the electric field in the sensor affects the distribution of charge rather than a physical voltage between the sensor strips and the back plane. For a given hit position the subsequent charge distribution is modelled by a normal distribution, the width of which is inversely proportional to the bias voltage parameter as indicated in Fig. 3.16.

The standard simulation has the bias voltage at 150 V. The effect of halving it is shown in Fig. 3.17. As expected the fraction of multi-strip clusters increases for the lower bias voltage for the smallest pitch bin at low projected angles. For larger pitches at lower angles there seems to be little difference indicating that the now wider charge distribution is not large enough to cover multiple strips with greater separations. The evolution of the distributions with projected angle is rather flatter for the lower bias voltage suggesting that the width of the Gaussian

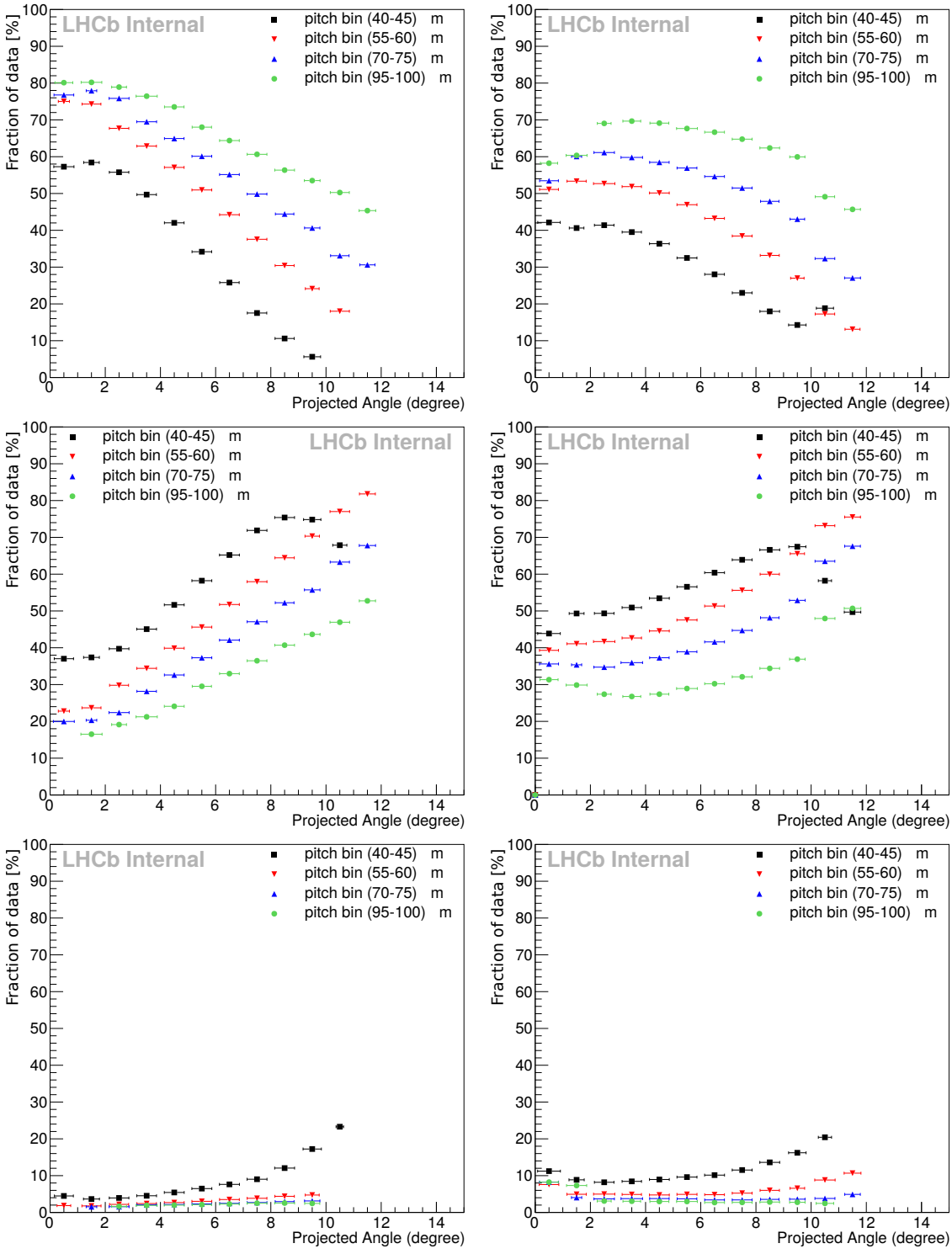


Figure 3.15: The percentage of the data that is comprised of one (top), two (middle) and three (bottom) strip clusters on r sensors for the 2010 (left) and November 2012 (right) samples.

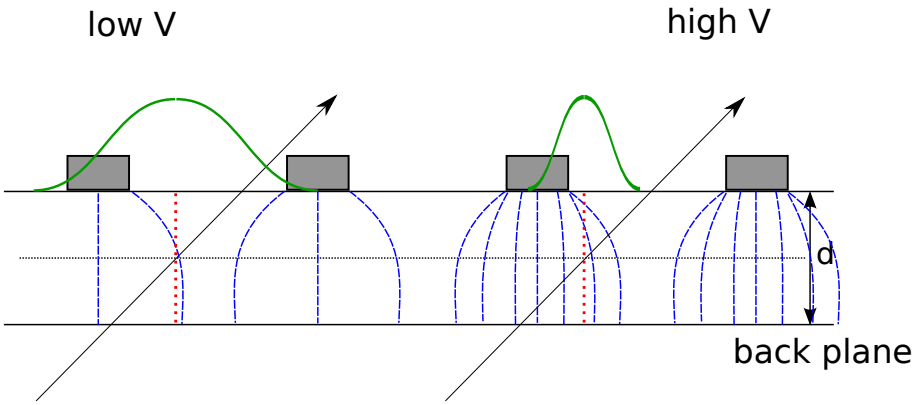


Figure 3.16: The simulation of the VELO with low (left) and high (right) bias voltages. The electric field is indicated in blue and the green Gaussian shapes represent the simulated distribution of the the charge from a particle hit.

function that describes the simulated field and subsequent spreading out of the charge is having an effect of similar magnitude to the distribution of the charge due to the angle of the particle path through the detector.

3.4.3 Capacitive coupling

The capacitive coupling between the strips is simply modelled in the simulation. In the default configuration a coupling of 1% is simulated; 2% of that strip's charge is taken off and evenly distributed to the two adjacent strips. Radiation damage my cause the surface properties of the strips to change, affecting the capacitance between adjacent strips. Fig. 3.18 shows the simulation reconstructed with 15% coupling between adjacent strips applied.

As for the change in bias voltage, the capacitive coupling qualitatively reproduces some of the changes seen in the data between 2010 and 2012. At low angles and small pitches there is an increase in the percentage of multi-strip clusters. Again the distributions as a function of projected angle are flatter with the increase in capacitive coupling with the change being most noticeable for the smallest pitch bin.

As mentioned previously the detector simulation is somewhat simplified, in particular with regards the bias voltage and how charge moves through the detector bulk. Additionally the changes are applied to all sensors evenly whereas in the real VELO the radiation has a radial dependence and the damage would not be evenly distributed.

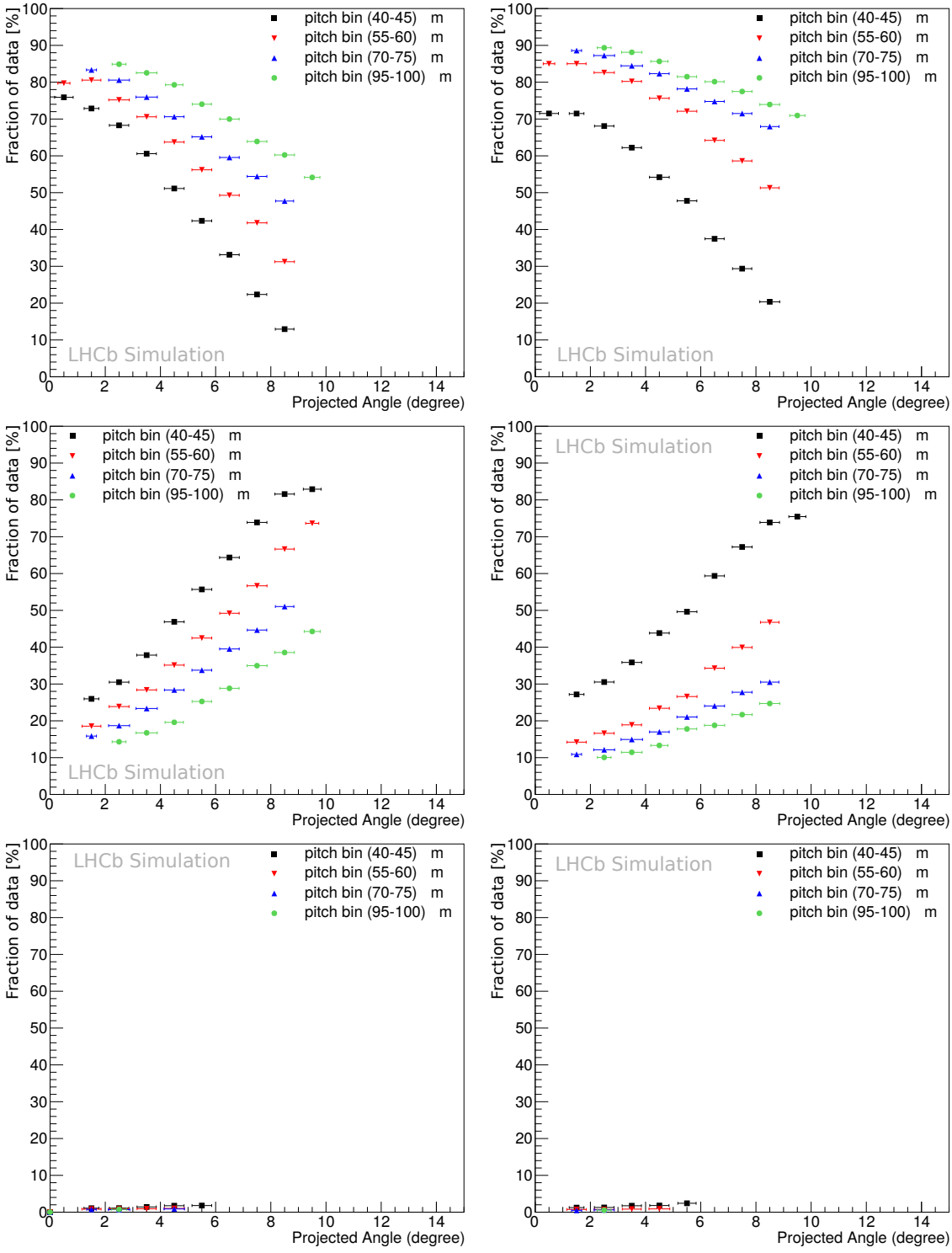


Figure 3.17: The effect of the bias in the VELO simulation. On the left is shown the standard 150 V bias, 75 V on the right for one (top), two (middle) and three (bottom) strip clusters.

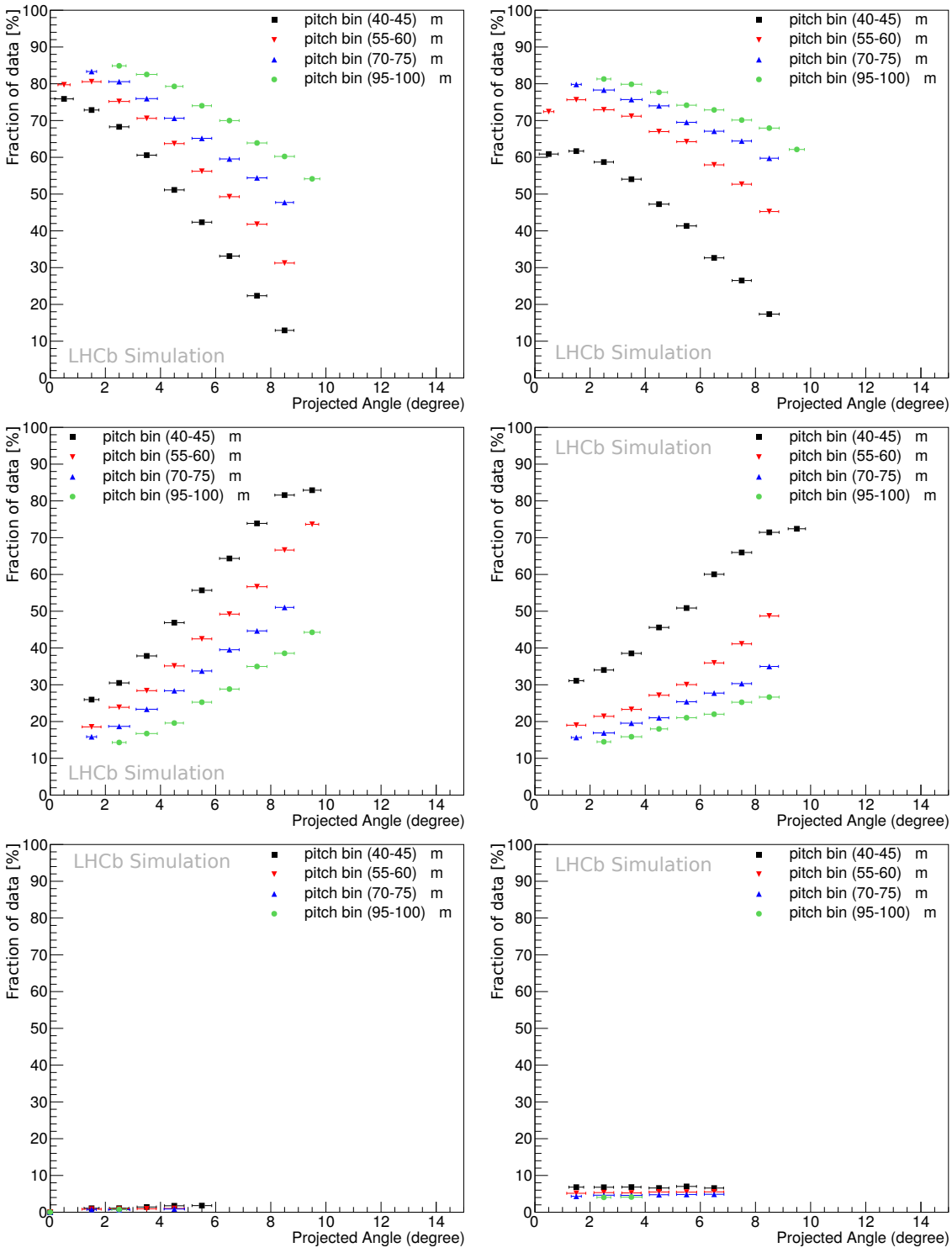


Figure 3.18: The effect of the coupling between adjacent sensor strips in the VELO simulation. On the left is shown the standard 1% coupling, 15% on the right for one (top), two (middle) and three (bottom) strip clusters.

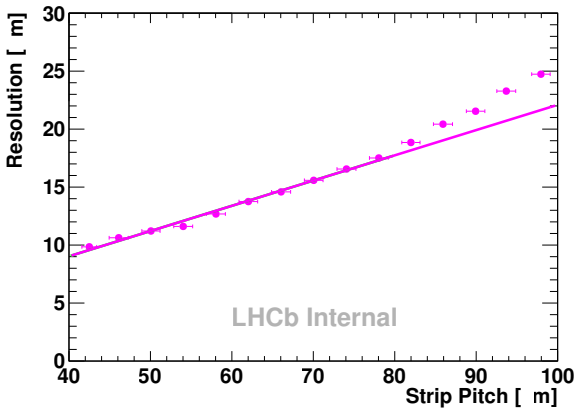
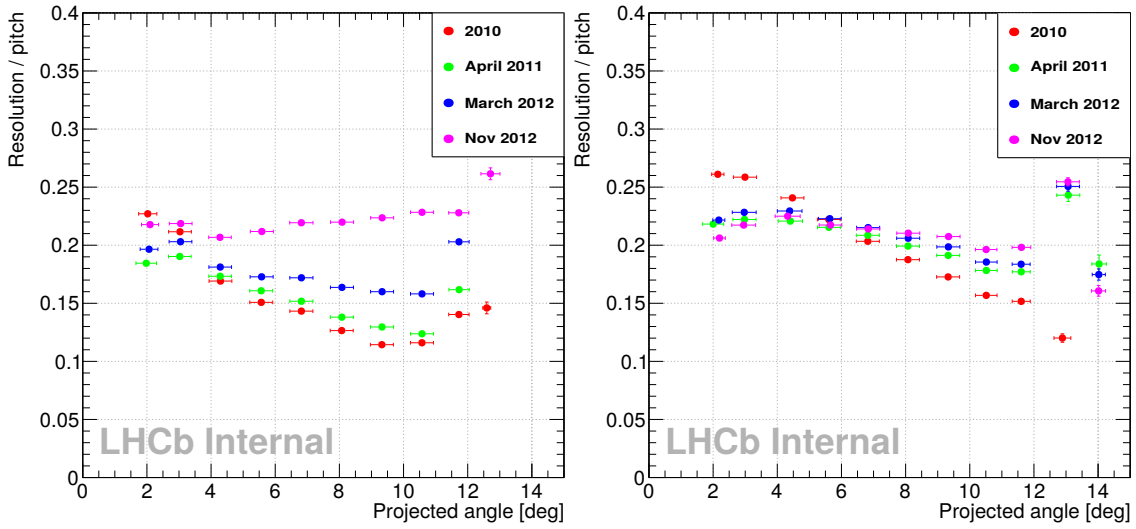


Figure 3.19: An example of the linear parametrisation of the resolution with strip pitch for the November 2012 data sample. Shown is the resolution of 1-strip r clusters for tracks with projected angles in the range 8–9 deg. A linear fit is overlaid, where only points in the range 40–80 µm were considered in the fit. The deviation from the linear parametrisation at large strip pitches can be clearly seen.

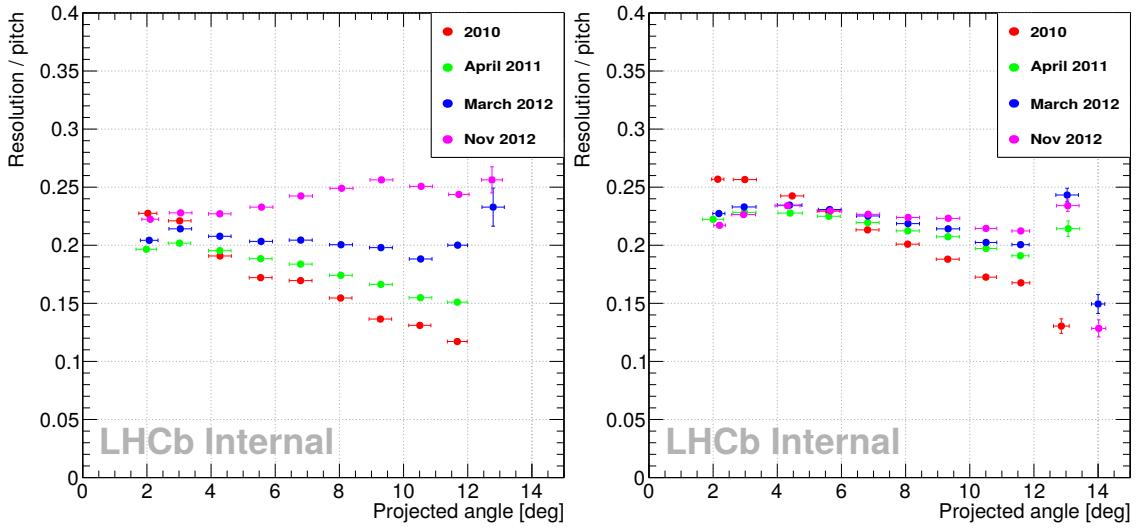
3.4.4 Resolution

The resolution of the VELO was measured for the four samples of data from Run 1 (2010, April 2011, March 2012 and November 2012) using the iterative method outlined earlier. From the April 2011 sample onwards it was noticed that the parametrisation of resolution with pitch was deviating from a linear description for larger pitches in some projected angle bins (for example see Fig. 3.19), probably due to the large variation in damage with radius from the beam. To avoid these larger pitches altering the results the parametrisation was fitted only between 40 and 80 µm, the larger pitches being ignored for the purposes of this study. As before the r and ϕ sensors were considered simultaneously though for brevity only the r cluster results are presented here. The results plotted against projected angle are shown in Fig. 3.20; the changing resolution with time is apparent. Confirmation of the November 2012 measurement can be seen in the example pull widths in Fig. 3.21 produced using the reconstruction with the error parametrisation taken from the results shown in Fig. 3.20. The deviation from a linear parametrisation is apparent at the lowest and largest pitches where the under-estimation of the true error leads to widths greater than one.

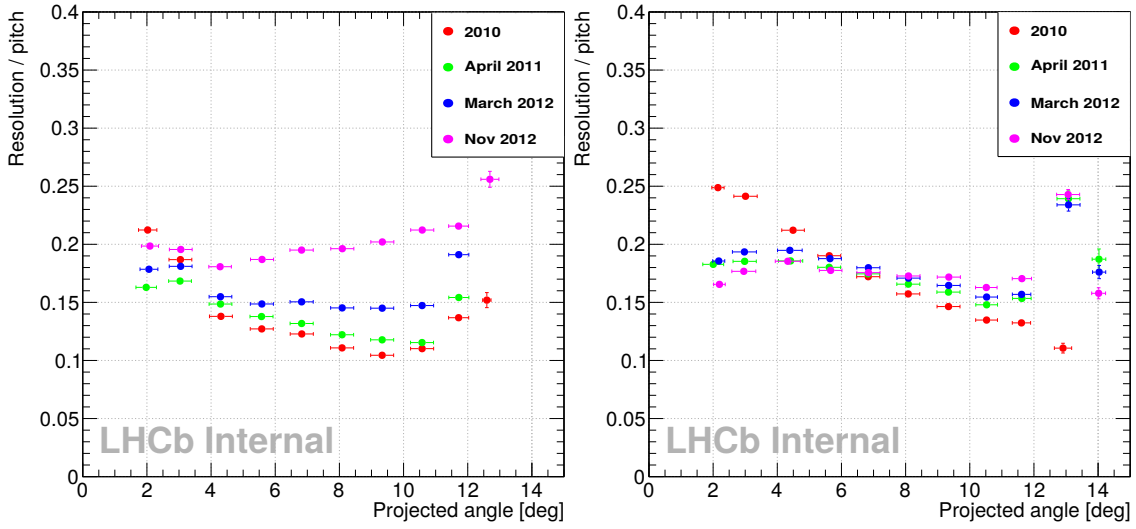
There appear to be two changes occurring. At the smaller angles the resolution appears to improve between 2010 and April 2011. Subsequently in the lower pitch bin the resolution degrades again towards the end of the data taking period. In the larger bin the resolution progressively degrades throughout Run 1. At larger angles the best resolution worsens over time, in particular during 2012 for the smaller pitch bin. This suggests that initially the increased charge sharing in the bulk that one expects due to the radiation damage is leading to an increase



(a)



(b)



(c)

Figure 3.20: The resolution divided by pitch plotted against projected angle for the four data samples for the r sensors. Shown are all clusters sizes (a), one (b) and two-strip (c) clusters in two bins of pitch: 40–45 µm (left) and 70–75 µm (right).

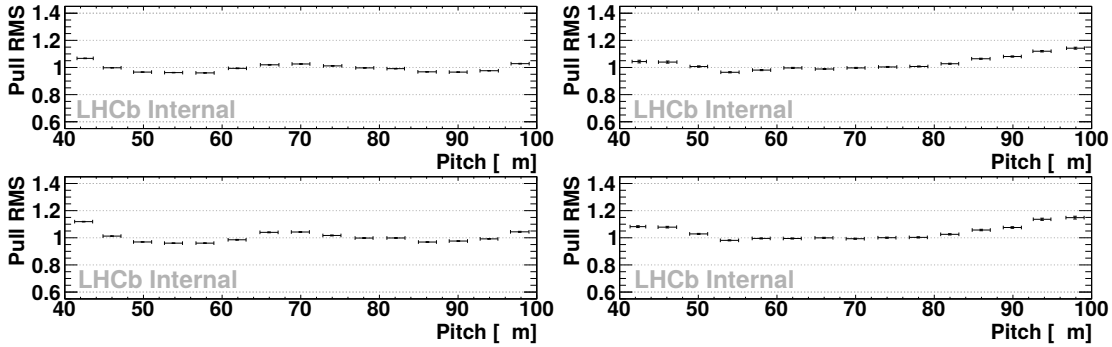


Figure 3.21: Examples of the widths of the pull distributions for the error parametrisation that was arrived at for the November 2012 data sample. Shown are the projected angle 3–4 (left) and 8–9 (right) degrees for one (top) and two (bottom) strip r clusters. The deviation of the true error from the linear parametrisation can be seen in the lowest and highest pitch bins where the error has been underestimated.

of the worst resolutions. On the other some other effect, potentially an increase in the coupling between the strips, is decreasing the resolution. Eventually after enough fluence the damage causing this degradation of the resolution dominates leading to the trends seen.

The radiation damage is not uniform over the whole of the VELO; the sensors closest to the interaction point receive the highest doses and the fluence falls off roughly as inverse square law with radius [77] as shown in Fig. 3.22. The results shown in Fig. 3.20 are integrated over all sensors so the variation in damage with pitch and projected angle is causing the resolution plotted against pitch to move away from being linear.

3.5 Effect on the reconstruction

As noted in section 3.1.1 the aim of this study was to update the error parametrisation used in the Kalman filter to reflect the true resolution of the VELO. To that end the resolution measured at the end of 2012, as described in the previous section, has been parametrised. The effect of updating the parametrisation on the physics performance has been assessed with data taken in November 2012. The reconstruction was run with the standard parametrisation and the new one and the results compared. Additionally a comparison has been made between the data taken in 2010 and in November 2012, both reconstructed with updated error parametrisations.

One would expect that the effects of the error parametrisation would be most apparent in the χ^2 quantities, in particular the track χ^2 and the impact parameter χ^2 as they both directly depend on the estimated errors of the track fit. Many

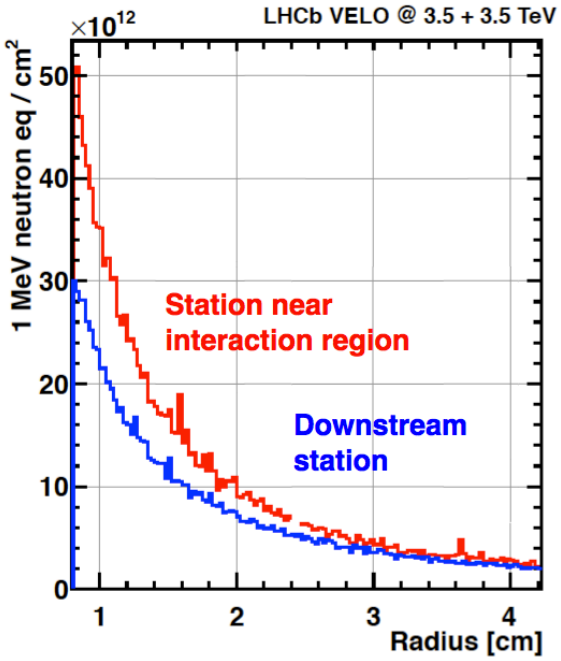


Figure 3.22: The estimated received fluence against radial distance from the beam axis after $1\,\text{fb}^{-1}$ of integrated luminosity for collisions with 7 TeV centre of mass energy. Two sensors are shown, one close to the interaction region and one downstream. The plotted curves are from a simulation prediction. From [92].

analyses cut directly on the track χ^2 to ensure good quality tracks as well as the track impact parameter χ^2 to look for a significantly displaced vertex. Furthermore the track fit employs an outlier rejection, based on the distance between a measurement and the track projection. The cut is imposed on the number of standard deviations between these quantities. Therefore one may expect that the number of outliers included or rejected would change with the updated parametrisation.

As the errors used in the track fit are essentially only a weighting between the predictions and measurements they should have minimal impact on the absolute measured quantities (those that are not χ^2) if all the errors scale by the same amount. The effects of the error parametrisation would be most readily observed for high momentum tracks as the uncertainty on those with low momentum will be dominated by the scattering effect of the detector material.

3.5.1 Tracks

For each track the hit efficiency can be estimated by the ratio of the number of hits included in the track and the expected number of hits that the track should have from the number of sensors traversed. A comparison of the efficiency is shown in Fig. 3.23. There is no significant change in the performance of attaching hits to

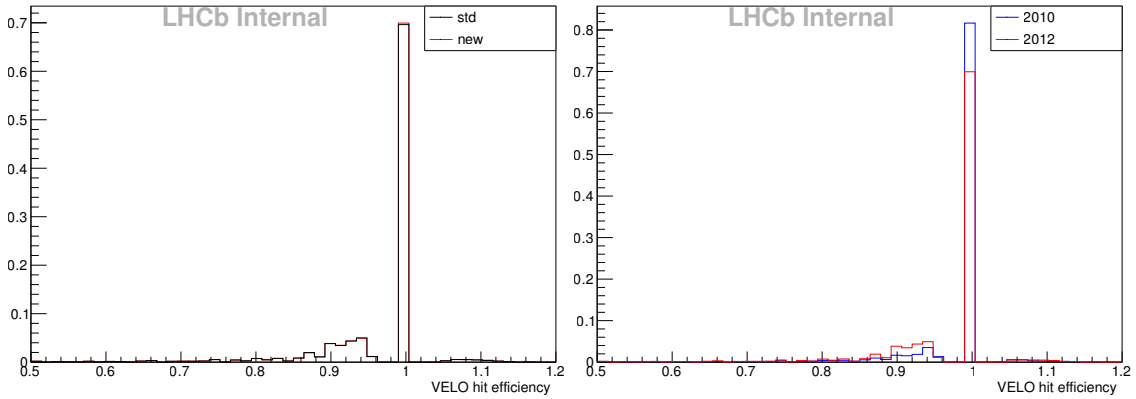


Figure 3.23: The pseudo-hit efficiency of the VELO. On the left is a comparison between the new error parametrisation (red) and the standard parametrisation (black) for the data taken in November 2012. On the right is a comparison between data taken in 2010 (blue) and November 2012 (red), both reconstructed with their respective updated error parametrisations. This efficiency compares the number of VELO hits in the track with the expected number of hits for the particle path.

tracks when using different error parametrisations. However between the start and end of Run 1 the efficiency has decreased, implying that hits are being missed. This decrease is expected; the cluster finding efficiency has been found to decrease with delivered luminosity to the detector [77].

The distribution of the χ^2 probability of the VELO segment of tracks is shown in Fig. 3.24. Ideally the distribution should be flat, resulting in a mean probability of 0.5. The difference between the two parametrisations for the November 2012 data sample can be seen in the left plot, particularly at higher probabilities, with the new parametrisation producing a significantly improved, flatter distribution. There is also a clear change between 2010 and November 2012 (the right plot) as a sign of the changing resolution, although using the updated error parametrisation has lessened the change compared to persisting with the standard parametrisation.

The change in the total track χ^2, which includes contributions from the other tracking detectors (TT and the three downstream tracking stations), is shown in Fig. 3.26. As can be seen the difference between the two parametrisations and the difference between the start and end of Run 1 is much reduced when considering the full tracks compared to just the VELO segments as expected. The implication is that the total errors for long tracks are dominated by other sources, most likely the resolutions of the other detectors.

From 2015 the primary vertex fit will be performed with just the VELO segment of tracks. Therefore to investigate the primary vertex and impact parameter quantities the VELO parts of fitted long tracks have been taken,

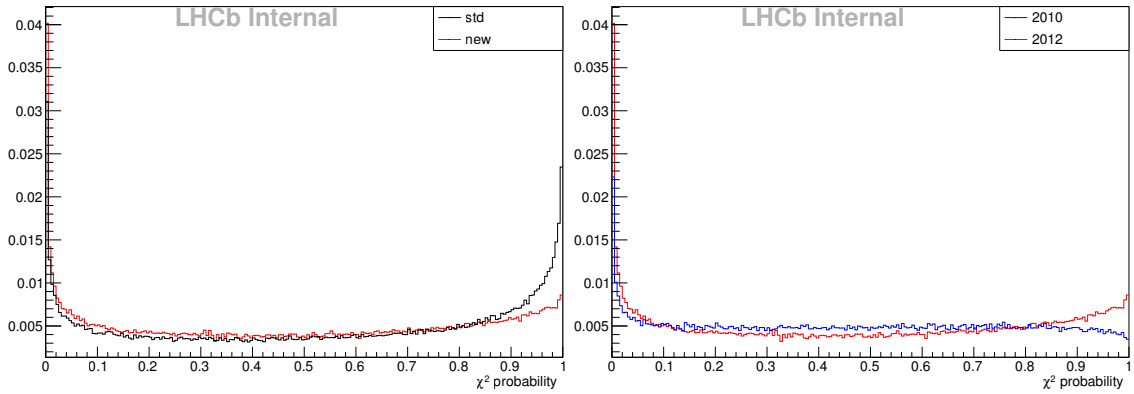


Figure 3.24: The χ^2 probability distribution of the VELO segment of tracks. On the left is a comparison between data taken in November 2012 reconstructed with the standard (black) and updated (red) error parametrisations. On the right is the comparison between data taken in 2010 (blue) and November 2012 (red), both reconstructed with their respective updated error parametrisations.

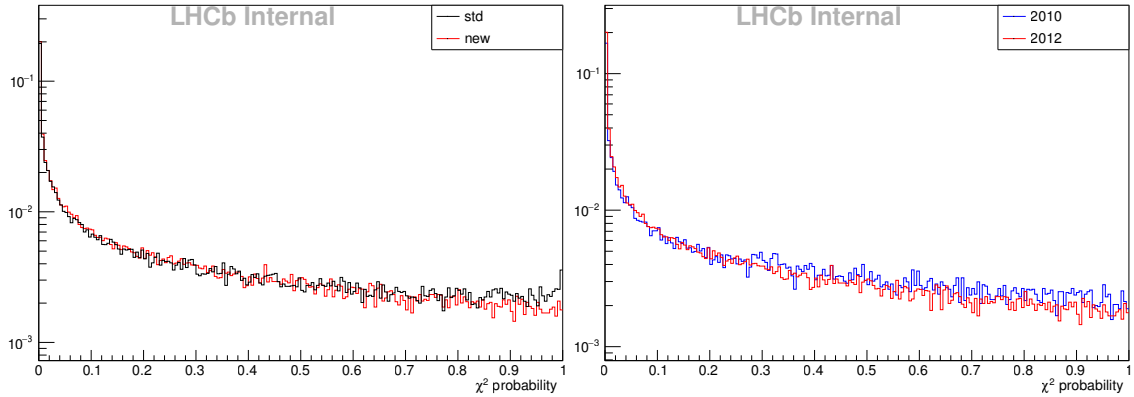


Figure 3.25: The χ^2 probability distribution of long tracks. On the left is the November 2012 data sample reconstructed with the standard (black) and updated (red) error parametrisations and on the right is a comparison between the 2010 data (blue) and the November 2012 sample (red). These distributions included contributions from all tracking stations.

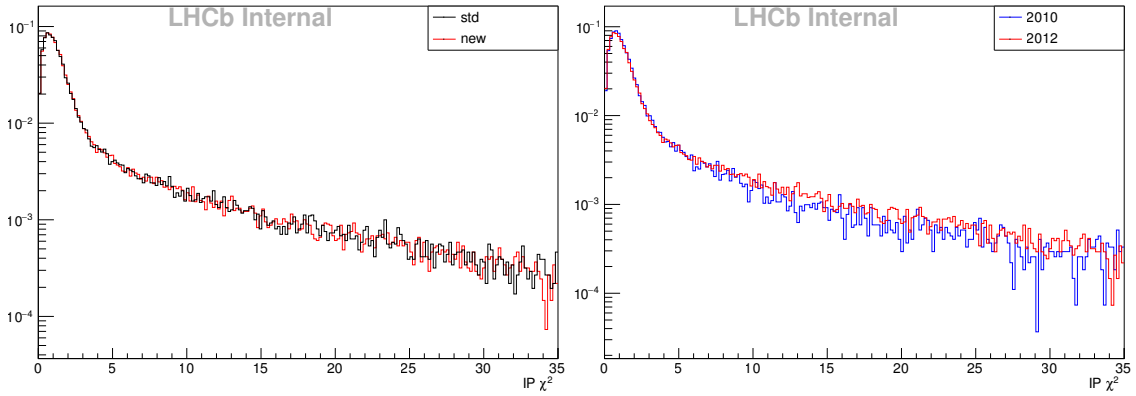


Figure 3.26: The impact parameter χ^2 distributions of tracks. On the left is the November 2012 data sample reconstructed with the standard (black) and updated (red) error parametrisations and on the right is a comparison between the 2010 data (blue) and the November 2012 sample (red).

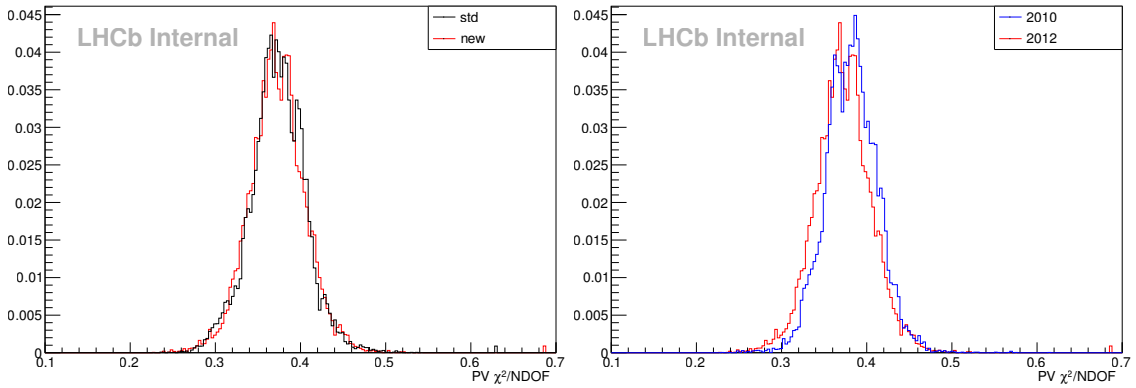


Figure 3.27: The primary vertex χ^2 per degree of freedom.On the left is the November 2012 data sample reconstructed with the standard (black) and updated (red) error parametrisations and on the right is a comparison between the 2010 data (blue) and the November 2012 sample (red).

refitted with just the VELO hits but using the full momentum information and then primary vertex fitted with these VELO tracks. When calculating the impact parameter quantities the primary vertex has been re-fitted with the track under consideration excluded from the fit.

The change in the impact parameter χ^2 distributions between reconstructions with the standard and updated parametrisations and between the start and end of Run 1 is small. Again there are several contributing uncertainties beyond those of the VELO hit positions, such as the uncertainty on the primary vertex fit and the errors of the other detectors.

The primary vertex χ^2 per degree of freedom is shown in Fig. 3.27. There is a slight shift between 2010 and 2012 but very little difference between the standard and updated error parametrisations with the 2012 sample.

Overall it therefore seems as if the error parametrisation has had very little effect on the reconstructed quantities for each track. As noted previously this is to be expected due to the other potential sources of errors.

3.5.2 Selected candidates

Beyond the individual particle tracks the effect on reconstructed short lived decays has been investigated. A sample of previously selected and reconstructed charmed hadron candidates was taken from a run in November 2012. The reconstruction with the updated error parametrisation was run on this sample and then a selection made for the $D^0 \to K^-\pi^+$ decay. This was compared to the same selection on the data reconstructed with the standard error parametrisation. As the data had been previously selected the purpose was not to examine the efficacy of cuts but to see if the distributions of any of the reconstructed quantities, in particular the combined D^0 candidate, had been affected by the change in error parametrisation.

No significant changes were found between the two reconstructions. The tracks of the decay products had χ^2 distributions very similar to that in Fig. 3.26. Other reconstructed quantities, such as the D^0 mass, momentum or impact parameter χ^2 showed no significant changes.

3.6 Conclusions

A data driven method to measure and parametrise the hit resolution of the VELO has been tested on simulated data and shown to reproduce the true distribution to a good degree of accuracy. The cluster resolution was measured on a sample of data taken in 2010, at the start of the LHC Run 1 data taking period, with the results being consistent with expectations. The method was subsequently applied to samples of data from different periods of Run 1. The resolution was found to change greatly over the data taking period. Initially the resolution improved in some angular regions but overall the resolution measured at the end of 2012 was somewhat worse than in 2010.

The reasons for the degradation of the sensor resolution have been briefly explored with simulated data. Increases in capacitive coupling between the strips and charge diffusion in the bulk both suggest the trends seen though they are not conclusive. Of course the effects of radiation damage are not universal throughout the detector and a more careful study is needed to assess the changing resolution with fluence. However the relevant output of this study for the physics reconstruction is obtaining a new error parametrisation. This has been extracted and more accurately reflects the true performance of the sensors. This has been

implemented in the LHCb reconstruction package for study by the tracking group.

Both the degrading resolution and the updating of the error parametrisation has been shown to have some effect on the track reconstruction, particularly for the VELO segments of tracks. However, for whole tracks and subsequent reconstructed quantities (such as the impact parameter χ^2) the effects become rather minimal as other contributions to the uncertainties dominate.

The resolution change with fluence merits some further comment as it was not entirely expected. Of the two mechanisms suggested in this chapter, the changes in electric field in the bulk silicon is the one that has been studied through measurements of the effective depletion voltage [92] and was expected. The suggestion of increased capacitive coupling between the strips is however less-well understood. Test-beam data on irradiated pixel sensors [109] has not shown such a large increase as would be required to explain these results. If the mechanism is due to effects on the surface of the sensors, reproducing them in a test environment will be difficult as one will not be able to easily reproduce the conditions for irradiation that the sensors in the experiment experience.

The second metal layer effect may also be a contributing factor. It has been found that after irradiation the read out lines become coupled to the silicon bulk, reducing the charge collected by the strips [92]. This was also an unexpected observation. The effect would mostly dominate at large pitches in the r sensors where there is a large number of routing lines perpendicular to the sensor strips. There is some evidence of the resolution at the largest pitches degrading considerably (see for example Fig. 3.19). Of course there could be changes in the bulk silicon, for example due to increased trapping of charge carriers [110] although a large effect is not expected.

It is possible that the bias voltages applied to the sensors at the end of Run 1 (150 V) was not high enough to fully deplete the sensors. Such a situation could be expected to severely impinge on the resolution as the charge collection efficiency would be rather compromised. In the future it would be worthwhile to study the resolution as a function of the applied bias voltage to assess the effect of the sensors not being fully depleted.

For Run 2 the error parametrisation will need to be measured again. The bias voltage of the most irradiated sensors has been changed to account for the expected increase in the effective depletion voltage [111]. After some data-taking further updates will be needed due to the continued radiation damage. A more pressing concern is the simulation which must be tuned to more accurately reflect the changing condition of the detector.

Chapter 4

Lifetime Measurements

The time-dependent observables A_Γ and y_{CP} were introduced in chapter 1. The former can be determined by the comparison of the effective lifetimes of the D^0 and $\overline{D}^0$ decaying to a CP eigenstate K^+K^- or $\pi^+\pi^-$; y_{CP} is the ratio of the effective lifetimes of the Cabibbo favoured $K^-\pi^+$ and singly Cabibbo suppressed modes. For both cases one is making precision lifetime measurements via a fit to measured lifetime distributions.

The overall analysis strategy that was common to each of these analyses is outlined in this chapter, prefaced with a brief description of some statistical techniques that feature heavily. The techniques described are sWeights and kernel density estimation. A description of the analysis method will be given, including how biasing effects of the selection are corrected in the fit. Some non-standard functions used in the analysis will also be introduced.

In this thesis a result for each observable is presented. Although the analyses are very similar they were completed at separate times; the first was the result for A_Γ using data collected in 2011, followed by y_{CP} with the same data. The details of the fits changed between these results and are included in chapters 5 and 6.

4.1 Tools

The analyses make extensive use of sWeights and kernel density estimation for a variety of purposes. As these subjects are not necessarily common knowledge they will be outlined in the following sections.

4.1.1 sWeights

The aim of the sWeight [112] technique is to unfold the distribution of some contribution to the data (signal or a background) in a variable, given a maximum likelihood fit to a group of other discriminating control variables. This is necessary

if one does not have a prior with which to fit that variable, for example if its shape is unknown. For example in the analyses presented in this document the shapes of the distributions of the backgrounds in decay time (the variable to be measured) are not known and cannot be discriminated directly in the fit. However in the D^0 mass variable the background probability density function is known and so provides a control with which to unfold the background distribution in decay time. The interested reader can find the derivation and a detailed discussion of the sWeight technique in [112]. The sole purpose of this section is to introduce the sWeight and demonstrate its efficacy.

The sWeight for a single event for the class n, calculated using the group of discriminating variables m is

$$ {}_sP_n(m) = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(m)}{\sum_{k=1}^{N_s} N_k f_k(m)}. \tag{4.1} $$

The sums are over the number of species contributing to the data (signal or background), with N_k being the estimated number of events of that species. $f_j(m)$ is the probability function in the variable m for the class j resulting from a maximum likelihood fit. The matrix V_{nj} distinguishes the sWeight from a simple class probability as it contains information about the overlap of the classes. It can be introduced as the covariance matrix of the likelihood fit to the discriminating variable,

$$ V_{nj}^{-1} = \frac{\partial^2(\mathcal{L})}{\partial N_n \partial N_j} = \sum_{e=1}^{N} \frac{f_n(m_e) f_j(m_e)}{(\sum_{k=1}^{N_s} N_k f_k(m_e))^2}. \tag{4.2} $$

As noted in the paper [112] it is more accurate to calculate V_{nj} explicitly, replacing the sum over the events with an integral, rather than simply take the covariance matrix of the fit.

As an example consider two independent variables, the discriminating variables m and control t with two contributions; signal and background. The data for these two variables are shown in Fig. 4.1 with the functions used to generate signal and background overlaid. In t both signal and background are exponential lifetime decays. A fit to the signal generated data (with no background) yields $\tau = (404.6 \pm 4.0)$fs. The aim is to fit the signal component of t in the total sample with a single exponential but the precise background shape is unknown to the experimenter so it cannot be included in a simple fit to the data; the signal distribution must be unfolded.

Initially one performs a maximum likelihood fit to the m variable using some

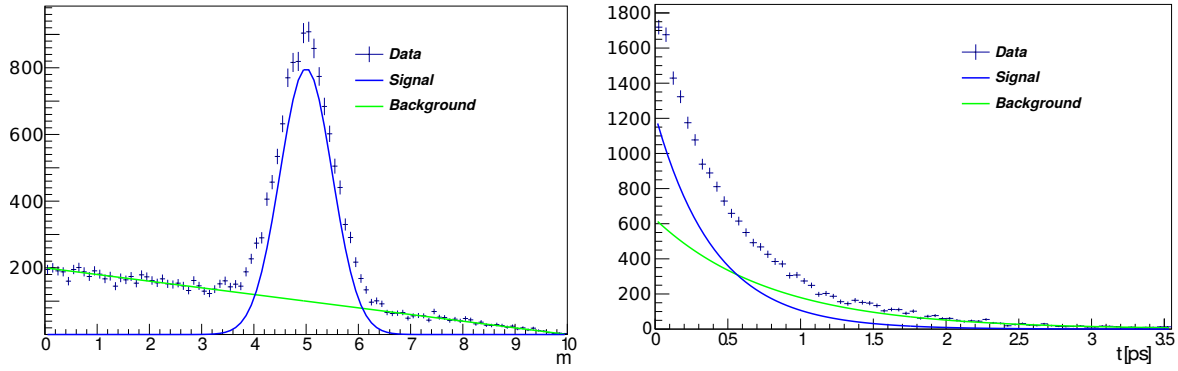


Figure 4.1: Generated data for two variables m (left) and t (right) with the true distributions for the signal and background components shown.

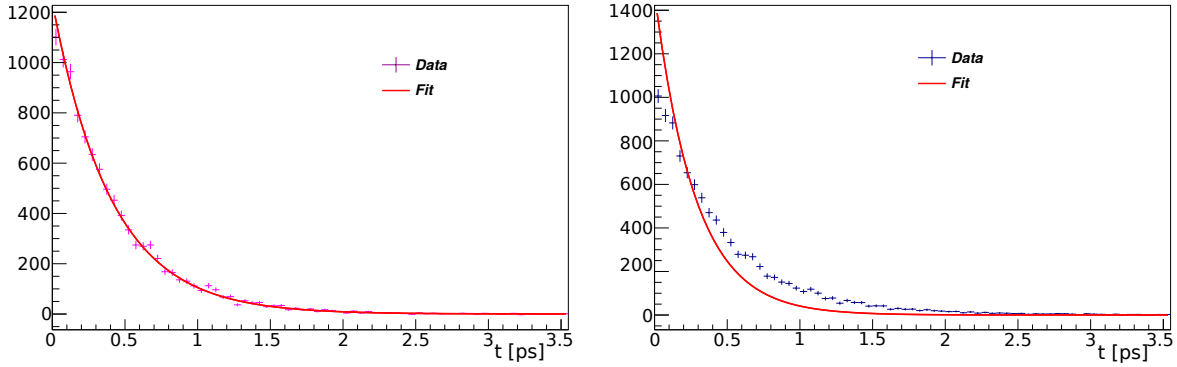


Figure 4.2: The data weighted with sWeights (left) and probability (right) plotted in the variable t. Overlaid are fits of single exponential functions. The fit to the sWeights matches the generated signal data, that to the probabilities does not.

sensible priors that describe the data. The data are then weighted by sWeights for the signal component. The t variable can then be fitted with just the signal PDF. The sWeighted t distribution is shown in Fig. 4.2 with the fit result. The fitted lifetime is $\tau = (405.3 \pm 3.9)$fs agreeing with the fit to the generated pure signal data. On the right of Fig. 4.2 is the data weighted by the probability of being signal as determined by the fit to the variable m. That fit is rather unsuccessful, yielding $\tau = (279 \pm 1.1)$fs; the probability distribution is contaminated by the background data.

One has to be aware that the discriminating variables (in this case m) must be uncorrelated with the distribution of the variable to be found. If this is not the case then some prior knowledge of that variable is required and the sWeight technique is not appropriate.

Finally one can see the technique applied to the analysis fit itself in Fig. 4.3. The plot is taken from a simulated experiment used to test the fitter. Data was generated using the fit configuration for the $K^-\pi^+$ final state with the unknown

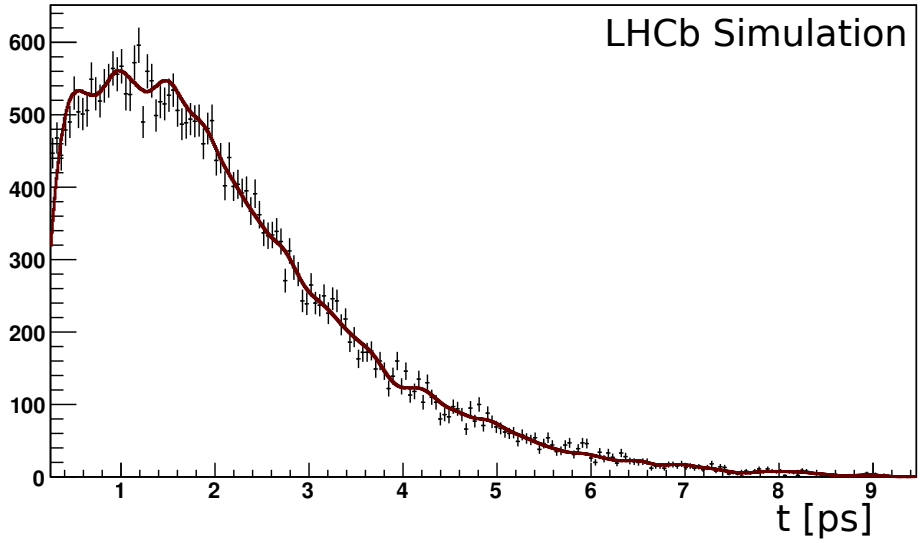


Figure 4.3: A comparison of an sPlot with generated data from a pseudo-experiment produced for the 2011 y_{CP} analysis. The data points in black are the generated data. The biased lifetime distribution of the combinatoric background which is unfolded from the data using the D^0 mass and Δm as discriminating variables is shown in brown. The sWeights have been interpolated with kernel density estimates, described in section 4.1.2, to produce a smooth distribution.

PDFs approximated with a combination of parametric PDFs. The fitter was then run on this generated data in the normal manner producing the sPlots (sWeighted distributions) of the backgrounds. Fig. 4.3 shows a comparison between these backgrounds and the PDF used in the generation. The sPlot closely matches the generated data.

4.1.2 Kernel Density Estimators

Having obtained a distribution from the sPlots one might wish to produce a continuous function that can be used in an unbinned fit. To this end kernel density estimators [113] are used. A more comprehensive overview can be found in [113].

To each data point is attached some kernel function. One simply sums these kernels to arrive at a continuous smooth PDF. The choice of function is up to the analyst; for the works presented in this document Gaussian kernels are used.

Fig. 4.4 shows the idea applied in a simple example. A Gaussian function (in blue) of mean 0 and width 1 has been randomly measured 10 times leading to the distribution of measurements on the axis. To each of these is applied a kernel of suitable width which are summed to produce the continuous function in green, an estimation of the true distribution; the accuracy here is limited by the sample size.

As the size of the bins in a histogram can distort the visualisation of data so the width of the kernels can impinge on the veracity of the resultant function. An

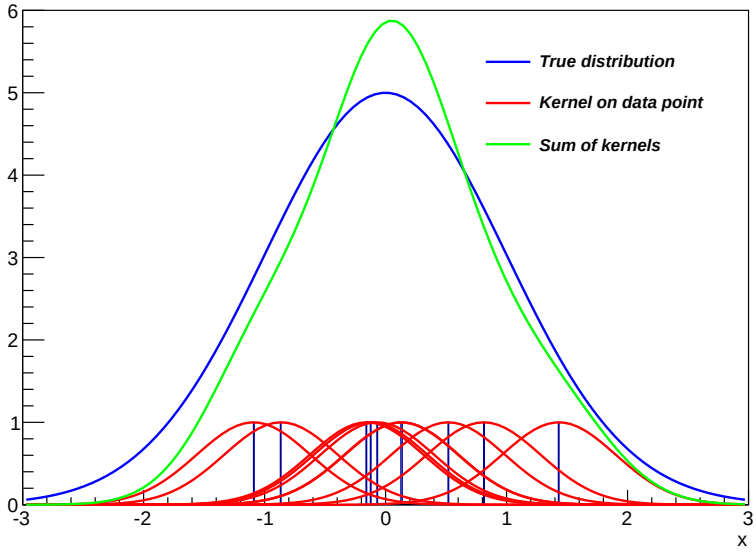


Figure 4.4: The Kernel Density Estimator technique. To the ten measurements (bars on the axis) are applied kernels (red). The sum is shown in green along with the true distribution in blue.

example is shown in Fig. 4.5 with a demonstration of over or underestimating the kernel width. For a normally distributed set of data the optimum width (standard deviation of the Gaussian kernel) h is given by [113]

$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}}, \tag{4.3}$$

where $\hat{\sigma}$ is the standard deviation of the data sample, n is its size.

For a more general distribution the optimum bandwidth is given by h^*

$$h^* = \left(\frac{R(K)}{\sigma_K^4 R(f'')n}\right)^{\frac{1}{5}}, \tag{4.4}$$

where K is the kernel function, σ_K its variance and f'' the second derivative of the true distribution of the data. R is the "statistical roughness" of the data defined as

$$R(f) = \int_{-\infty}^{+\infty} f(x)^2 dx. \tag{4.5}$$

As derivatives of the true distribution are used to ascertain the statistical roughness one is compelled to start with an initial estimate of the true shape. Equation 4.3 can be used to produce said estimate, with its derivatives then appearing in equation 4.4 to create a more optimal function. One can then iterate

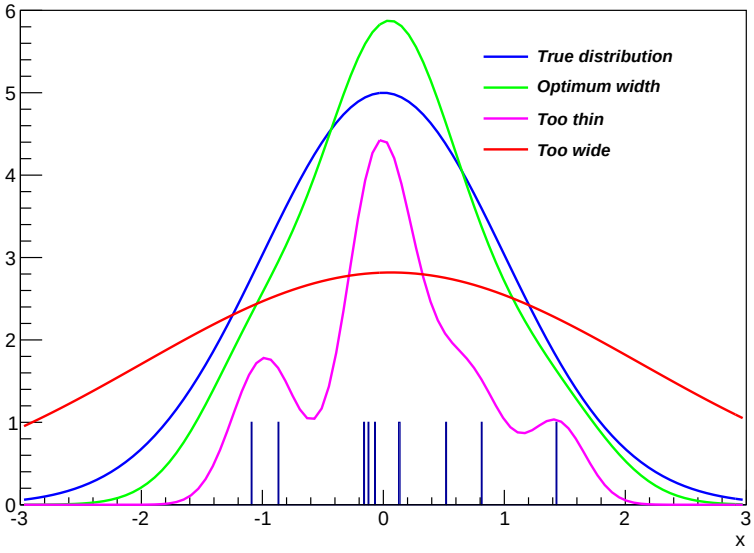


Figure 4.5: The importance of the kernel bandwidth. The true distribution is in blue, with the function using the optimum bandwidth (equation 4.4) in green. The red and magenta curves are the functions resulting from the kernels being too wide and narrow respectively.

until a reasonable representation of the true distribution has been achieved.

In reality the optimum bandwidth varies across the data. For example in the tails of distributions there are few data points so wider kernels are more appropriate than at the centre. Adaptive bandwidths and the relevant machinery to implement them are introduced in [113]. However, they were not used in the analyses presented here so will not be discussed further.

Applying kernels at a border where the data has been cut poses a problem as the kernel distribution should reflect this boundary. An example of the problem is shown in Fig. 4.6 For each individual kernel one could apply the cut at the boundary and renormalise the remaining distribution. This has the advantage of keeping a smooth distribution at the border but it pushes the mean of the distribution away from the boundary, distorting the total estimated distribution around this region; the effect is apparent in the left plot of Fig. 4.6. Another option would be to reflect the cut part of the kernel at the boundary back into the rest of the kernel. The disadvantage of this approach is that the first derivative of the kernels is zero at the boundary, shown on the right in Fig. 4.6, which again would not reflect the data. For these analyses the former option was chosen; the dataset is sufficiently large that the kernels widths are small so the distortions at boundaries due to applied cuts are small.

Thus far only one dimensional kernels have been considered. For a two

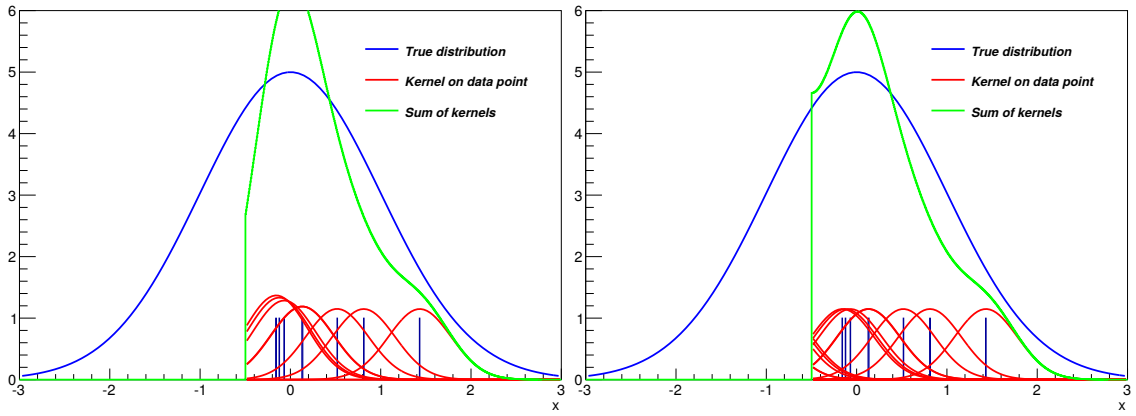


Figure 4.6: The treatment of boundaries in kernel density estimation. Shown is the same data as in Fig. 4.4 but a cut has been applied. On the left the kernels have been renormalised leading to the total distribution being distorted away from the boundary. On the right the kernels have been reflected at the boundary, causing the total distribution to have a first derivative of zero at the border.

dimensional kernel one can could treat the two variables independently, in which case a width for each is sufficient. However if there is strong correlation between the two variables it would be more accurate to use a complete two dimensional treatment. The result is that the optimum bandwidth is given by a matrix which describes a rotation of the kernel. However in the work presented here the variables of any two dimensional distributions to which kernels are applied are only mildly correlated so the simple treatment is sufficient.

An example of the kernel density estimators used in the analysis can be seen in Fig. 4.3. An sPlot has had kernels applied to it to form a smooth function for use in an unbinned fit. The resulting function matches the generated data closely, demonstrating the efficacy of both the sPlot and kernel techniques.

4.2 Analysis Method

4.2.1 Introduction

The measurement of the effective D^0 lifetime τ as it decays to a specific final state is achieved via the fitting of a single exponential decay function, $e^{-t/\tau}$, to the measured decay times t. The extracted lifetimes are then used to calculate the observables of interest, y_{CP} and A_Γ

The initial D^0 flavour is ascertained by searching for the strong decay $D^{*+} \to D^0\pi_s^+$, the charge of the pion indicating D^0 or $\overline{D}^0$. The pion is referred to as 'slow' (π_s^+) as it is of low momentum compared to the D^0 decay products due to the small Q value of the transition. This method of tagging the D^0 flavour

introduces a background of incorrectly tagged D^0 candidates due to the potential of selecting a pion that originated in the proton-proton collision. Such an occurrence is a mis-tag of the D^0 candidate and is referred to as the "random slow pion" background.

An additional source of background in the analysis is from D^0 candidates formed by combinations of tracks that are not from the signal decay. This component is referred to as combinatoric background. Contributions also arise from real decays that have been mis-reconstructed such that the candidate resembles the signal decay. For example multi-body decays may have been reconstructed with one or more daughter particles emitted or the daughter particle may have been mis-identified. Finally a background arises due to charm produced in the decay of a long lived particle (B) which is not reconstructed resulting in an inaccurate measurement of the D^0 candidate's decay time.

The analysis is performed using an unbinned maximum likelihood fit in two stages, each fitting two variables. The first stage is to fit the D^0 mass, m_{D^0}, and the difference between the D^{*+} and D^0 masses, Δm. Fig. 4.7 shows the distribution of these variables. m_{D^0} allows for the separation of real D^0 from combinatoric background and mis-reconstructed decays. The Δm variable can be used to preferentially select correctly tagged D^0 candidates. If the π_s^+ that is combined with the D^0 does originate from a D^{*+} decay then the distribution will peak in Δm at 146 MeV, the difference in rest masses of the D^{*+} and D^0. Other random pions from the interaction point will lead to a continuous distribution that cuts off at the pion rest mass. In Fig. 4.7 the π_s^+ background can be seen in regions A, with combinatoric background dominating in regions B. Contributions to the data of real D^0, combinatoric, mis-reconstructed and π_s^+ background are henceforth referred to as *classes* (*cl*). The outcome of the first stage of the fit is to assign a probability to each event that it is from a particular class.

Having separated the signal from the non-D^0 backgrounds in the first stage, the second stage is used to extract the effective lifetime of the signal. Two more variables are fitted to do so: the D^0 decay times and the natural logarithm of the χ^2 of the D^0 impact parameter with respect to the primary vertex, $ln(\chi^2_{IP})$.

D^0 candidates (both signal and slow pion background) can be produced at the interaction point from the collision, called *prompt* candidates, or from decays of B particles, which are given the sobriquet *secondaries*. Prompt and secondaries are the *subclasses* (*scl*) of the signal and random slow pion classes. A summary of the classes and subclasses is shown in table 4.1. Secondary candidates have longer measured decay times than the true D^0 decay time and thus would bias the measurement as the intermediate B decay vertex is not reconstructed. Consequently a variable is required that will allow the separation of the prompt

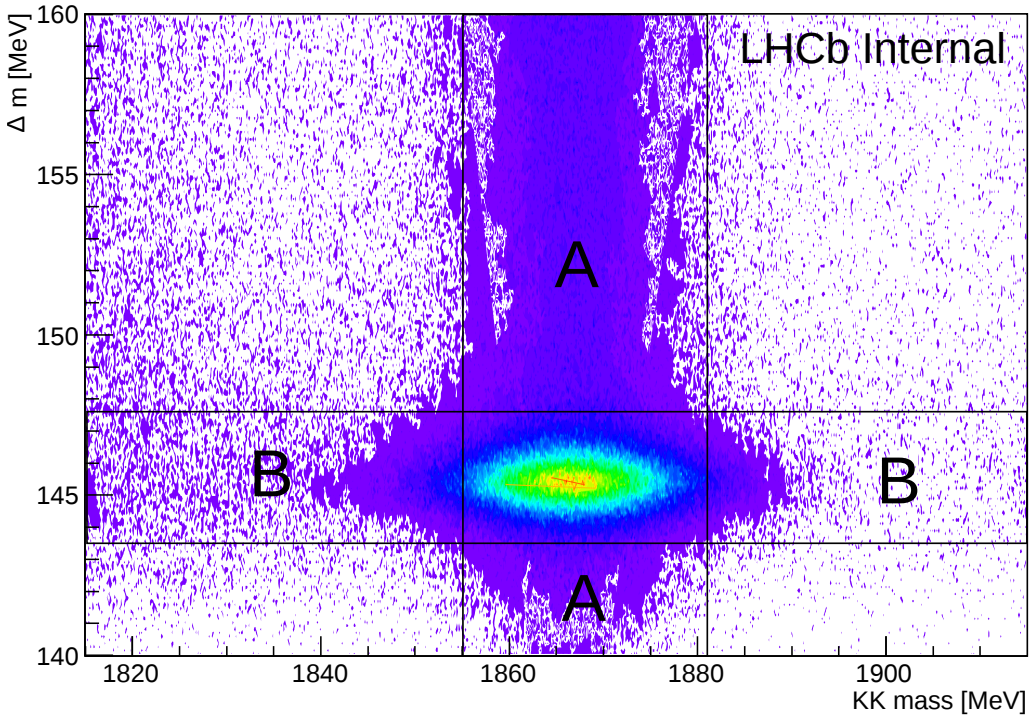


Figure 4.7: Mass and Δm distribution for a sample of $D^0 \to K^+K^-$ in the 2011 data. The signal peak can be clearly seen in the centre.

and secondary subclasses.

As the secondary D^0 candidates come from the decay of another particle they have a large impact parameter as shown in Fig. 4.8. A D^0 produced at the interaction point (left) should have an impact parameter of 0 as its momentum vector points back to the primary vertex. A candidate produced by a B decay (right) however points away from the primary vertex giving a large impact parameter. A large impact parameter implies a similarly large χ^2_{IP}, defined as the measured impact parameter divided by the uncertainty on the quantity (essentially an indication of how significant the measured impact parameter is). The natural logarithm is taken to give a shape that is readily fitted. Thus the variable to discriminate prompt and secondary is $ln(\chi^2_{IP})$.

The $ln(\chi^2_{IP})$ of the secondary exhibits a profound decay time dependence; the longer lived the B, the greater on average the D^0 impact parameter is. For this reason the second stage fit is in these two dimensions with the $ln(\chi^2_{IP})$ variable being described by a PDF that is conditional on proper time.

2010 analysis

The first measurements of y_{CP} and A_Γ at LHCb were made using $29\,\text{pb}^{-1}$ of data with collisions at 7 TeV collected in 2010. That analysis served as a proof of the principal of the method and the analysis of the 2011 data broadly follows the

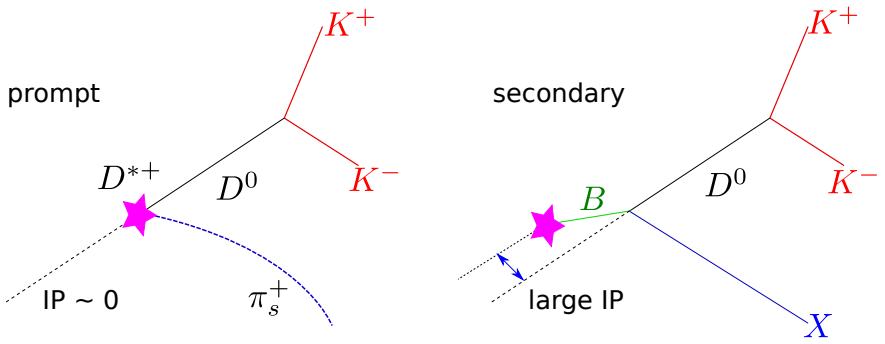


Figure 4.8: The separation of prompt and secondary. On the left a prompt D^0 candidate points back to the primary vertex. On the right a secondary candidate has a large impact parameter.

Class	Subclass
Signal	prompt secondary
slow pion	prompt secondary
Combinatoric	-
Mis-reconstructed	-

Table 4.1: The classes and subclasses of the fit.

template of that effort [114]. The results are [115]

$$y_{CP} = (5.5 \pm 6.3_{stat} \pm 4.1_{syst}) \times 10^{-3}$$
$$A_\Gamma = (-5.9 \pm 5.9_{stat} \pm 2.1_{syst}) \times 10^{-3}. \quad (4.6)$$

The statistical precision of these numbers was already competitive with that of the HFAG averages at the time [41]. The largest uncertainties were due to combinatoric and secondary backgrounds.

The 2010 data sample was small compared to the dataset that has been analysed here. In that analysis very tight cuts were applied around the correctly tagged D^0 peak in the Δm variable and combinatoric and mis-tag backgrounds were assumed to be negligible. The much larger numbers of candidates that followed required a much more thorough fitting method to be employed in the subsequent analyses. A proper description of the backgrounds had to be included, as did more carefully considered models for each of the variables.

4.2.2 Data and selection

The data set analysed was collected in 2011 and comprises of $1\,\text{fb}^{-1}$ of proton-proton collisions at 7 TeV centre of mass energy. Section 5.1 lists the cuts applied to the data in the trigger and stripping (see sections 2.2.8 and 2.2.9 for descriptions of both).

The data are selected from exclusive D to two-body trigger and stripping lines. Initially there are no requirements on the L0 decision. In HLT1 the candidate must be TOS (the trigger must be fired by the signal candidate) on the Hlt1TrackALLL0 line. This line reconstructs the tracks of the decay products and imposes some minimum momentum and track quality cuts (track fit χ^2 and minimum number of hits in the detector). It also requires a minimum impact parameter (the closest distance between the track and the interaction point) for the track to select those that likely originate from the decay of a long-lived particle. From HLT2 the candidate must come from one of the exclusive D to two body lines which combine the tracks of HLT1 into the decaying particles. The trigger lines come in pairs, a default and a wide-mass line. The latter allows for a wider range of D^0 candidate masses and is pre-scaled by 0.1 so as not to waste the trigger bandwidth. The wider mass range can be used to study background events which do not peak at the D^0 mass. The trigger aims to produce D^0 candidates from pairs of tracks using a fast reconstruction compared to that performed offline. After the triggering the data is processed offline where tighter cuts are applied with the aim of producing high purity samples of the decays of interest. It is at this stage that a π_s^+ candidate is selected in order to tag the initial D^0 flavour.

For selecting D^0 candidates (or any other long lived particle) from proton-proton collisions, where potentially many different particles and tracks are produced, one primarily seeks a displaced decay vertex. This is done quickly by applying impact-parameter (IP) cuts on the daughter tracks in HLT1 and flight distance cuts on the reconstructed D^0 candidate in HLT2 and offline. This search for a displaced vertex introduces a time dependent acceptance function which is discussed further in section 4.2.3. The acceptance function must be taken into account as the selection has a higher efficiency for longer lived candidates.

Finally some extra selection cuts are applied after the stripping. The most pertinent is to select one candidate from each event. For some collisions more than one D^0 or D^{*+} candidate may be constructed (for example combining the D^0 with more than one possible π_s^+ candidate). Only one randomly selected candidate per event is used in the fit. A minimum decay time cut of 0.25 ps is used to remove a region where very few events are selected. This prevents the fit from becoming unstable in the low lifetime region where statistics are low. For each final state the

particle identification requirements on the daughters are adjusted to further increase the purity of the data set. An upper limit on the radial (perpendicular to the beam direction) flight distance of 4 mm is imposed. This value is just less than the distance between the LHC beam and the inside edge of the VELO RF foil. The cut therefore prevents very long lived D^0 candidates that have originated from interactions between particles and the VELO material.

The data are split into samples that reflect the experimental conditions, the divisions being motivated by the magnet polarity or the occurrence of LHC technical stops. During these technical stops the alignment or trigger changes could influence aspects of the fit. As the A_Γ measurement is a flavour tagged analysis the samples are also split into D^0 and $\overline{D}^0$ to be fitted separately. The same splits were adopted for the y_{CP} analysis.

4.2.3 Swimming

In order to select long lived particles (be it a B, D or anything else) from proton collisions, where most interactions do not produce something of interest to the analysis, one usually searches for a displaced vertex from which two or more tracks originate. For reasons of speed in the HLT this is achieved initially by imposing a minimum impact parameter requirement on individual tracks; combining tracks to try and fit a vertex takes time, especially if several combinations have to be tried. Once a more complete reconstruction can be done, such as in HLT2 and offline, more accurate vertexing requirements can be imposed. These kinds of requirements invariably create some lifetime bias in the collected data. In particular, the trigger requirements on the daughter IP χ^2 and D^0 flight distance favour a longer lived D^0. To measure the unbiased mean lifetime of the D^0 one therefore has to include the effects of selection in the fit by means of an acceptance function $f(A)$. The acceptance is found on a per-event basis using the data driven 'swimming' [116, 115] method.

To obtain $f(A)$ with this technique one has to be able to rerun the trigger and stripping in the same conditions as for the online data taking and subsequent processing. Having the LHCb HLT implemented in software allows for this. The L0 hardware trigger only has very basic requirements on the calorimeter and muon systems and should not introduce any lifetime bias.

The principle of the method is shown in Fig. 4.9. For a single event the acceptance resembles a simple top hat with values of 0 or 1; for a given decay time it is either accepted or rejected. The D^0 production vertex is moved along the flight direction, changing the D^0 decay time for the same kinematics. The trigger and stripping is rerun separately at each point along this variation and the

decisions are recorded, constructing the top hat as shown. On the left the primary vertex has been moved such that the event fails the impact parameter requirement on one of the daughters. In the middle is the point where the decay time is such that the event is at the cut values and the selection decision changes. This decay time, t_{min} is the first *turning point*. On the right is shown the measured decay time where the event has passed the selection. The acceptance is stored as a series of turning points where the decision changes from rejected to accepted and vice versa.

To account for finite computing resources the search for the turning points is done in steps. Initially a rough search is performed; a granularity of 4mm for the region ±200mm around the PV and a granularity of 40mm for larger regions up to ±600mm. For those areas where a turning point is found (where the selection decision has changed) the precision is improved with smaller search regions over four iterations giving a final turning point resolution of $\frac{15.625\mu m}{\sqrt{12}}$ [72]. This represents the resolution of the smallest search interval. The initial search region does imply that top hats of less than 4mm width might not be picked up though the effect of this is believed to be negligible. Such a small top hat implies that the D^0 has flown far enough to be accepted but then fallen into the region of another primary vertex to be rejected very quickly. However the interaction region is comparatively large ($\sim$ 30mm) and the interaction multiplicity small ($\sim$2 on average) that such an occurrence is unlikely.

As mentioned previously the swimming must be performed with exactly the same trigger setup as the original selection. The same trigger software versions and TCKs (trigger configuration keys that record the setup of the trigger) are used as in the data taking periods, along with the DDDB (the description of the detector geometry) and LHCBCOND (the LHCb conditions database which records any time dependent changes to the detector such as alignment configurations) tags, which together describe the layout of the detector and running conditions. The trigger and offline stripping are swum separately. Additionally an upper lifetime acceptance effect is taken into account through an analytical calculation. A very long lived D^0 may lead to its daughter final state tracks no longer having the minimum number of hits in the VELO (5 in the HLT and 3 in the stripping), something that cannot be tested by only moving the primary vertex. However considering the event geometry and flight direction allows for the upper lifetime limit of an event to be calculated.

Finally the acceptance from each stage (trigger, stripping and VELO) are combined by a simple multiplication of the top hat functions for each event to give the overall per-event acceptance function which can be used in the fit.

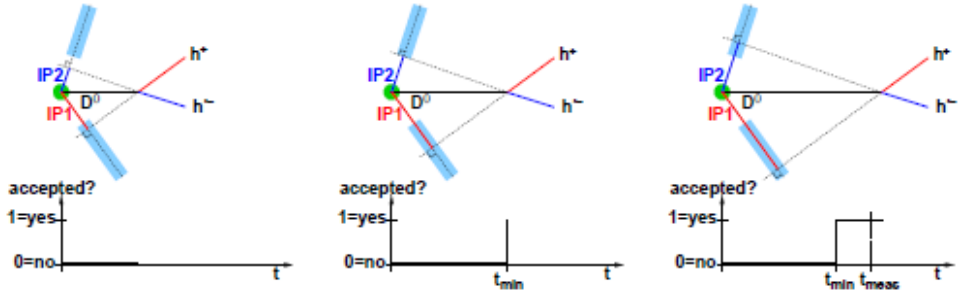


Figure 4.9: The swimming method. The primary vertex is moved along the D^0 flight direction. On the left the decay time is such that one of the daughters (h^-) fails the IP cut. In the middle is the decay time, t_{min} at which it first passes the cut and there is a turning point in the acceptance function. On the right is the measured decay time showing the event passing the selection and the subsequent top hat shape of the function. Reproduced from [115]

4.2.4 Total Fit PDF

In all, four variables are fitted to extract the signal lifetime: m_{D^0}, Δm, $ln(\chi^2_{IP})$ and t. The first two are assumed to be uncorrelated with the second two so the fit is performed in two stages. m_{D^0} and Δm are fitted together to separate the classes and the results used in the fit for t and $ln(\chi^2_{IP})$ to separate the subclasses. There are approximately 42 fitted parameters, depending precisely on the configuration of the fit which varies slightly for the different final states and between the A_Γ and y_{CP} analyses. The true numbers are given in sections 5.2.4 and 6.1.5.

For each stage the total PDF is the sum of the PDFs of each class (and subclass) weighted by their respective fractions p_{cl}. The first stage is given by

$$f(m, \Delta m) = \sum_{cl} p_{cl} f(m, \Delta m | cl). \tag{4.7}$$

As the two variables are considered independent $f(m, \Delta m|cl)$ is the product of the PDFs of the two variables for each class, $f(m|cl)$ and $f(\Delta m|cl)$. The specific configuration of the PDFs can be found in the chapters pertaining to each individual analysis. Using Bayes' theorem one finds that for a single candidate the probability to belong to a given class is

$$P(cl|m_{D^0}, \Delta m) = \frac{f(m_{D^0}, \Delta m|cl) p_{cl}}{f(m_{D^0}, \Delta m)}. \tag{4.8}$$

In the second stage the total fit PDF for a single event of given mass and Δm is described by

$$f(t, ln(\chi^2_{IP}), A|m_{D^0}, \Delta m) = \sum_{cl} f(t, ln(\chi^2_{IP}), A|cl) \cdot P(cl|m_{D^0} \Delta m), \tag{4.9}$$

where A describes the acceptance.

For combinatoric and mis-reconstructed background components $f(t, ln(\chi^2_{IP}), A|cl)$ is not an analytic function but instead an interpolated numerical distribution. The method for ascertaining these PDFs is described in section 4.2.7. The signal and π^+_s background must be further deconstructed into their subclasses for prompt and secondary

$$f(t, ln(\chi^2_{IP}), A|cl) = \sum_{scl} P(scl|cl) \cdot f(t, ln(\chi^2_{IP}), A|cl, scl), \tag{4.10}$$

where $P(scl|cl)$ is the fraction of the subclass within the class. Breaking down the second term gives

$$f(t, ln(\chi^2_{IP}), A|cl, scl) = f(ln(\chi^2_{IP})|t, A, cl, scl) \cdot f(t|A, cl, scl) \cdot f(A|cl, scl). \tag{4.11}$$

The first function is a conditional PDF describing the $ln(\chi^2_{IP})$ variable as it varies with decay time. The second is the decay time distribution for a given acceptance function, the third term being that acceptance function which is found by the swimming. Together the last two give the *biased* lifetime PDF. The content of $f(t|A, cl, scl)$ is expounded upon in section 4.2.6.

4.2.5 Acceptance functions for analytic PDFs

The output of the swimming provides a per event acceptance function from which the total acceptance of the data sample can be obtained. However one needs to find the individual acceptance functions for each class and subclass $f(A|cl, scl)$ to insert into equation 4.11. An iterative method is used to unfold the distributions.

$f(A)$ is the probability density of finding a particular acceptance function. A more convenient way to present this in the fit is to break the function down into some constituents. The top-hats for each candidate that emerge from the swimming can be described as points in time where the trigger decision has been found to change (rejected to accepted and vice versa), the turning points. For n turning points the acceptance can therefore be described by the first turning point (TP_1) and subsequently the difference between successive points

$$\begin{aligned} f(A) &= f(TP_1, TP_2, ...TP_n) \\ &= f(TP_1) \times f(TP_2 - TP_1) \times ...f(TP_n - TP_{n-1}) \\ &= f(TP_1) \times \prod_{i=2}^{n} f(TP_i - TP_{i-1}). \end{aligned} \tag{4.12}$$

These turning point distributions can be parametrised in any way the analyst

deems to be most convenient, in this case decay time t. For the purposes of keeping the analysis manageable only candidates with two turning points are used (those with a single top hat acceptance). Few events are cut out from this as the radial cut tends to be the upper limiting factor.

For a given subclass of a class the per event probability is

$$P(scl|cl, m, \Delta m, ln(\chi^2_{IP})\, t, TP_1) = \frac{f(m, \Delta m|cl)g(t|A, cl, scl)g(ln(\chi^2_{IP})|t, A, cl, scl)TP_1(t|cl, scl)P(scl|cl)}{\sum_{scl} f(m, \Delta m, cl)g(t|A, cl)g(ln(\chi^2_{IP})|t, A, cl)TP_1(t|cl, scl)P(cl)}. \quad (4.13)$$

If the turning point distribution $f(TP_1)$ on the right is correct then it should match the distribution of probabilities produced by equation 4.13 when plotted in the TP_1 variable. Starting from some estimate of $f(TP_1)$ the probabilities are calculated and compared to the estimate. If different this new probability distribution is plotted in $f(TP_1)$, kernels applied to make a continuous function and then inserted back into the equation to calculate a new set of probabilities. This process is iterated until successive $f(TP_1)$ estimates are similar at which point it is thought to have converged and a good approximation of the true distribution has been found.

The initial estimate of the turning point function is taken from data. Prompt candidates mostly lie at small $ln(\chi^2_{IP})$ and should not exhibit significant time dependence. Their turning point distributions are therefore estimated from the data with a cut applied at $ln(\chi^2_{IP}) < 1$. As secondaries do have significant time dependence and occupy a large range of $ln(\chi^2_{IP})$ all candidates are used to produce the estimate $f(TP_1)$.

The parametric functions describing the lifetime and $ln(\chi^2_{IP})$ on the right in equation 4.13 are of course not known (they are to be fitted). However to produce the turning point functions it is sufficient to use estimates given by the results of a previous fit. The full fit can be run several times and with each iteration the starting values for these functions are updated, improving the accuracy of the generated turning point PDFs.

The turning point difference function is calculated in exactly the same manner, just inserting $f(TP_2 - TP_1)$ into equation 4.13 instead of $f(TP_1)$. The sum of the individual turning point PDFs can be compared to the data as in Figure 4.10. As can be seen good agreement is observed suggesting the method has produced something reasonable.

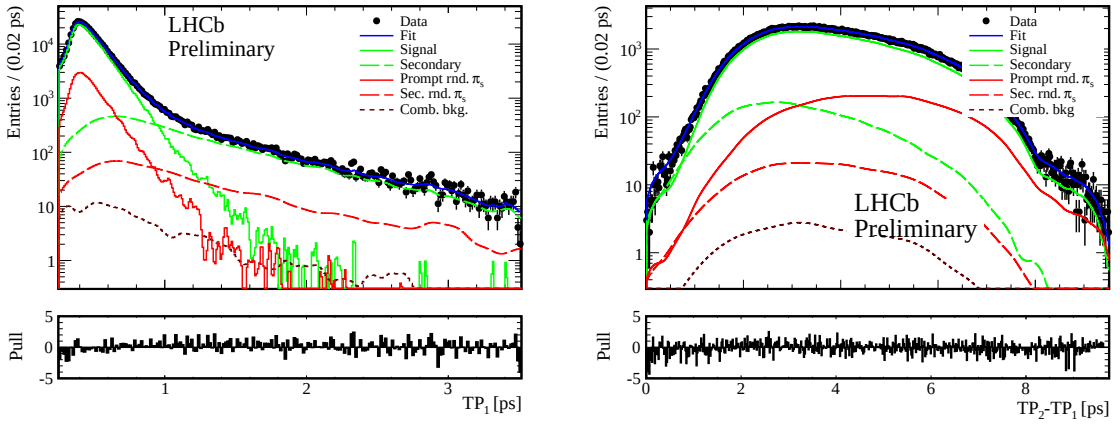


Figure 4.10: Plots of the distributions of TP_1 (left) and $TP_2 - TP_1$ (right) overlaid with the fit PDFs $f(TP_1)$ and $f(TP_2 - TP_1)$ for each component and their sum for the $K^-\pi^+$ final state in the 2011 y_{CP} analysis. The data is shown with the subclass PDFs and their sum overlaid. As can be see the resultant total function for these two variables matches the data to a good degree.

4.2.6 The lifetime PDF

Prompt

For a single prompt event the normalised lifetime PDF is

$$f(t|A, cl, prompt) = \frac{\frac{1}{\tau}e^{-t/\tau}\Theta(t)}{\int_{-\infty}^{+\infty} \frac{1}{\tau}e^{-t/\tau}\Theta(t)\mathrm{d}t}. \tag{4.14}$$

The Heaviside function $\Theta(t)$ is 1 for the times at which the candidate is accepted, 0 otherwise. Performing the integral in the denominator gives

$$f(t|A, cl, prompt) = \frac{\frac{1}{\tau}e^{-t/\tau}}{\sum_i \left[e^{-t_{min,i}/\tau} - e^{-t_{max,i}/\tau}\right]}. \tag{4.15}$$

The index i represents the number of top hats in the acceptance for the event. In principal a candidate could have several top hats though in practice only those with one are used in the analysis (rendering the sum obsolete) to make the calculation of the turning point functions rather more straightforward.

The detector decay time resolution is taken into account with a Gaussian function of width $\sigma = 50$fs

$$R(t, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t'-t)^2}{2\sigma^2}}. \tag{4.16}$$

This choice of model is motivated by the study in [91]. As noted in that study a

more complex resolution model could be used. The veracity of the chosen single Gaussian is tested and a systematic uncertainty assigned for it (see section 4.2.11).

The convolution of the Gaussian with the single exponential is

$$\begin{aligned}\left[\frac{1}{\tau}e^{-t/\tau}\Theta(t)\right] \otimes R(t,\sigma) &= \frac{1}{\tau}e^{-t/\tau}e^{\sigma^2/2\tau^2} \cdot \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{t}{\sigma}-\frac{\sigma}{\tau}} e^{-\frac{y^2}{2}}\mathrm{d}y \\ &= \frac{1}{\tau}e^{-t/\tau}e^{\sigma^2/2\tau^2}F\left(\frac{t}{\sigma}-\frac{\sigma}{\tau}\right). \end{aligned} \tag{4.17}$$

The function $F\left(\frac{t}{\sigma}-\frac{\sigma}{\tau}\right)$ is related to the error function by

$$F(x) = \frac{1}{2}\left[1 + erf\left(\frac{x}{\sqrt{2}}\right)\right], \tag{4.18}$$

where

$$erf(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-r^2}\mathrm{d}r \tag{4.19}$$

is the cumulative probability of a normal distribution between $\pm x$.

The signal lifetime PDF may also include a variation of the efficiency of the track reconstruction with decay time. The track finding algorithms prefer tracks that originate from the axis of the beam line. Therefore longer lived D^0 candidates that tend to decay further from the beam line will have daughters that are less likely to be reconstructed. The details of this efficiency are outlined in [91]. The key result is that this efficiency can be parametrised as a quadratic function of the particle's distance of closest approach to the beamline,the DOCAZ

$$\epsilon(\mathrm{DOCAZ}) = \alpha + \beta \times \mathrm{DOCAZ}^2. \tag{4.20}$$

Each track's DOCAZ can be expressed in terms of the particle decay lifetime via the linear dependence

$$\mathrm{DOCAZ} = \gamma + \delta t. \tag{4.21}$$

For each event there are two tracks with displaced vertices giving

$$\begin{aligned}\epsilon(t) &= (\alpha + \beta \cdot \mathrm{DOCAZ}_1^2)(\alpha + \beta \cdot \mathrm{DOCAZ}_2^2) \\ &= (\alpha + \beta(\gamma_1 + \delta_1 t)^2)(\alpha + \beta(\gamma_2 + \delta_2 t)^2). \end{aligned} \tag{4.22}$$

The efficiency is included in the per-event probability

$$f(t|A, sig, prompt) = \epsilon(t) \left(\frac{1}{\tau} e^{-t/\tau} \otimes R(t, \sigma) \right) \frac{1}{N}, \tag{4.23}$$

where N is a normalisation factor. The parameter β takes negative values so the efficiency of each track being reconstructed is parabolic with decay time; there are regions of time where the efficiency becomes negative according to the parametrisation. At these times the efficiency is assumed to be 0, so the limits of the integral in the normalisation is given by the roots of the efficiency for each track. Labelling each track r_1 and r_2 and the first and second roots as $r_{1,1}$ and $r_{1,2}$ the normalisation becomes

$$N = \int_{max(r_{1,1}, r_{2,2})}^{min(r_{1,2}, r_{2,2})} \epsilon(t) \left[\frac{1}{\tau} e^{-t/\tau} \otimes R(t, \sigma) \right] \Theta(t) \mathrm{d}t, \tag{4.24}$$

where again the Heaviside function describes the acceptance top hat function.

The parameterisation of DOCAZ (γ_i and δ_i) is worked out for each track so the efficiency correction is applied on a per-event basis. This is computationally intensive so is not applied as standard in the fit. The effect is tested and if non-negligible a systematic assigned to it.

Secondary

The secondary lifetime PDF does not have the resolution or efficiency corrections applied to it; it comprises a small part of the data (<10%) and is long-lived. The measured decay times of secondary candidates could naively be expected to be the convolution of two single exponential decays of lifetimes τ_1 and τ_2

$$f(t|A, cl, sec) = \left(e^{-t/\tau_1} \otimes e^{-t/\tau_2} \right) \Theta(t). \tag{4.25}$$

The convolution of two exponentials gives their difference

$$f(t|A, cl, sec) = \frac{\tau_1 \tau_2}{\tau_2 - \tau_1} \left(e^{-t/\tau_1} - e^{-t/\tau_2} \right) \Theta(t). \tag{4.26}$$

As the intermediate vertex is not reconstructed and the B and D flight directions are not parallel the lifetimes τ_1 and τ_2 are not the physical B and D lifetimes but effective lifetimes. Fig. 4.11 shows a fit to simulated data for secondaries.

It was found that the low lifetime region of secondaries is very hard to fit in the data. This region is dominated by prompt decays and in the $ln(\chi^2_{IP})$ variable the secondaries mostly lie under the prompt peak. For this reason one of the

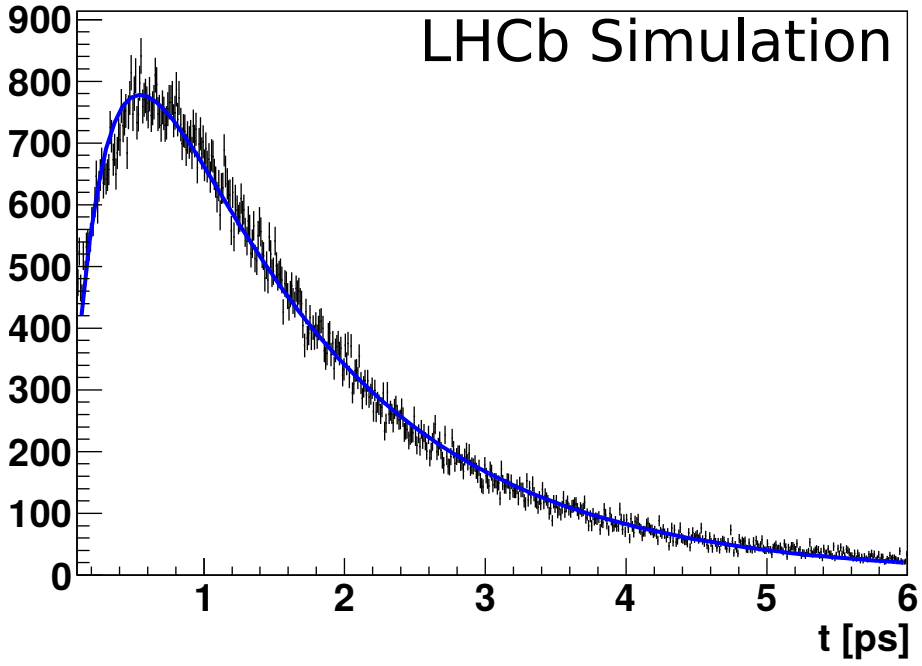


Figure 4.11: The decay-time distribution of secondary D^0 decays from simulated data. The result of the fit of equation 4.26 is overlaid. Reproduced from [114].

secondary lifetime parameters, τ_1 is fixed from a fit to MC data as in Fig. 4.11. The possible bias due to this fixing of a parameter is included in the systematic uncertainties.

An average acceptance function

In order to visualise the fitted PDF it is convenient to define a biased decay time function. This can be simply taken as the sum of the per-event functions over the data sample. Doing so one recovers the purely unbiased exponential decay $\frac{1}{\tau}e^{-t/\tau}$ multiplied by an average acceptance function $A(t)$,

$$A(t) = \sum_{candidates} \frac{1}{\int\limits_{acc} e^{-t/\tau}\mathrm{d}t} f(A), \tag{4.27}$$

where the integral is over the accepted times of that candidate. An example of the average acceptance is shown in Fig. 4.12. The turn on of the acceptance at low lifetimes due to the lifetime biasing cuts of the selection is clearly seen as well as the reduction of the acceptance at larger lifetimes due to the radial cut.

Cuts associated with the lifetime PDF

Some selections have been mentioned in the previous sections and they are summarised here.

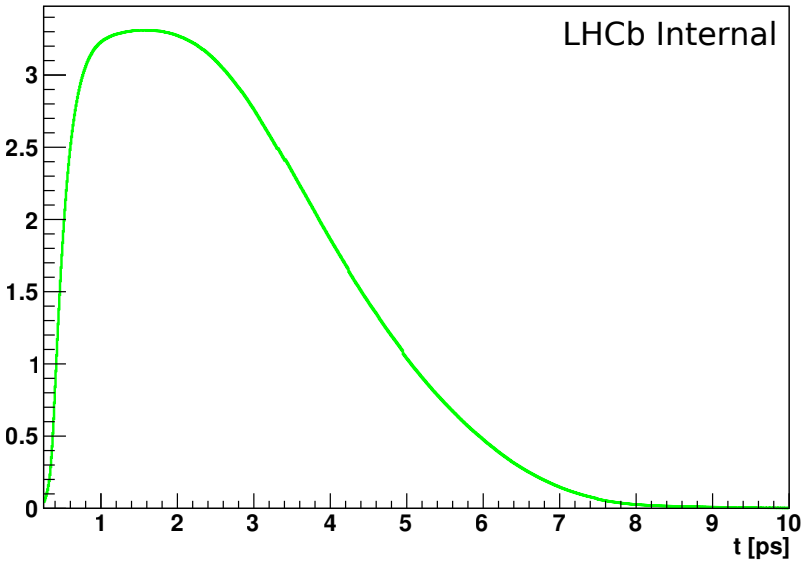


Figure 4.12: An example of the average acceptance function. Plotted is the generated average acceptance function for the prompt signal component of a fit to $K^-\pi^+$ data from the 2011 y_{CP} analysis. The fit was to magnet polarity down data taken during the run before the July technical stop.

There is a minimum decay time cut imposed to try and avoid regions where very few events are expected. The acceptance here should be effectively 0 so in the normalisation integral a small number results if the second turning point is also at low times and the acceptance function may become unreasonably large. The effect of this cut on the measurement is tested by altering the cut and a systematic uncertainty is assigned. Additionally, due to the resolution of the swimming and decay time measurements any events with their measured decay time outside of the two turning points are rejected.

Only candidates with two turning points are used as the extra complications from those with more greatly impact the time required to run the fit. As a test the number of turning points accepted is increased to four and the results compared with the standard fit to ensure no bias is introduced by the cut.

4.2.7 Non-analytic background PDFs

The shapes of the combinatoric and mis-reconstructed backgrounds in the t and $ln(\chi^2_{IP})$ variables are not known so they cannot be included in the fit by analytic PDFs. For this reason the functions in the fit are kernelised sPlots. One needs to find the unbiased time, conditional $ln(\chi^2_{IP})$ and acceptance function for these classes.

The turning point and conditional $ln(\chi^2_{IP})$ distributions ($f(TP_1)$, $f(TP_2 - TP_1)$ and $f(ln(\chi^2_{IP})|t, A, cl)$) can all be found using the techniques

already described. From the mass fit the probabilities and sWeights are calculated. Initially the data are assumed to be normally distributed and using equation 4.3 Gaussian kernels are applied to the probability distribution and a continuous function formed. Using this the first and second derivatives are calculated at each point to give the optimum bandwidth of the kernels for the data (equation 4.4). The kernels are then applied to the sPlot giving these three PDFs.

The biased decay time PDF is also found by this method, leaving only the unbiased PDF and average acceptance function to be found using an iterative approach. The biased function is

$$f(t, A|cl)_{bias} = f(t|cl)_{unbias} \times f(A|cl). \tag{4.28}$$

Initially an estimate is used for the unbiased time PDF; a single exponential decay with the lifetime parameter taken to be the mean of the sWeighted decay time distribution for that class. The acceptance function is given by

$$f(A|cl) = \sum_{candidates} sWeight(cl)_{cand} \times f(TP_1) \times f(TP_2 - TP_1) \times \frac{1}{\int\limits_{acc} f(t|cl)_{unbias} \mathrm{d}t}. \tag{4.29}$$

Using the estimate $f(A|cl)$ can be calculated, multiplied by the unbiased time distribution estimate and compared to the biased PDF of the kernelised sPlot. If the unbiased time estimate is the correct distribution the two should be the same. If they are not then the biased time sPlot is divided by the new acceptance to give a second estimate of the unbiased time PDF and the process is iterated. Eventually the product of the unbiased time PDF and acceptance function should match the sPlot at which point the method has converged. The unbiased time PDF can then be inserted into the fitter.

4.2.8 Unusual PDF definitions

For the most part the analysis used standard PDFs; combinations of Gaussian, exponentials etc. However some special PDFs were employed and they are listed here. The derivations of the first two are in [72].

The phase space PDF

For the Δm variable a PDF is required for the π_s^+ background. This is a distribution that has a cut off at the π mass and models the energy released in the

D^{*+} decay based on the two body Lorentz invariant phase space

$$LIPS(m_{D^{*+}}) = \frac{\sqrt{\left[m^2_{D^{*+}} - (m_{D^0} + m_\pi)^2\right]\left[m^2_{D^{*+}} - (m_{D^0} - m_\pi)^2\right]}}{m^2_{D^{*+}}}, \tag{4.30}$$

where in this instance $m_{D^{*+}}$ and m_{D^0} are the invariant masses of the D^{*+} and D^0 that have been reconstructed (not the PDG [117] average masses). Defining the energy released as $Q = m_{D^{*+}} - m_{D^0} - m_\pi$ and inserting into the equation, one arrives at the PDF

$$f(Q|\pi^+_s bkg.) = \begin{cases} Q^{1/2} + aQ^{3/2} & Q \geq 0 \\ 0 & Q < 0. \end{cases} \tag{4.31}$$

In the fit the π is assumed to be a true pion at the PDG average mass. The variable Δm is used and so Q is redefined for $Q = \Delta m - m_\pi$.

Smeared $ln(\chi^2_{IP})$ PDF

This PDF is used to model the prompt and secondary $ln(\chi^2_{IP})$ shapes; an example of the distribution of the prompt component is in Fig. 4.13. It allows for the translation of the secondary shape (as it evolves with decay time) and some smearing to take into account the wider shape of the secondaries. It is based upon taking the natural logarithm of a χ^2 distribution with two degrees of freedom and introducing some smearing into the resulting function:

$$P(y) = \begin{cases} \exp\{\alpha_L y - \exp\{\alpha_L(y-\mu)\}\} & y \leq \mu \\ \exp\{\alpha_L \mu + \alpha_R(y-\mu) - \exp\{\alpha_R(y-\mu)\}\} & y > \mu \end{cases}. \tag{4.32}$$

Crystal Ball PDF

The Crystal Ball PDF [118] is designed to model distributions of processes with energy loss. It is comprised of a Gaussian distribution with the lower tail becoming a power law distribution below a certain point. It is defined as

$$f(x) = \begin{cases} N \times e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \frac{x-\mu}{\sigma} \geq \alpha \\ A \times (B - \frac{(x-\mu)}{\sigma})^{-n} & \frac{x-\mu}{\sigma} < \alpha, \end{cases} \tag{4.33}$$

Where μ and σ are the mean and standard deviation of the Gaussian distribution respectively. The parameter α determines the number of deviations from the mean at which the distribution transitions between the Gaussian and power law

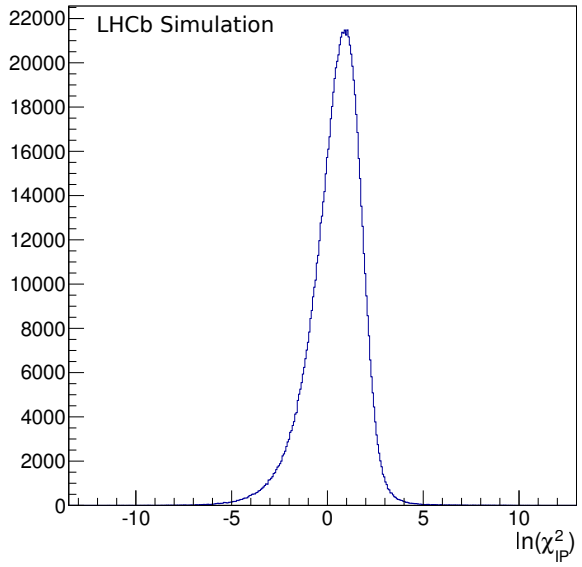


Figure 4.13: The distribution of $ln(\chi^2_{IP})$ for simulated prompt $D^0 \to K^-\pi^+$ decays.

components. The factors N, A and B ensure the total distribution is correctly normalised and that the function and its first derivative are continuous at the transition.

For these analyses the Crystal Ball is used to model the energy loss of the pion in $D^0 \to K^-\pi^+$ and $D^0 \to \pi^+\pi^-$ decays. As it is light the pion radiates significantly such that the measured momentum as shifted to lower values than the initial momentum from the decay. This produces a long tail in the reconstructed invariant mass distributions for these modes.

4.2.9 Blinding

The measurements were carried out in a blinded fashion in order to avoid biases due to the expectations of the analysts when setting up the fit. The prompt signal lifetime parameters τ for the K^+K^- and $\pi\pi$ fits were multiplied by a blinding factor at the interface between the minimiser and the fitting classes. Thus the true values are fitted to the data but the user will see the blinded lifetimes.

The $K^-\pi^+$ final state was not blinded as the comparison of the measured lifetime with the PDG was used to validate the fit. The two final states from which the physics results were obtained and the two initial states (D^0 or $\overline{D}^0$) were blinded separately. The blinding factor b for a given final state is the sum of a tagged and untagged part

$$b_{D^0} = b_{untagged} + b_{tagged}, \qquad b_{\overline{D}^0} = b_{untagged} - b_{tagged},$$

where $b_{untagged}$ and b_{tagged} are different for K^+K^- and $\pi\pi$. The A_Γ measurement

can therefore be unblinded without revealing the true lifetimes and unblinding y_{CP} by removing the b_{tagged} parameter.

4.2.10 Combining results

When quoting results for the whole sample the following method is used. For combinations of D^0 and $\overline{D}^0$ lifetimes the un-weighted mean of the individual numbers is taken. When combining the results from the splits by magnet polarity or technical stop the weighted mean is used, the weighting being the signal yield.

4.2.11 Checks and Systematics Overview

The veracity of the results of the fitter and the systematic uncertainties are evaluated both through tests on data and with pseudo-experiments. The latter involves generating several data sets, fitting them and comparing the results with the parameters used in the generation. They are henceforth referred to by the sobriquet *toy* experiments.

The most straightforward test of the fitter compares $K^-\pi^+$ fitted lifetime with the PDG [117] average of (410.1 ± 1.5) fs. The analyses are not absolute lifetime measurements and a lot of the biases cancel in the ratios to produce y_{CP} and A_Γ. However for the results to be believable they should agree with the world average within the uncertainties. Additionally one hopes that the quality of the fits would instil confidence that the results are accurate.

The fit configurations for each of the final states are tested using toys to check they do not introduce biases. Many (of the order of 330 though the precise number varies depending on individual job successes) data sets are generated using the default fitter configurations and then fitted; the results of the fits are compared with the generating parameters and checked for biases. The data sets are generated using the current world averages for the D^0 lifetime and y_{CP}. In the case of A_Γ the D^0 and $\overline{D}^0$ lifetimes are the same for the standard generation and to check the sensitivity to that observable a set it also generated with a non-zero A_Γ.

The sources of systematic uncertainties are the same for y_{CP} and A_Γ. They are summarised in table 4.2 which also lists how they are tested.

For the tests using fits to data (for example the effect of the minimum decay time cut) the assigned uncertainty is half of the full variation (plus and minus) of the quantity in question (y_{CP} or A_Γ). In the case of tests using toys the uncertainty is taken as the observed bias on the mean of the pulls resulting from the changes to the fitter. The larger of the absolute bias or the difference between the biases of the test and the standard configuration is used. The toy data sets have similar statistics to the $\pi^+\pi^-$ data sample, so to produce an absolute

Source	Test
Secondaries	vary fixed τ_1 parameter in time PDF toys with varied secondary fraction
Acceptance functions	Scale turning points Offset turning points
Decay-time resolution	Change resolution model Change resolution width
Mis-reconstructed bkg.	Vary lifetime of unbiased estimate PDF Change fractions around the nominal fit values
Combinatoric bkg.	Toys with different generating PDFs in m_{D^0} and Δm Toys with varied combinatoric time generation Toys with increased combinatoric fraction, $\pi^+\pi^-$ only
Minimum decay-time cut	Vary the cut
Reconstruction efficiency	Apply efficiency correction
No. of TPs	Include up to 4 TPs

Table 4.2: A summary of the systematic checks for y_{CP} and A_Γ. The precise variations are described in the text in sections 5.4.2 and 6.4.

uncertainty the bias is multiplied by the $\pi^+\pi^-$ statistical uncertainty (1.06×10^{-3}) to give an absolute uncertainty. In the circumstances where more than one test is carried out to measure a source of uncertainty the larger value is taken. For the toy tests the two final states have the same figure for the uncertainty assigned to them unless a specific test has been carried out for that final state (for example when checking the combinatoric background).

Backgrounds not from D^0

Combinatoric background is present in all three final states. As its time PDF is obtained by producing the sPlot extracted from the mass fit its systematic is investigated by examining the effect of using an incorrect model in the mass fit. To that end toys are generated with the combinatoric component using a different PDF in m_{D^0} and Δm to the standard fit. These are then fitted with the regular PDFs and the resulting bias taken. The time distribution of the generated combinatoric background is varied in a separate test and the bias again measured.

An additional test of the combinatoric background is performed for the $\pi^+\pi^-$ final state. There is a far greater amount of it in that mode so toys are generated with a larger combinatoric fraction to reflect that and compared with the standard configuration.

The mis-reconstructed backgrounds of the K^+K^- mode are examined with a combination of fits to data and toys. Fits are run with their fractions varied to be $\pm$the best fit values. The initial estimate unbiased time PDF that is used to extract the acceptance function is also varied.

Secondaries

An additional major source of background are candidates originating from B decays. These are particularly troublesome as their separation in the $ln(\chi^2_{IP})$ fit is tricky. The result of the fit to simulated data used to fix the τ_1 in the unbiased time PDF (see section 4.2.6) is checked by examining the effect of varying that parameter.

Additionally toys are generated with half as much and twice as much secondary as in data and the bias examined. A test is also performed with the generated data having the $ln(\chi^2_{IP})$ time evolution different to that found in the fit.

Turning points

The accuracy with which the turning points are determined by the swimming technique is tested using data through two methods.

In one instance the turning points all have an offset applied to them, referred to as introducing a turning point bias (TP bias). The values of the biases tested varied over the three analyses (see the individual sections for the details). This was due to the analysts gaining a better understanding of the TP resolution.

The second test is to apply a scale to the turning points by multiplying them by some constant factor. Again for reasons of a change in understanding the details changed between the measurements and are discussed in the relevant sections.

Decay time resolution

As noted in section 4.2.6 the exponential decay time PDF is convolved with a Gaussian resolution function. The estimated decay-time resolution of LHCb is taken to be 50 fs, motivated by the study in [91]. To test the effect of the decay time resolution model, this uncertainty parameter in the PDF is varied between 30 fs and 100 fs. Additionally it was shown that a more complex model that is the sum of three Gaussian functions may give a slightly more accurate description of the detector performance. the three widths are $\sigma_1 = 32$ fs, $\sigma_2 = 65$ fs and $\sigma_3 = 380$ fm. The respective fractions are $f_1 = 0.714$, $f_2 = 0.282$ and $f_3 = 0.004$. This more complex model is also trialled and compared to the standard model. The largest variation is taken as the systematic uncertainty.

Checks

As well as the numerical uncertainties described above some checks are carried out to ensure the stability of the results. In particular one hopes to find that the results from the individual data splits are consistent between the magnet polarities

and either side of technical stops. These are then split further into shorter run periods to see if there is any systematic trend over the course of the data taking.

The data is also split by some kinematic variables: D^0 p, p_T and η. Track multiplicity, the direction of the flight of the D^0 and L0 trigger decisions are also checked.

4.3 Summary

Both the A_Γ and y_{CP} analysis follow the methods presented in this chapter. Some details changed between the two results being obtained and they are detailed in the following chapters along with the results and detailed descriptions of the assessments of the systematic uncertainties.

Chapter 5

2011 A_Γ analysis

The measurement of A_Γ on the data collected by LHCb in 2011 was published in Physics Review Letters at the start of 2014 [119]. The Cabibbo favoured $K^-\pi^+$ final state is included in the analysis as it provides a check for the fitter by comparison of the measured lifetime with the world average and through an A_Γ pseudo-measurement with the expectation of a null result.

The details particular to this measurement, in particular the selection of the data and configuration of the fitter are expounded upon in sections section 5.1 and section 5.2. The blinded fit results are then presented in section 5.3, followed by the results of the systematics tests in 5.4. Finally the un-blinded measurement of A_Γ is given.

5.1 Selection

5.1.1 Trigger

The data is selected by two exclusive D to two-body hadron high level trigger lines: a signal and wide-mass line. The latter is pre-scaled by a factor of about 0.1 and covers a wider mass range than the signal line. The trigger cuts are summarised in table 5.1. The trigger configuration was altered slightly in the July 2011 technical stop (TS) where some cuts were tightened. The most pertinent cuts are those related to the distance of the D^0 candidate decay with respect to the primary vertex. In particular there is a minimum requirement on the impact parameter χ^2 with respect to the primary vertex of 2.0 before the TS, 9.0 after. The impact parameter (IP) is defined as the closest distance between the track direction and the interaction point; the χ^2 is the IP divided by the uncertainty on the impact parameter, where the uncertainty is a combination of the results of the track fit and the vertex fit. There is also a minimum requirement for the D^0 flight distance χ^2 (the χ^2 is defined analogously as for the IP χ^2) of 25.0 and 40.0 before

Particle	Quantity	Cut
Daughter	p_T	>800 MeV/c with one >1500 MeV/c
	p	>5.0 GeV/c
	Track χ^2/DoF	<3.0
	IP χ^2 w.r.t. PV	>2.0 (9.0)
	max DOCA	<0.1 mm
D^0	p_T	>2.0 GeV/c
	vertex χ^2/DoF	<10.0
	flight distance χ^2 w.r.t. PV	>25.0 (40.0)
	PV DIRA	<0.99985
	mass signal	$1815.0 < m_{D^0} < 1915.0$ MeV/c^2
	mass wide mass	$1715.0 < m_{D^0} < 2015.0$ MeV/c^2
Global	No. of tracks	<180
	Hlt1TrackAllL0 TOS	True

Table 5.1: The trigger cuts for the 2011 data taking. Those numbers in parentheses are the cuts after the July technical stop.

and after the technical stop.

In addition there are several cuts relating to the quality of the reconstruction, both for the fit of the tracks of the D^0 candidate decay products and the vertex fits. Finally the ranges of the signal and wide mass trigger ranges are defined as 1815–1915 MeV/c^2 for the former and 1715–2015 MeV/c^2 for the latter.

The output of the trigger gives a data set that is as close to being ready for analysis as possible. In essence the selection for the whole analysis has been performed in the HLT, with the exception of the particle identification requirements that are not available in the trigger.

5.1.2 Stripping

For the most part the cuts applied in the offline processing reflect those in the trigger, although the offline reconstruction is more complete than that in the trigger reflecting the greater time available to it. The differences in selection relate to the extra information available offline for the π_s^+ and the particle identification PID. In the stripping selection a π_s^+ candidate is combined with a D^0 candidate to form the D^{*+}. Requirements on the quality of the π_s^+ track and the resulting vertex fit are imposed for the D^0 and π_s^+ to form the D^{*+}. The difference in D^{*+} and D_s^+ mass, Δm, is required to be in the range Δm<160.0 MeV/c^2.

The particle identification is also applied at this stage. Any tracks that have associated hits in the muon system are vetoed. Additionally selection cuts are applied to the hadron PID likelihoods to separate pions and kaons. These cuts are initially fairly loose in order to retain the maximum number of signal candidates.

Particle	Quantity	Cut
Daughter	$\pi^{\pm}$ DLL(K-π)	$<$ -5
	$K^{\pm}$ DLL(K-π)	$>$5
D^0	Radial decay distance	$< 4\,\mathrm{mm}$
	Decay time	$>0.25\,\mathrm{ps}$
	m_{D^0} wide-mass	1815–1820 MeV/c^2 , 1910–1915 MeV/c^2
	m_{D^0} signal	1820–1910 MeV/c^2
D^{*+} and π_s^+	Δm	140–152 MeV/c^2
Global	No. of TPs	2
	t outside swimming acc.	False
	No.of candidates in event	1

Table 5.2: A summary of the cuts made after the stripping.

5.1.3 Futher cuts

After the stripping some additional tightening of the cuts was enacted, summarised in table 5.2. A cut on the number of candidates in the event was made such that only one was chosen from events with multiple reconstructed candidates. The chosen candidate from the was selected at random and the effect of this was checked to ensure no biases were introduced.

The daughter PID cuts were tightened to improve the pion-kaon separation and reduce the possible backgrounds. The fit range in D^0 mass and Δm were defined. Finally candidates with un-physical acceptance functions (the measured decay time is outside the acceptance from the swimming due to the finite resolution) were rejected as were those events with more than two turning points (more than a single top hat acceptance for the event) in the acceptance (see section 4.2.5).

5.1.4 Data splits

The data set was split into eight sub-samples for fitting. The division was by magnet polarity (up or down), D^0 flavour and by data-taking period before and after the technical stop in July. The sample sizes that were fitted (the number of candidates after all selections) are in tables 5.3 ($\pi^+\pi^-$), 5.4 (K^+K^-) and 5.5 ($K^-\pi^+$).

5.1.5 Decay Tree Fitter

In this analysis the Decay Tree Fitter (DTF) algorithm [120] was used in the reconstruction of the D^{*+} candidate. This algorithm is designed to reconstruct decays using a more holistic view than the traditional method of combining final

	Magnet polarity	D^0	$\overline{D}^0$
Pre TS	Up	135144	137979
	Down	164764	165041
Post TS	Up	195874	200543
	Down	318083	316866

Table 5.3: The number of fitted $\pi^+\pi^-$ candidates in each sub-sample.

	Magnet polarity	D^0	$\overline{D}^0$
Pre TS	Up	379227	391689
	Down	461698	461139
Post TS	Up	558690	575832
	Down	963432	965379

Table 5.4: The number of fitted K^+K^- candidates in each sub-sample.

state particles to make candidates of decaying particles. Instead the whole decay chain is fitted together. In the process one can impose constraints on the decay vertices, the directions and the invariant masses of the particles. Information about the reconstructed particles with and without the fit was kept.

The D^0 candidates were made from combining two kaons without applying any constraints. The DTF was not used for the D^0 to avoid any possible effects the vertex constraint may have on the measured decay time. A π_s^+ was added which was constrained in the fits so its direction of flight points back to the primary vertex. The 4-vectors of the D^0 (without DTF) and π_s^+ (after applying the DTF constraints) were then combined to form a new D^{*+} candidate.

The effect of this algorithm is to improve the Δm resolution. The π_s^+ is a low momentum track and consequently has poor relative momentum and direction resolution; it is this particle that contributes most to the Δm uncertainty. The improvement in Δm resolution can be seen in Fig. 5.1. The stripping cut on Δm at $160\,\mathrm{MeV}/c^2$ leads to the distribution with the constraint decreasing at higher values; this motivates the tighter Δm cut listed in table 5.2.

Using DTF in this manner introduces some time dependency in the Δm variable. This is due to the effect on background events originating from B decays,

	Magnet polarity	D^0	$\overline{D}^0$
Pre TS	Up	2829271	2964950
	Down	2900389	2966165
Post TS	Up	4205761	4426319
	Down	6850383	6990802

Table 5.5: The number of fitted $K^-\pi^+$ candidates in each sub-sample.

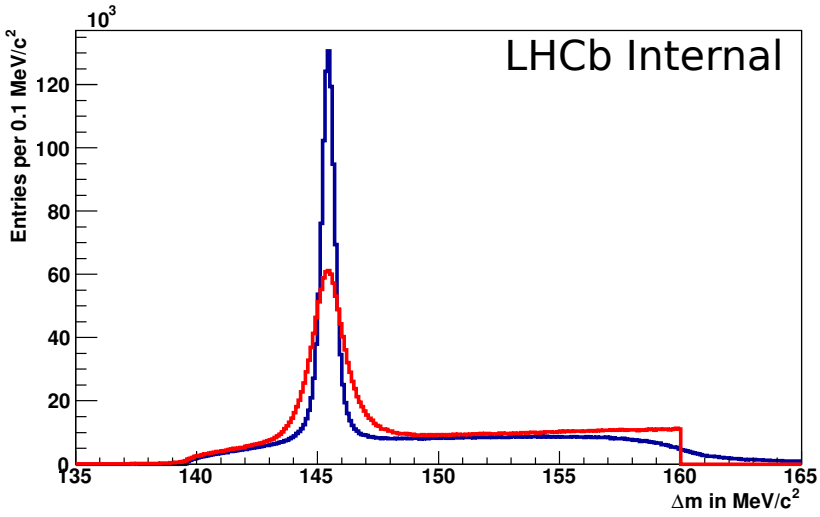


Figure 5.1: Comparison of Δm with (blue) and without (red) a constraint on the π_s^+ vertex using DTF. From [72].

known as the secondary component. For those secondary events that are correctly tagged the attached π_s^+ has its momentum vector altered significantly to point back to the primary vertex in the reconstruction. This correction is time dependent; for apparently longer lived D^0 candidates the correction to the π_s^+ momentum is greater than for those that are shorter lived. This manifests itself in a time dependent Δm distribution for the secondary component. In the fit Δm and decay time are considered completely independent. This assumption could potentially bias the fit. For this reason the tests with simplified simulations that are used to test the fitter are generated with a time-dependent Δm distribution for the secondary component and then fitted with the standard configuration. To estimate the time dependence a sample of $B \to D$ decays that had previously been selected and reconstructed are taken. These are then reconstructed again without the intermediate B and assuming that the D has originated at the collision point. These events are binned in decay time and fitter in Δm with a Gaussian function. Fig. 5.2 shows how the mean and width of the Δm distribution varies with time for the K^+K^- final state. The overlaid red lines are linear fits that describe how the variables evolve and which are used when generating the toys.

5.2 Fitter configuration

5.2.1 Mis-reconstructed backgrounds

A study was carried out to examine the possibility of mis-reconstructed events appearing as backgrounds in the selected data. These backgrounds are from decay

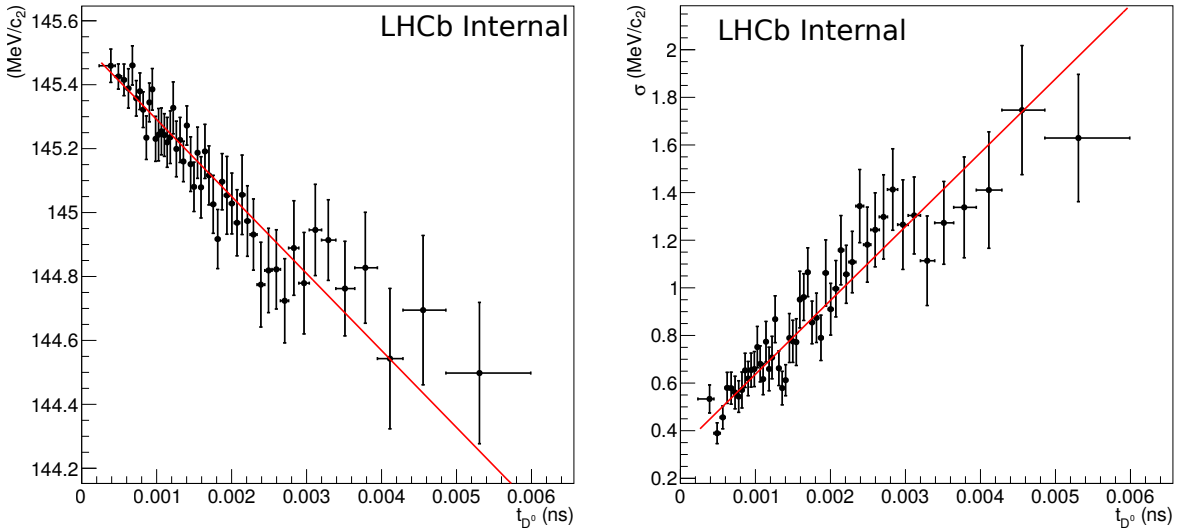


Figure 5.2: The time evolution of secondary decays in the Δm variable. The data has been binned in proper time and the Δm variable of each bin fitted with a Gaussian function. On the left is plotted the mean of the Gaussian, on the right the standard deviation. The overlaid red lines are linear fits. The final state is K^+K^-. From [72].

Decay	Branching fraction
$D^0 \to K^-\pi^+$	$(3.89 \pm 0.05)\%$
$D^0 \to K^+K^-$	$(3.93 \pm 0.08) \times 10^{-3}$
$D^0 \to \pi^+\pi^-$	$(1.40 \pm 0.03) \times 10^{-3}$

Table 5.6: The branching fractions of the three signal modes, from PDG [117].

modes, other than the one of interest, that have been reconstructed as the signal. Such candidates may affect the shape of the mass and Δm distributions and affect the quality of those fits. They may also extend under the signal peak in mass and Δm and be mistaken for signal, biasing the eventual lifetime measurement. They therefore must be considered in the fit.

Potential background sources were examined that have a branching ratio of comparable size or larger than the signal modes and are listed in table 5.7. They also must be reconstructed to lie in the D^0 mass range of interest, either due to some mis-ID changing the reconstructed mass or due to a particle not being reconstructed. The most pernicious of these backgrounds may peak in either D^0 mass or Δm in which case they would be mistaken for signal if not taken into account. The modes considered are listed in table 5.7 along with how they may be mistaken for signal.

Simulated data samples were obtained for each of these as well as for the signal decays. Using the MC truth information only true signal for each decay was

Signal mode	Background mode	Branching fraction %	Mis-reconstruction
$K^-\pi^+$	$D^0 \to \pi^+\pi^-\pi^0$	1.44	$\pi^- \to K^-$, π^0 lost
	$D^0 \to K^-\mu^+\nu_\mu$	3.3	$\mu^+ \to K^+$
	$D^+ \to \pi^+\pi^+\pi^-$	0.3	$\pi^- \to K^-$
	$D_s^+ \to K^+K^-\pi^+$	5.5	π^+ lost
K^+K^-	$D^0 \to K^-\pi^+\pi^0$	13.9	$\pi^+ \to K^+$, π^0 lost
	$D^0 \to K^-\mu^+\nu_\mu$	3.3	$\mu^+ \to K^+$
	$D^+K^-\pi^+\pi^+$	9.4	$\pi^+ \to K^-$, π^+ lost
	$D_s^+ \to K^+K^-\pi^+$	5.5	π^+ lost
$\pi^+\pi^-$	$D^0 \to \pi^-e^+\nu_e$	0.28	$e^+ \to \pi^+$

Table 5.7: Possible mis-reconstructed backgrounds, their branching fractions and the mis-reconstruction.

taken. The stripping selections were then applied to each assuming the signal hypothesis and the expected yield of the backgrounds calculated. The results for the K^+K^- final state are in table 5.8 and $K^-\pi^+$ in table 5.9.

Channel	Events in MC	Candidates selected by stripping	Expected fraction w.r.t. signal
$D^0 \to K^+K^-$	10060944	41441	–
$D^0 \to K^-\pi^+\pi^0$	2027984	8	3
$D^0 \to K^-\mu^+\nu_\mu$	6380471	9	0.3
$D^+ \to K^-\pi^+\pi^+$	10150948	32	3
$D_s^+ \to K^+K^-\pi^+$	4014973	44	2

Table 5.8: The results of the Monte Carlo background study for the K^+K^- final state. The results of the stripping selections are compared between each background and the signal and combined with the relative fractions to give the estimate of the amount of background expected with respect to signal.

In the $K^-\pi^+$ mode only the $D^0 \to \pi^+\pi^-\pi^0$ decay was found to be non-negligible. In K^+K^- the four listed backgrounds were thought to be significant. For these the MC was fitted in m_{D^0} and Δm assuming the signal hypothesis. Some of the stripping cuts had to be removed to increase the sample sizes for the purposes of fitting, in particular those of the PID. Additionally the mass windows were widened for the fits in order to obtain reasonable statistics. As examples the fits for the $K^-\pi^+\pi^0$ component can be see in Fig. 5.3 and for the D_s^+ in Fig. 5.4. These shapes were then fixed and included in the fitter as new classes whose fractions were free to float.

Trial fits determined that the $K^-\pi^+$ fit was not sensitive to the mis-reconstructed background. The fraction fitted to be negligible and so it was left out of the final fit configuration. Similarly in K^+K^- the $D^+ \to K^-\pi^+\pi^+$ and $D^0 \to K^-\mu^+\nu_\mu$ were found to be too similar to the other backgrounds shapes and the fraction fitted zero. In the case of the former it was very similar to the

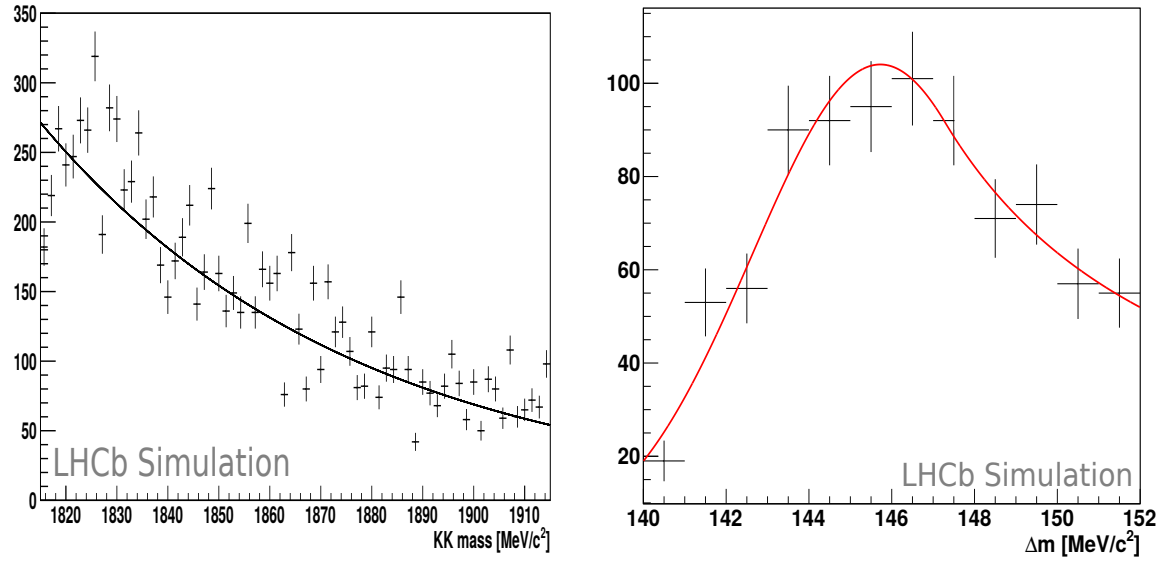


Figure 5.3: Simulation of $D^0 \to K^-\pi^+\pi^0$ reconstructed under the K^+K^- hypothesis. Shown are the m_{D^0} (left) and Δm (right) variables with fit results overlaid.

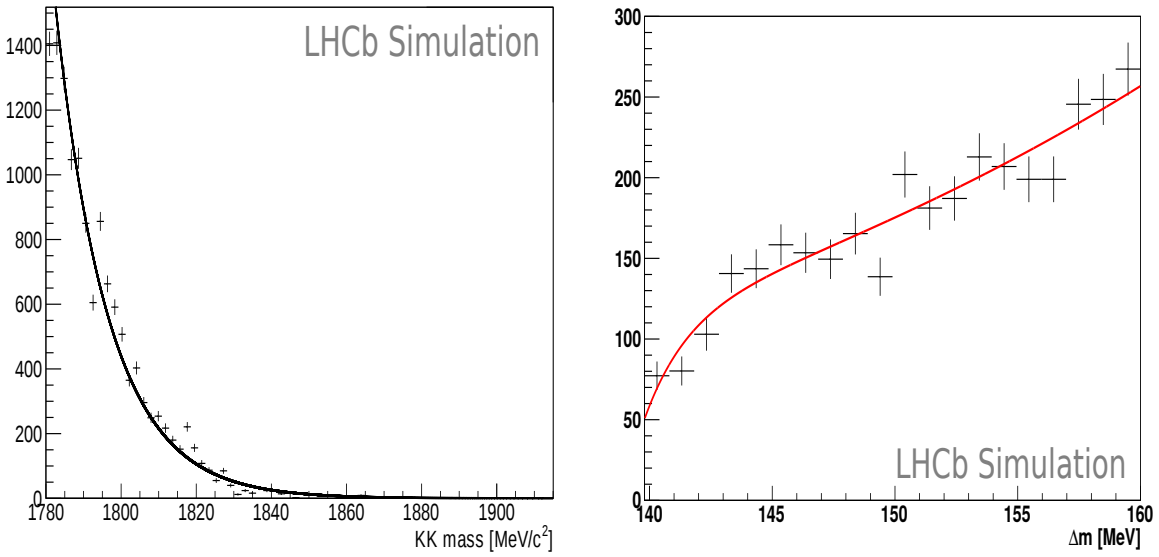


Figure 5.4: Simulation of $D_s^+ \to K^+K^-\pi^+$ reconstructed under the K^+K^- hypothesis. Shown are the m_{D^0} (left) and Δm (right) variables with fit results overlaid.

Channel	Events in MC	Candidates selected by stripping	Expected fraction w.r.t. signal
$D^0 \to K^-\pi^+$	10141454	303147	–
$D^0 \to K^-\mu^+\nu_\mu$	6380471	0	$< 4 \times 10^{-3}$
$D^0 \to \pi^+\pi^-\pi^0$	–	–	0.03
$D^+ \to \pi^+\pi^+\pi^-$	2038992	57	7×10^{-3}
$D_s^+ \to \pi^+\pi^+\pi^-$	2053988	0	$< 4 \times 10^{-3}$

Table 5.9: The results of the Monte Carlo background study for the $K^-\pi^+$ final state. The results of the stripping selections are compared between each background and the signal and combined with the relative fractions to give the estimate of the amount of background expected with respect to signal. In the case of $D^0 \to \pi^+\pi^-\pi^0$ the result of $D^+ \to \pi^+\pi^+\pi^-$ has been combined with the relative branching fractions of that decay for the estimate due to the lack of suitable Monte Carlo data.

combinatoric background in both fitted variables. The latter distribution had a strong resemblance to the larger $K^-\pi^+\pi^0$ mode. In reality it does not matter if these specific decays are not included explicitly as long as they are accounted for in some kind of general background model which is propagated to the time fit via the kernelised sPlots.

The determination of the mis-reconstructed backgrounds, in particular by taking the fit shape from MC with limited statistics, could provide a source of systematic uncertainty. The effect of this was evaluated and is described in section 5.4.

5.2.2 Mass fit

The mass signal peak was described by the sum of two Gaussian functions and a Crystal Ball function (see section 4.2.8) with common means but independent widths. Together these describe a mass peak with some per-event measurement errors that are not known but can be described by the combination of Gaussian functions. The Crystal Ball function models the low tail of the mass peak due to the radiation of the final state pions. In K^+K^- this radiative tail is negligible so the Crystal Ball turnover parameter, α was fixed to some arbitrarily large value as was the power law of the tail, N.

In $K^-\pi^+$ and $\pi^+\pi^-$ the radiative tail is significant. For $\pi^+\pi^-$ both the turn over and power were fixed to the values obtained from a fit to simulation as the data set was not large enough to constrain them. In the large $K^-\pi^+$ data set the power was similarly fixed but the α parameter was allowed to float. Again due to the limited statistics, the contribution of the third Gaussian in the $\pi^+\pi^-$ was fixed as was its width, this time to the result of the $K^-\pi^+$ fit.

The π_s^+ background PDF was the same as the signal PDF in m_{D^0} as it

comprised of true D^0 candidates. The combinatoric was a floated exponential function whilst the mis-reconstructed background shapes were fixed to the results of fits to simulation as described in section 5.2.1. The fractions describing the contributions of each background was left to float.

In Δm the correctly tagged signal was described by three independent Gaussian functions for K^+K^- and $K^-\pi^+$. As with the mass projection these functions model the peak and a resolution function. Additionally as the decay tree fitter affects the Δm of the secondary it was important that these functions had freedom to take that into account. In $\pi^+\pi^-$ however the three means were forced to be the same due to the limited size of the data set. The combinatoric and π_s^+ backgrounds were described by the phase space PDF as described in Eq. 4.31. As in the mass variable the mis-reconstructed background shapes were fixed to the results of the fits to simulated data.

These PDFs are summarised in section 5.2.4.

5.2.3 Time fit

A substantial challenge for this analysis was obtaining a good two-dimensional fit to decay time and $ln(\chi^2_{IP})$. In particular the time dependence of the secondary component was open to question. A study with simulated data was performed in order to estimate the evolution of the shape of $ln(\chi^2_{IP})$ with proper time for the secondary component.

A sample of unbiased (no selections were applied) $D^0 \to K^+K^-$ MC simulation was taken. From these only those that had come from a B decay were selected to form a sample of secondary candidates. This sample was then broken into slices of time and the $ln(\chi^2_{IP})$ variable for each slice fitted with the secondary PDF, a smeared $ln(\chi^2_{IP})$ PDF from equation 4.32 (see section 4.2.8). The fit parameters at each time point were plotted and fitted with suitable functions. The results for the mean and right-hand smearing parameters (μ and α_R respectively) are shown in Fig. 5.5, where the fits have been performed with the left-hand smearing parameter (α_L) fixed to 1. The time evolutions were both fitted with functions of the form

$$p(t) = A + B \cdot t + C \cdot \arctan(D \cdot t). \tag{5.1}$$

The same study was done with data using a sample of reconstructed $B \to D^0X$ decays (see [72] for the complete list of modes used) with the D^0 decaying to either K^+K^-, $\pi^+\pi^-$ or $K^-\pi^+$. Similar results were obtained as those found with simulation. The results of study with data were used when examining alternative time evolutions for the systematics assessment.

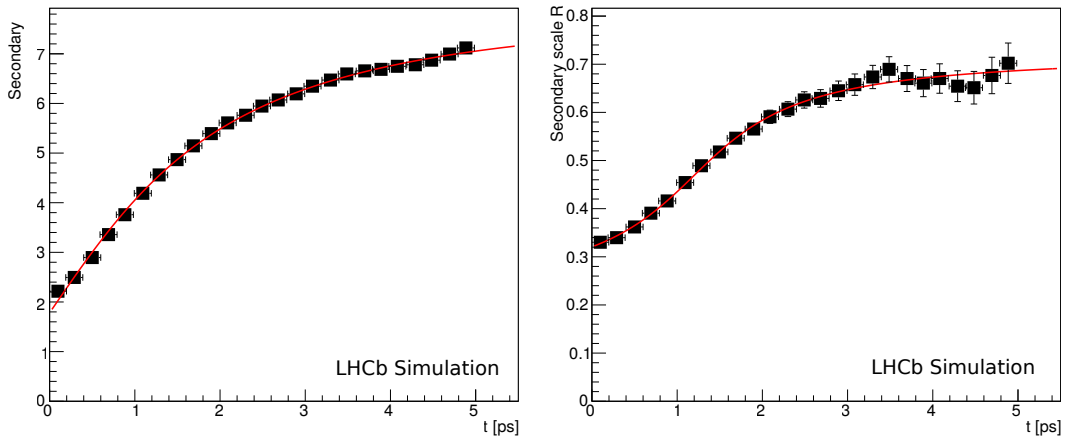


Figure 5.5: Evolution of the secondary fit parameters. Simulated data of the secondary background has been binned in decay time and fitted with the smeared $ln(\chi^2_{IP})$ PDF in equation 4.32. Plotted are the μ (left) and α_R parameters of the fits in each bin of decay time. Overlaid in red are fits of the function in Eq. 5.1.

In the analysis fit it was found that the time evolution of the secondary α_R parameter fitted to resemble a linear function. This was adopted as the standard configuration. The time dependence of α_L was also described by a linear function to allow for some variation.

Both signal and π^+_s background prompt $ln(\chi^2_{IP})$ components were made up of two of the smeared $ln(\chi^2_{IP})$ PDFs (equation 4.32) with a common α_L parameter, the signal and π^+_s background PDFs sharing the same parameters and the same time evolutions of those parameters. The combination of the two PDFs described the natural logarithm of a χ^2 distribution and allows for some distortion due to instrumental effects, such as the errors in the χ^2 calculation not necessarily being accurate.

The secondary components were modelled with a single smeared $ln(\chi^2_{IP})$ function, again with shared parameters. This component is small enough that a single PDF was sufficient to describe the data. As in the mass fit some simplifications were made for the $\pi^+\pi^-$ final state: the time evolution of the α_L parameter was fixed to the fit from the study with simulated data and the evolution of the mean parameter neglected the linear term.

The decay time PDFs were as described in section 4.2.6. The signal and π^+_s background prompt PDFs had independent decay times. The secondary signal and π^+_s PDFs however had common lifetime parameters to simplify the fit configuration. The differences between them were expected to be small and the secondary components comprise a small portion of the data. In K^+K^- and $K^-\pi^+$ the secondary lifetime PDF was the sum of two of the difference of exponential PDFs from Eq. 4.26 with one common fixed lifetime parameter and two floated.

For $\pi^+\pi^-$ just one difference of exponentials PDF was used again with a fixed parameter. For each component the turning point distributions were calculated separately giving different acceptance functions for each component of the fit.

Numerical PDFs

As noted in section 4.2.7 the PDFs for the turning points and backgrounds are numerical PDFs. It was initially thought that the PDFs produced by each of the eight samples alone would suffer from a lack of statistics in some of the smaller background classes. For this reason it was decided to apply the methods to the whole data set. The individual mass fits were performed first. The weighted averages of the fitted parameters were taken to produce an estimate PDF for the whole data sample. The total mass PDF was then constructed using these averaged numbers and the techniques of sections 4.2.5 and 4.2.7 were applied using the whole data set with this average.

5.2.4 Fit PDFs summary

The fit configuration has been outlined in the previous section. Here the model for each variable is summarised.

- m_{D^0}:
  - Correctly tagged D^0: The sum of two Gaussian functions and a Crystal Ball with common mean. The widths of the second and third are parametrised as ratios compared to the first. In the case of K^+K^- the transition to the power law is fixed to a large value as it should not contribute significantly to the mass distribution and the power itself is fixed to an arbitrary value. Therefore the $\pi^+\pi^-$ PDFs have **8** free parameters and the K^+K^- **6**.
  - π_s^+ background: The same PDF as correctly tagged.
  - Combinatoric background: A single exponential with floating decay constant. **1** free parameter.
  - Mis-reconstructed backgrounds: The shapes are fixed as in section 5.2.1; the fraction of their contributions are left to float.
- Δm:
  - Correctly tagged D^0: Three Gaussian functions with independent means. The second and third widths are parametrised as scales with respect to the first width. There are therefore **8** fitted parameters: three

means, three for the widths and two for the relative fractions of the Gaussian functions.

- π_s^+ background: Phase space PDF of Eq. 4.31. The factor for the power $\frac{1}{2}$ was fixed to 1. Higher orders were found to fit to 0 so were not included.
- Combinatoric background: Same PDF as the π_s^+ background with the same fit parameters.
- Mis-reconstructed background: Fixed shapes from MC.

- $ln(\chi^2_{IP})$ against t:
  - Prompt signal: The sum of two smeared $ln(\chi^2_{IP})$ PDFs of Eq. 4.32. The two have common mean and α_L parameters. The right hand side smearings, α_R, are independent. There are **5** fitted parameters (μ,α_L,α_{R1},α_{R2} and the relative fraction of the functions). All the parameters are given a linear time evolution, with the two α_R parameters having common evolutions

$$p(t) = A + B \times t. \tag{5.2}$$

    The time evolutions therefore introduce a further **3** parameters to be fitted.
  - Secondary signal: A single smeared $ln(\chi^2_{IP})$ PDF, giving **3** parameters. The mean μ has the time dependence

$$p(t) = A + B \cdot t + C \cdot \arctan(D \cdot t). \tag{5.3}$$

    Both the left and right hand side smear parameters, α_L and α_R followed a linear function

$$p(t) = A + B \times t. \tag{5.4}$$

    Therefore the time evolutions of the parameters introduced **5** parameters.
  - Prompt π_s^+ background: The same PDF as for the prompt signal with common parameters.
  - Secondary π_s^+ backgrounds: The same PDF as for the secondary signal with common parameters.
  - Combinatoric backgrounds: Kernelised sPlots.

  - Mis-reconstructed backgrounds: Kernelised sPlots.

- Decay time:
  - Prompt signal: A single exponential convolved with a Gaussian resolution function (Eq. 4.17). **1** free parameter.
  - Secondary signal: The sum of two secondary lifetime PDFs (the convolution of two single exponentials) as in Eq. 4.26. The two share a common τ_1 parameter of 0.27 ps, fixed from a fit to simulation. There are **3** parameters to be fitted: the two lifetime parameters and the relative fraction between them.
  - Prompt π_s^+ background: Another single exponential convolved with a Gaussian function. **1** fitted parameter.
  - Secondary π_s^+ background: The same PDF as for secondary signal with the same lifetime parameters.
  - Combinatoric background: Kernelised sPlots.
  - Mis-reconstructed backgrounds: Kernelised sPlots.

The fractions of the classes were allowed to float, giving **2** parameters for $\pi^+\pi^-$ and **4** for K^+K^-. Furthermore the fractions of the prompt sub-classes within the signal and π_s^+ background classes floated introducing **2** additional parameters. In total therefore there are **19** parameters in the mass - Δm fit, **23** in the decay time -$ln(\chi^2_{IP})$ fit, resulting in **42** fitted parameters in total for both the K^+K^- and $\pi^+\pi^-$ final states.

5.3 Fit results

The K^+K^- and $\pi^+\pi^-$ mass and Δm fit results are shown in Fig. 5.6 for an example data sample chosen at random. The fitted distributions for $ln(\chi^2_{IP})$ and decay time are in Fig. 5.7 and the $TP1$ and TP_{diff} plots are given in Fig. 5.8. Underneath the distributions pull plots are shown. For each plotted bin the pull is defined as

$$\frac{data - fit}{uncertainty}$$

where the uncertainty is the statistical uncertainty of the plotted bin. An example of the two-dimensional pull distribution for the $ln(\chi^2_{IP})$ and t fit is shown in Fig. 5.9 for one of the $\pi^+\pi^-$ fits.

As can be seen there are regions in the time fit where the fit model has not adequately described the data. In $ln(\chi^2_{IP})$ this was found to be mostly in the

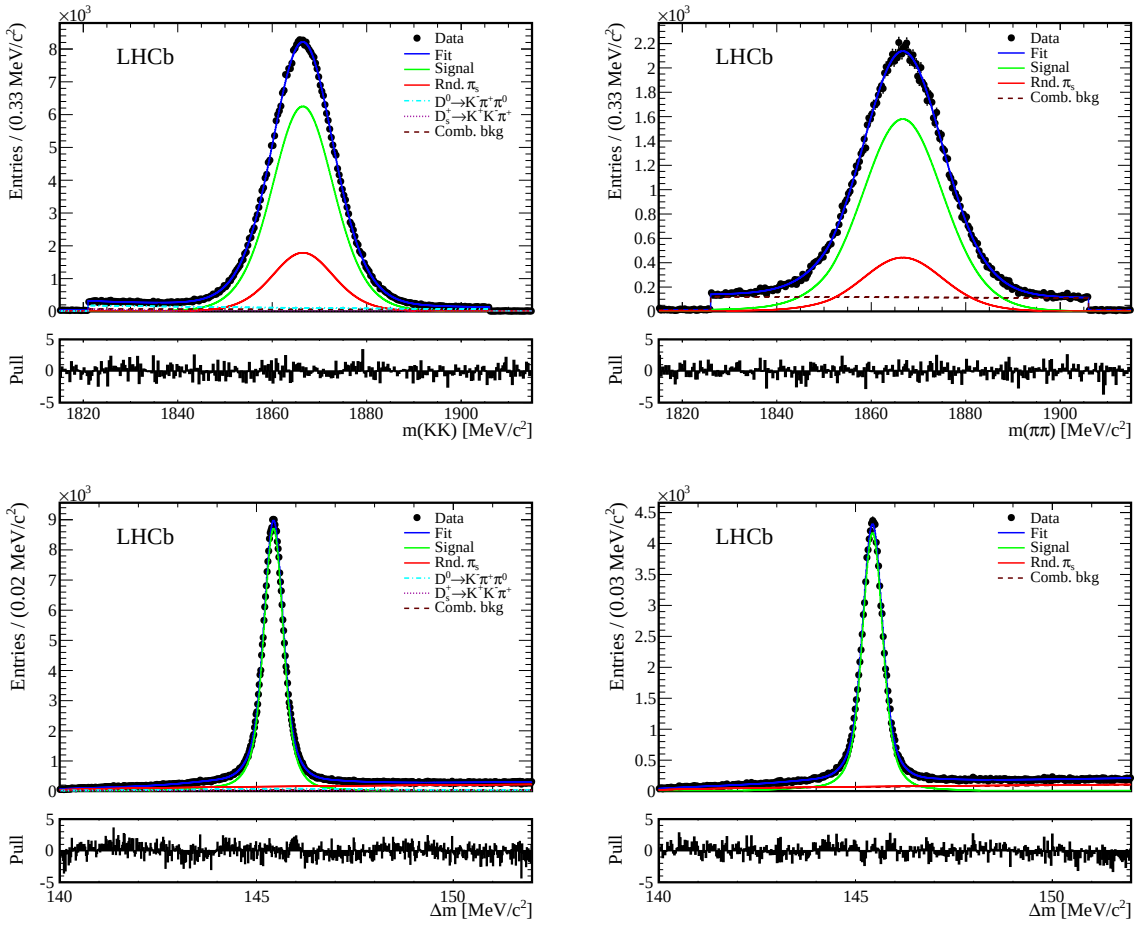


Figure 5.6: The m_{D^0} (top) and Δm (bottom) distributions for the K^+K^- (left) and $\pi^+\pi^-$ (right) final states with fitted components overlaid. The data sample shown here is D^0 candidates, with magnet polarity down for the run period before the July technical stop.

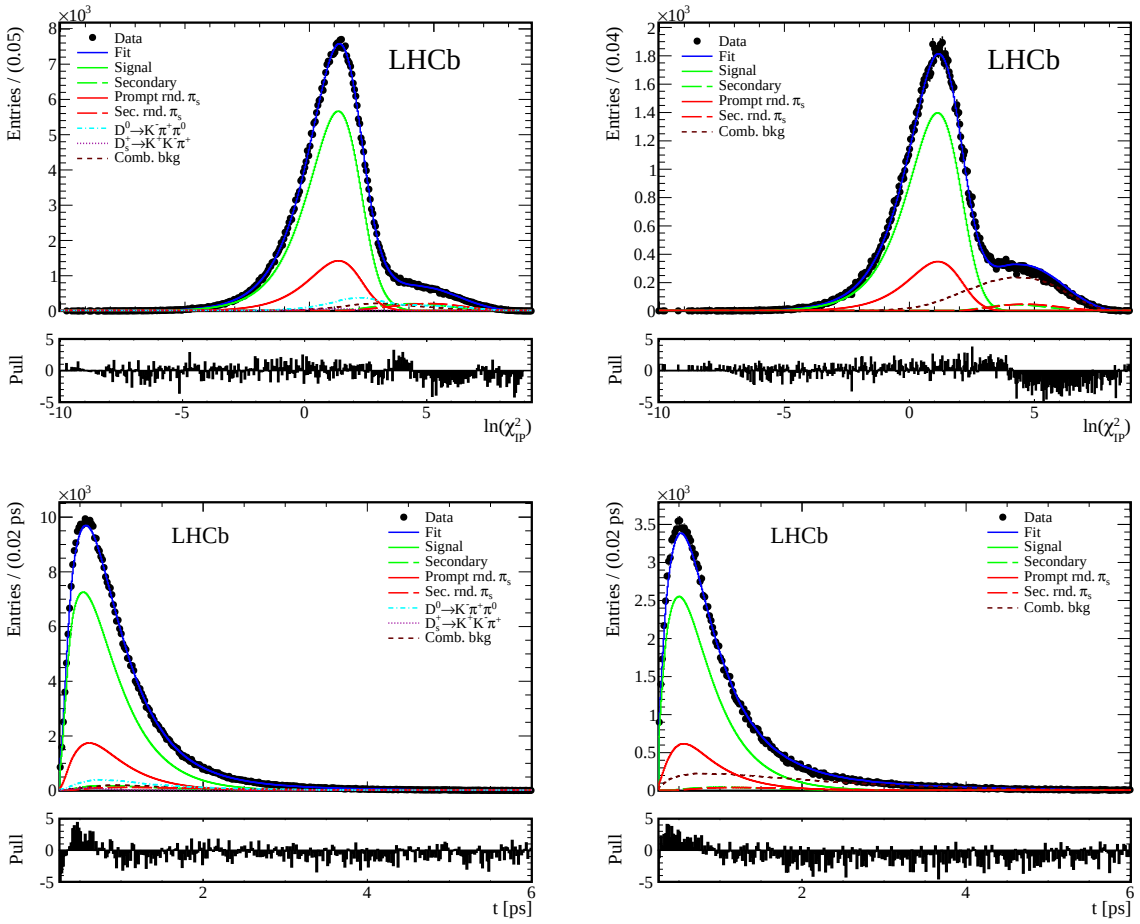


Figure 5.7: The $ln(\chi^2_{IP})$ (top) and decay time (bottom) distributions for the K^+K^- (left) and $\pi^+\pi^-$ (right) final states with fitted components overlaid. The data sample is D^0 candidates, with magnet polarity down for the run period before the July technical stop.

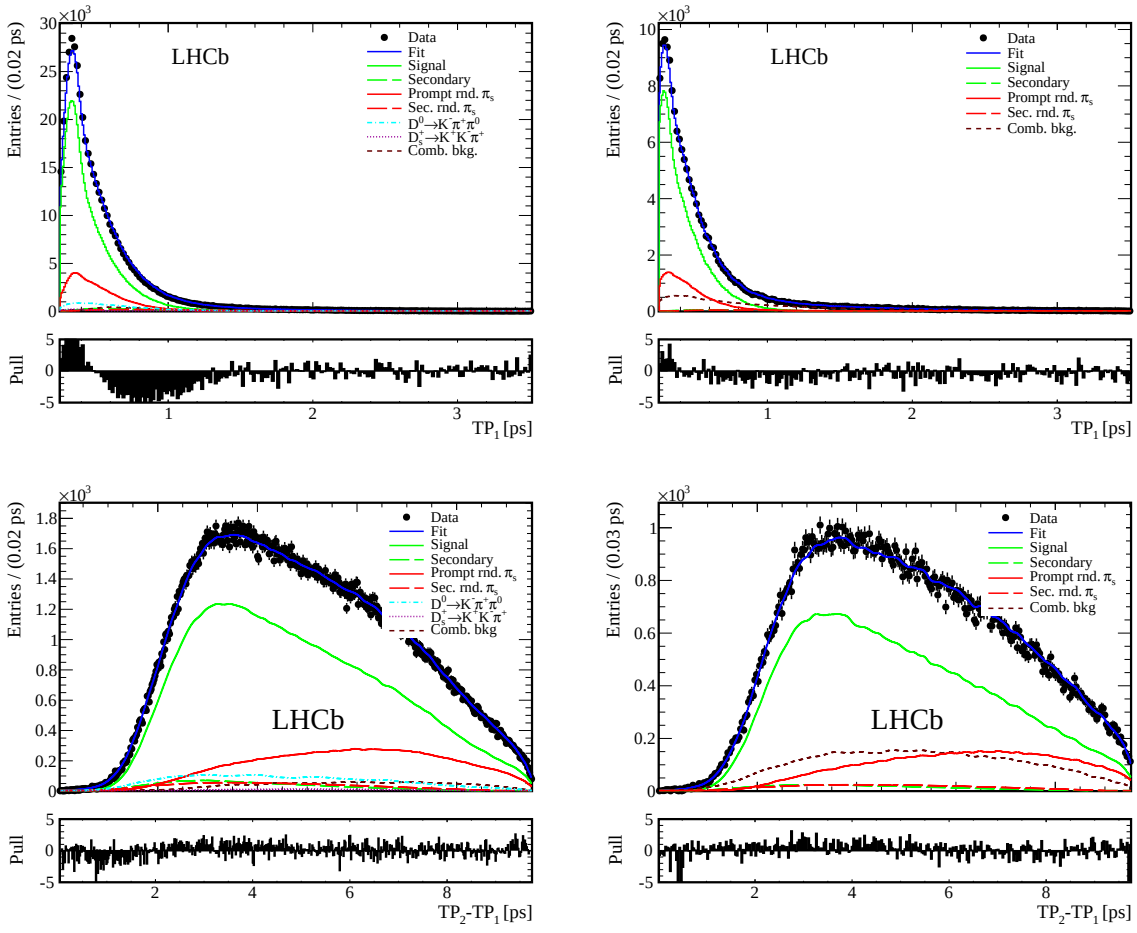


Figure 5.8: The TP_1 (top) and TP_{diff} (bottom) distributions for the K^+K^- (left) and $\pi^+\pi^-$ (right) final states. Overlaid are the generated functions used in the fits. The data sample is D^0 candidates, with magnet polarity down for the run period before the July technical stop.

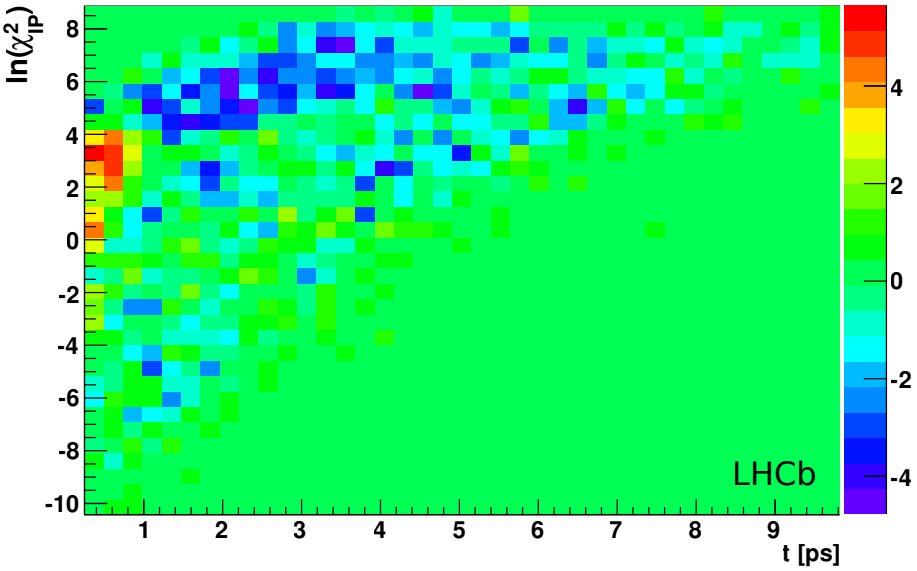


Figure 5.9: An example of the two-dimensional pull distribution. Plotted are the pulls for the fits to t and $ln(\chi^2_{IP})$ for the $\pi^+\pi^-$ final state, magnet polarity down for the run period before the July technical stop.

region where the secondary component dominates ($ln(\chi^2_{IP})$>3). Despite significant investigation into the secondary parametrisation no improvement could be made. The effect of this is included in the systematics studies as part of the assigned uncertainty due to the secondary component of the data. Most importantly the prompt peak looks to be well described and hence the fit was sufficient for the purposes of measuring A_Γ. Since the final states analysed are identical any potential biases in the lifetime due to the fit quality should be small and cancel when the resulting ratios are calculated for A_Γ. Section 5.4.1 gives details as to why the fit quality is considered to be sufficient for the measurement of A_Γ.

5.4 Systematic uncertainties and checks

5.4.1 Fitter Validation

The fitter was validated by determining the mean lifetime of the $K\pi$ mode; the measured lifetime value was (412.876 ± 0.084) fs. This shows reasonable agreement with the world average of (410.1 ± 1.5) fs [117]. As there is no physics information in the $K\pi$ lifetime this measurement was not blinded (whereas the K^+K^- and $\pi^+\pi^-$ states were). Furthermore no systematics studies were carried out for the $K\pi$ mode due to the long computing time required to analyse the large data-set. Therefore, although the statistical error is very small this is in no way a measurement of the mean D^0 lifetime beyond the fitter validation. The lifetimes of

	Pre TS		Post TS	
	MagUp	MagDown	MagUp	MagDown
$\tau(D^0)$ (fs)	414.70 ± 0.29	414.42 ± 0.29	412.07 ± 0.24	412.01 ± 0.18
$\tau(\overline{D}^0)$ (fs)	414.74 ± 0.29	414.92 ± 0.29	412.10 ± 0.24	411.66 ± 0.18
Average (fs)	414.72 ± 0.21	414.67 ± 0.21	412.09 ± 0.17	411.84 ± 0.13

Table 5.10: The measured $K^-\pi^+$ lifetime for the eight individual sub-samples. The average is un-weighted.

Split	Average D^0 and $\overline{D}^0$ τ (fs)
MagUp	413.14 ± 0.13
MagDown	412.68 ± 0.11
Pre TS	414.69 ± 0.15
Post TS	411.93 ± 0.10

Table 5.11: The average lifetimes for the run periods and magnet polarities.

the individual data samples can be seen in table 5.10 along with the D^0 and $\overline{D}^0$ un-weighted averages. The weighted lifetime averages for the data samples with different magnet polarities and before and after the technical stops are in table 5.11.

There is good agreement in the lifetimes for the split by magnet polarity but a discrepancy between the run periods of order 1%. For A_Γ such a lifetime difference should cancel out so the result would not be affected. As expected for the Cabibbo favoured mode ($K^-\pi^+$) the D^0 and $\overline{D}^0$ lifetimes agree giving $A_\Gamma(K^-\pi^+) = (-0.05 \pm 0.20) \times 10^{-3}$.

A null test was performed by measuring a pseudo-A_Γ with the K^+K^- and $\pi^+\pi^-$ final states using a random flavour tag. The fits were performed with samples that contained randomly selected D^0 and $\overline{D}^0$ candidates so each sample should have roughly equal numbers of each initial state. Measuring A_Γ should give 0 in this instance. For K^+K^- the measured value for the whole data set was $(1.0 \pm 0.6) \times 10^{-3}$, in $\pi^+\pi^-$ it was $(-0.4 \pm 1.1) \times 10^{-3}$; both consistent with 0.

Further validation was achieved with toy studies. 300 toy data sets of 1.5 million candidates were generated using the fitter configurations of the K^+K^- mode with the best fit results to data. The lifetime parameter was generated using what was the world average for y_{CP}, giving $\tau(K^+K^-) = 406.0\,\text{fs}$; A_Γ was assumed to be 0 so D^0 and $\overline{D}^0$ events were generated with the same lifetimes. The time dependence of the secondary shape in the Δm variable due to the DTF was included in the generation. Each individual 1.5 million candidates set was split into 8 to roughly match the statistics of the sub-samples in the real data (including the slight difference between D^0 and $\overline{D}^0$). The statistics of each roughly matches

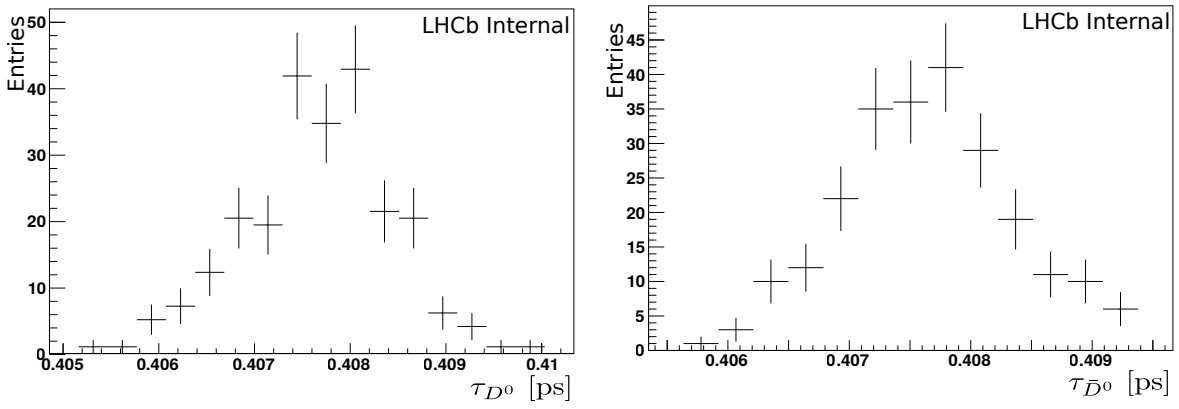


Figure 5.10: The results of fits to 300 toy simulation samples, generated with $A_\Gamma = 0$. Shown are the D^0 (left) and $\overline{D}^0$ (right) lifetimes. The means of the distributions are 407.7 fs (D^0) and 407.6 fs ($\overline{D}^0$).

the size of the $\pi^+\pi^-$ sample. Ideally larger toy sets would have been processed to reflect the larger K^+K^- and $K^-\pi^+$ modes but the available computing resources were prohibitive.

The lifetimes of both D^0 and $\overline{D}^0$ fits exhibited a noticeable bias of 1.7 fs as shown in Fig. 5.10. However the resulting A_Γ was unbiased (Fig. 5.11).

To check that the fit really is sensitive to measuring non-zero A_Γ values another set of toys was generated with $A_\Gamma = 5 \times 10^{-3}$ ($\tau(D^0) = 403.97$ fs and $\tau(\overline{D}^0) = 408.03$ fs). The result, shown in Fig. 5.12 shows no bias on the resulting A_Γ measurement.

Binned A_Γ

The A_Γ measurement was simultaneously made using the same data with an alternative method, known as the 'binned' method. The analysis was carried out completely independently and so it provides a cross check of the lifetime fit. The full details may be found in [119, 72].

The binned method measures A_Γ by measuring the ratio of the yields of signal D^0 and $\overline{D}^0$ candidates decaying to K^+K^- or $\pi^+\pi^-$ in bins of proper time. This ratio as a function of time, $R(t)$, for the K^+K^- mode can be expressed as

$$R(t) = \frac{N_{\bar{D}^0}}{N_{D^0}} \exp\left(1 + \frac{2A_\Gamma}{\tau_{KK}} t\right) \frac{1 - e^{-\Delta t/\tau_{\bar{D}^0}}}{1 - e^{-\Delta t/\tau_{D^0}}}, \tag{5.5}$$

where Δt is the width of the time bin. Thus by a fit of a straight line to a plot of $R(t)$ against t one may extract A_Γ from the gradient.

The blinding was implemented such that the two methods could be compared before unblinding. The blinded results for the total data sample are in table 5.12

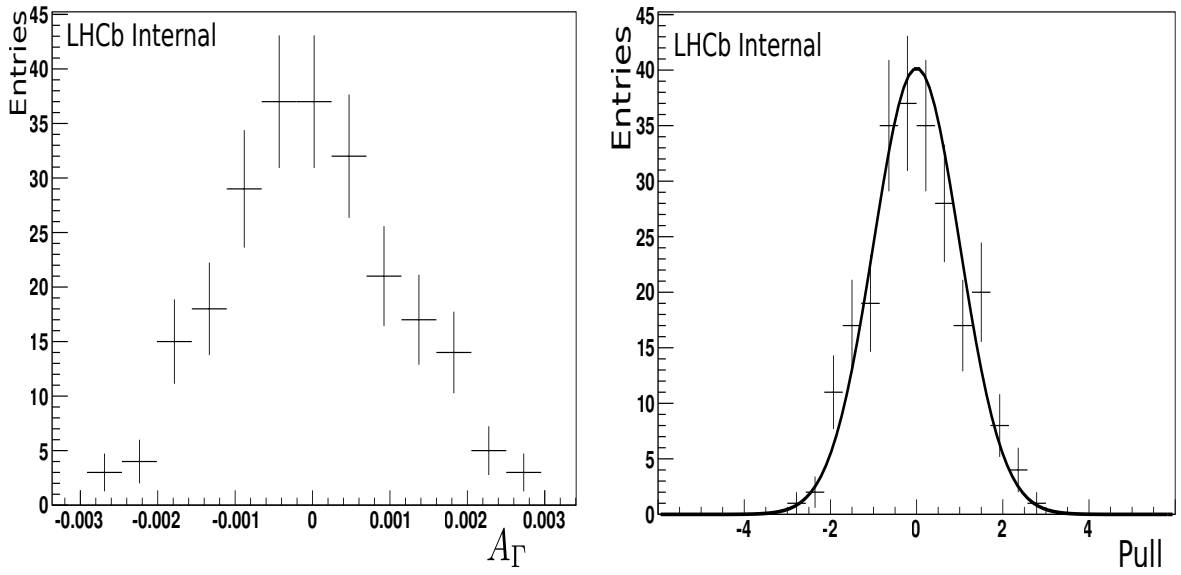


Figure 5.11: The results of fits to 300 toy simulation samples with a zero A_Γ value used in the data generation. On the left are the absolute measured values, on the right the pulls. The overlaid curve is a reference Gaussian with mean 0 and width 1. The mean values of each distribution are shown, along with the RMS of the pulls. The results are unbiased with the statistical uncertainty and the errors have been estimated correctly. The pulls have a mean of -0.03 ± 0.07 and RMS 1.1 ± 0.1; the result is unbiased.

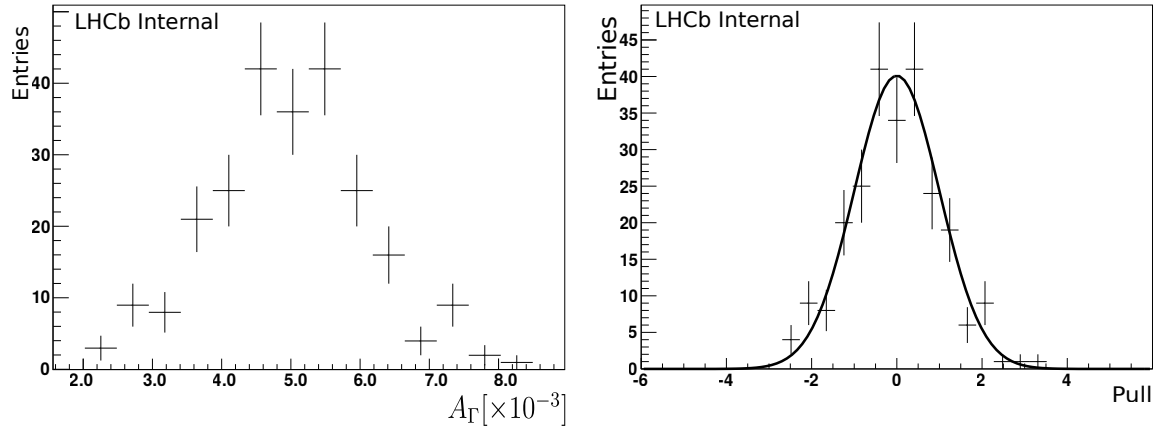


Figure 5.12: The results of fits to 300 toy simulation samples using a non-zero A_Γ in the data generation. On the left are the absolute measured values, on the right the pulls. The overlaid curve is a reference Gaussian with mean 0 and width 1. The pulls have a mean of -0.03 ± 0.07 and RMS 1.1 ± 0.1; the result is unbiased. The mean of the distribution of the measured A_Γ is $(4.97 \pm 0.07) \times 10^{-3}$, matching the value of A_Γ used in the generation.

	$A_\Gamma(K^+K^-)(\times 10^{-3})$	$A_\Gamma(\pi\pi)(\times 10^{-3})$
Unbinned	-16.4 ± 0.63	-22.9 ± 1.06
Binned	-15.68 ± 0.80	-21.73 ± 1.16

Table 5.12: The blinded results for the binned and unbinned methods. The quoted uncertainties are statistical.

and can be seen to be consistent.

5.4.2 Systematic uncertainties

The tests outlined in section 4.2.11 were performed. Typically the tests tend to show a pronounced effect on the measured D^0 and $\overline{D}^0$ lifetimes but the changes are very similar between the two. Therefore these shifts largely cancel in the ratio that is calculated for measuring A_Γ. For example Fig. 5.17 shows the lifetime deviations for one of the tests to ascertain the turning point accuracy, with the A_Γ results shown in Fig. 5.15. The lifetimes change in a similar manner leading to a small A_Γ variation.

The results of all the tests to determine the systematic uncertainties are summarised in table 5.13 with the sign (positive or negative) of the shifts where only one test has been compared to the standard fit. In the case where more than one test has been performed for a source the largest resulting uncertainty was assigned. The assigned uncertainties are summarised in table 5.14.

Decay-time resolution and efficiency correction

The decay-time resolution was tested through variations to the resolution model; either changing the Gaussian resolution width from the standard value (50 fs) between 30 fs to 100 fs or introducing the more complex function described in section 4.2.11. Both tests of the decay time resolution returned absolute uncertainties of the order of 10^{-7} on A_Γ for both final states. The uncertainty due to the resolution model was therefore considered to be negligible. Using the efficiency correction in the signal lifetime returned a similarly negligible effect on A_Γ.

π_s^+ background

The tests for the effects of inaccuracies in modelling the π_s^+ background were only carried out on the $\pi^+\pi^-$ sample, the same result being assigned to the two final states. Allowing for a fully independent description of the π_s^+ component in the m_{D^0} variable led to a negligible change in A_Γ, as did treating the π_s^+ lifetime parameter as a multiplication of the signal lifetime parameter. Allowing only the

Source	Test	$A_\Gamma(KK)(\times 10^{-3})$	$A_\Gamma(\pi\pi)(\times 10^{-3})$
Turning points	TP bias	0.09	0.11
	TP scale	0.02	0.05
Resolution	Vary σ	0.00	0.00
	Change model	0.00	0.00
Secondary	Vary τ_1	0.00	0.00
	Toys $2\times$ scndry. fraction	-0.03	-0.03
	Toys $0.5\times$ scndry. fraction	-0.07	-0.07
	Toys t dep. $ln(\chi^2_{IP})$	-0.05	-0.05
π^+_s bkg.	Independent m_{D^0} PDF	0.00	0.00
	Independent m_{D^0} width	+0.01	+0.01
	Scale π^+_s τ	0.00	0.00
Combinatoric bkg.	Toys change m_{D^0} and Δm	0.00	0.00
	Toys change t	+0.02	+0.02
	Toys larger fraction for $\pi^+\pi^-$	–	-0.04
Mis-reconstructed bkg.	Fix t PDF	0.02	-
	Fix component fraction	0.02	-

Table 5.13: The results of the systematic uncertainty tests. For those where the evaluated uncertainty was less than 1×10^{-5} the effect was considered negligible so 0.00 is quoted. The numbers quoted are the largest possible variations seen. For those tests to data the total $\pm$ uncertainty is given; for the toy tests the sign of the bias is indicated.

width of the π_s^+ mass peak to deviate from the signal introduced a change in A_Γ that gives an uncertainty of 0.01×10^{-3}. The sign of the change for the single test is positive but a symmetric uncertainty has been assigned.

Combinatoric background

Two toy tests were carried out, one varying the mass and Δm description, the other the unbiased time distribution, both in the generation of the data. The first test resulted in a bias of -0.01 ± 0.07 in the mean of the pulls. The second gave 0.01 ± 0.10. The default bias is -0.01, so the test varying the time distribution gave the biggest relative uncertainty of 2%. Multiplying by the $\pi^+\pi^-$ statistical uncertainty yields an absolute uncertainty on A_Γ of 0.02×10^{-3}.

In addition a $\pi^+\pi^-$ specific test with toys was carried out whereby the generated data had a greatly increased combinatoric component to more closely reflect that in the data for this final state. The fits to these data sets exhibited a bias in the mean of the pulls of -0.04 ± 0.06 leading to an uncertainty of 0.04×10^{-3} for the $\pi^+\pi^-$ final state.

Mis-reconstructed backgrounds

The uncertainty due to the effect of the mis-reconstructed backgrounds in the K^+K^- mode involved fixing the unbiased lifetime estimate PDF instead of iterating with the baised distribution from the sWeights. The fixed PDF was a single exponential the lifetime parameter of which was varied. For the $D^0 \to K^-\pi^+\pi^0$ backgrounds the lifetimes tried were 550 fs, 740 fs and 1200 fs; for $D_s^+ \to K^-K^+\pi^+$ they were 520 fs, 860 fs and 1250 fs. These numbers were motivated by fits to simulated data with and without the stripping selections and the mean of the likelihood distributions that are produced from the first stage of the fit (1226 fs and 1250 fs for $K^-\pi^+\pi^0$ and D_s^+). The alteration was done in pairs, 550 fs with 520 fs and so on. The systematic uncertainty was evaluated to be 2.4×10^{-5} (see Fig. 5.13).

A second test was carried out whereby the fractions of the each mis-reconstructed background component was fixed at to be at one standard deviation above or below the standard fit results to mimic the effect of a possible incorrect shape leading to their contributions being under or ever estimated. Each of the two contributions was varied over the possible nine combinations with the resulting uncertainty being 1.7×10^{-5}.

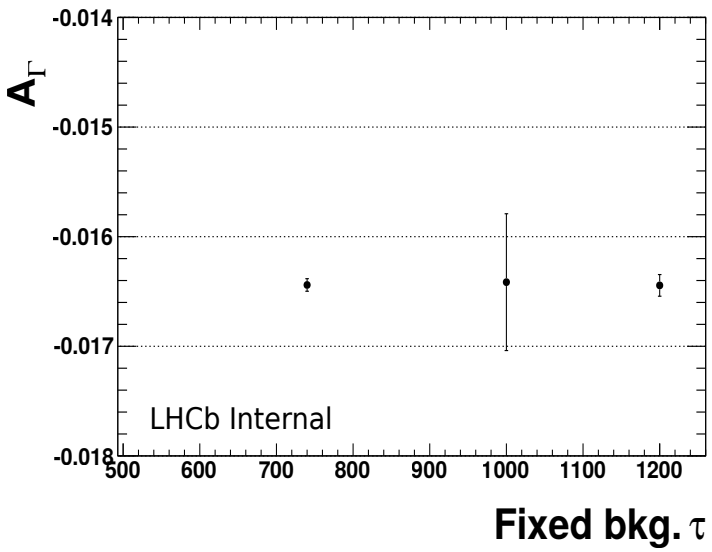


Figure 5.13: The variation of A_Γ due to varying the lifetime of the unbiased PDFs of the mis-reconstructed backgrounds. Shown is the average value of A_Γ for all the subsamples.

Secondaries

Fits were performed with the fixed lifetime parameter of the secondary unbiased lifetime PDF being varied between 0.15 ps and 0.35 ps (the standard fit fixes it at 0.27 ps). The variation in A_Γ caused by this was 2.7×10^{-6} for $\pi^+\pi^-$ and 2.0×10^{-6} for the K^+K^- mode.

Toy tests were also carried out as outlined in section 4.2.11. Fits to the toy data sets with double and half the standard secondary fractions yielded biases in the mean of the pulls for A_Γ of -0.03 ± 0.09 and -0.06 ± 0.08 respectively.

The second test to change the time dependency of the mean parameter of the secondary $ln(\chi^2_{IP})$ distribution was designed to examine the effect of an inaccurate description of that evolution in the fit. This was believed to have lead to the poor fit quality in Fig. 5.7. The time dependency was taken from fits to an enriched sample of secondary decays. These were events that had already been reconstructed as $B \to D$ decays which were then reconstructed as prompt decays. They were binned in decay time and fitted with the $ln(\chi^2_{IP})$ PDF in order to extract an estimate of the time evolution of the $ln(\chi^2_{IP})$ variable (similarly to the plots in Fig. 5.5 but using data instead of simulation). A spline was applied to those fit points and used as the time dependency when generating the toys. The toys were then fitted with the standard configuration with the exception that the fixed secondary lifetime parameter was changed to 0.2 ps in the fit to try and mimic the poor pulls at low decay times seen in the fits to data. Such a test yielded a bias on the pulls of -0.05 ± 0.07.

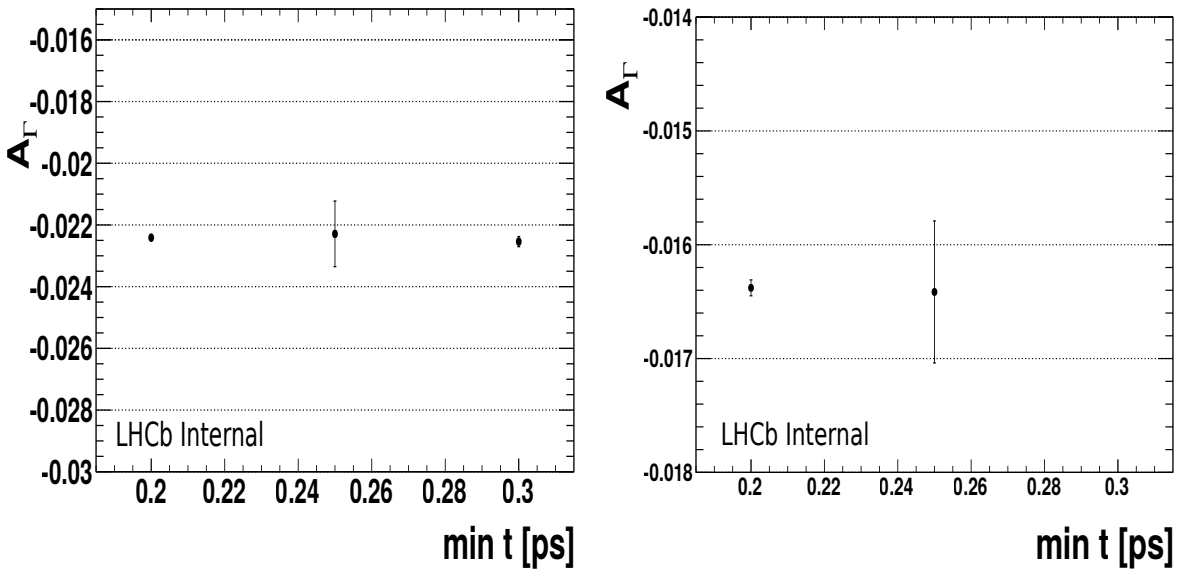


Figure 5.14: The effect on the measured A_Γ of varying the minimum time cut. On the left is the $\pi^+\pi^-$ final state and the K^+K^- is on the right.

For evaluating the toy uncertainties whichever was the larger of the difference in the biases between the test and the standard configuration (-0.01), or the absolute value of the bias was taken. In this case the largest pull was -0.06, or a relative uncertainty of 6%. The toy data sets were of similar size to the $\pi^+\pi^-$ sample so the relative uncertainty (6%) was multiplied by the $\pi^+\pi^-$ statistical uncertainty (1.06×10^{-3}) to yield the absolute systematic uncertainty on A_Γ of 0.07×10^{-3} which was assigned for the effect of the modelling of the secondary component of the data in the fit.

Minimum decay time cut

Any change in A_Γ due to a change in the value of the minimum decay time cut should be statistical in nature. The fits were performed with the minimum time cut varied from 0.15 ps to 0.35 ps. To test whether the resulting variation was statistical in nature, a χ^2 probability of the variation being consistent with a statistical variation was calculated. For the two final states the probability was large; 50.4% for K^+K^- and 9.9% for $\pi^+\pi^-$. It was therefore decided that the variation observed was consistent with statistical variation and so a systematic uncertainty was not assigned for this effect. The deviations can be seen in Fig. 5.14.

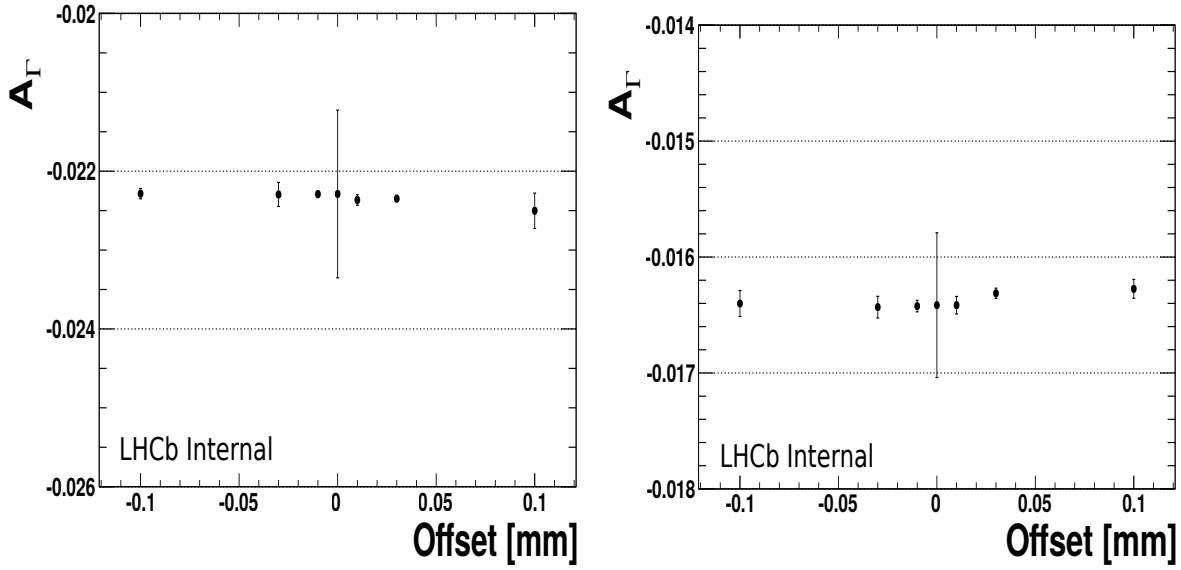


Figure 5.15: The variation of A_Γ due to adding an offset to each turning point. The $\pi^+\pi^-$ final state is on the left, K^+K^- on the right.

Turning point accuracy

The effects of the limited turning point accuracy was tested through altering the turning points in one of two ways. The first was to apply a constant offset to each turning point of up to $\pm 0.1\,\text{mm}$. This covers a range of two multiples of the smallest step size used in the swimming method when searching for turning points. The second alteration was to apply a scale to each turning point of up to $\pm 1\%$.

The former test produced the biggest deviation for both modes and the values were assigned for the systematic uncertainties due to the limited turning point accuracy; the values were 0.11×10^{-3} in $\pi^+\pi^-$ and 0.09×10^{-3} for K^+K^- from applying the offset. Applying the scale resulted in 0.05×10^{-3} and 0.02×10^{-3} respectively. The results of the first test is in Fig 5.15 the second in Fig 5.16. Fig. 5.17 shows how the D^0 and $\overline{D}^0$ lifetimes are affected by the offset test for the K^+K^- final state; as can be seen, to a large degree the systematic variation has cancelled when A_Γ is calculated. Nevertheless the uncertainty due to the turning point precision remains significant.

Summary

A summary of the assigned systematic uncertainties for each source is in table 5.14 along with the total systematic and statistical uncertainties. All of the systematic uncertainties are assumed to be symmetric in nature. They are so small relative to the statistical uncertainty that this conservative estimate has been used. In particular for the toys, although an individual bias may appear in one direction

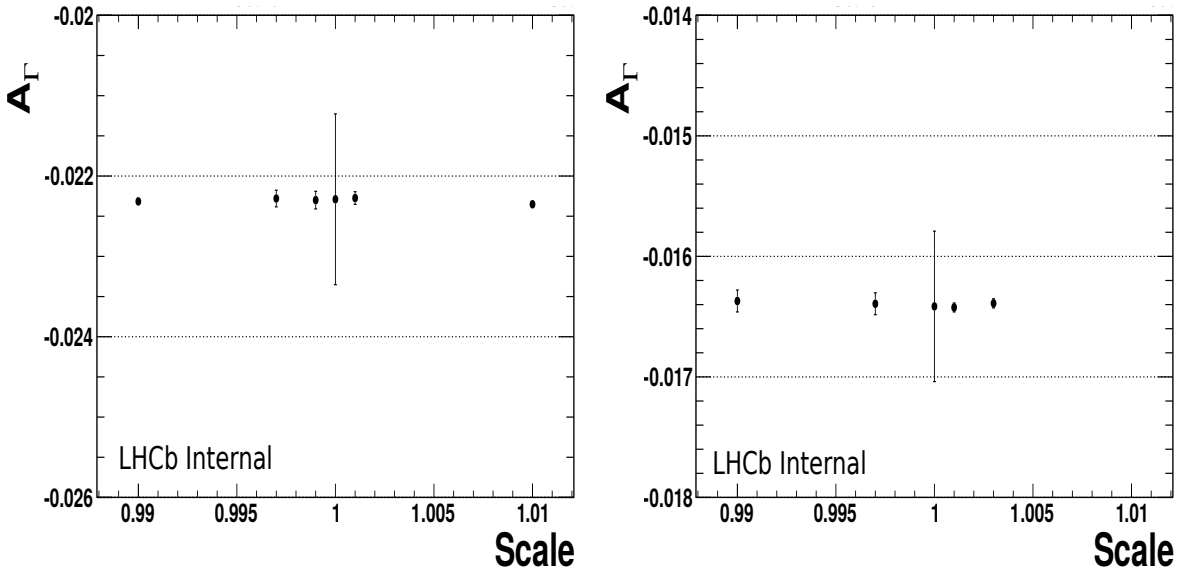


Figure 5.16: The variation of A_Γ due to applying a scale to each turning point. The $\pi^+\pi^-$ final state is on the left, K^+K^- on the right.

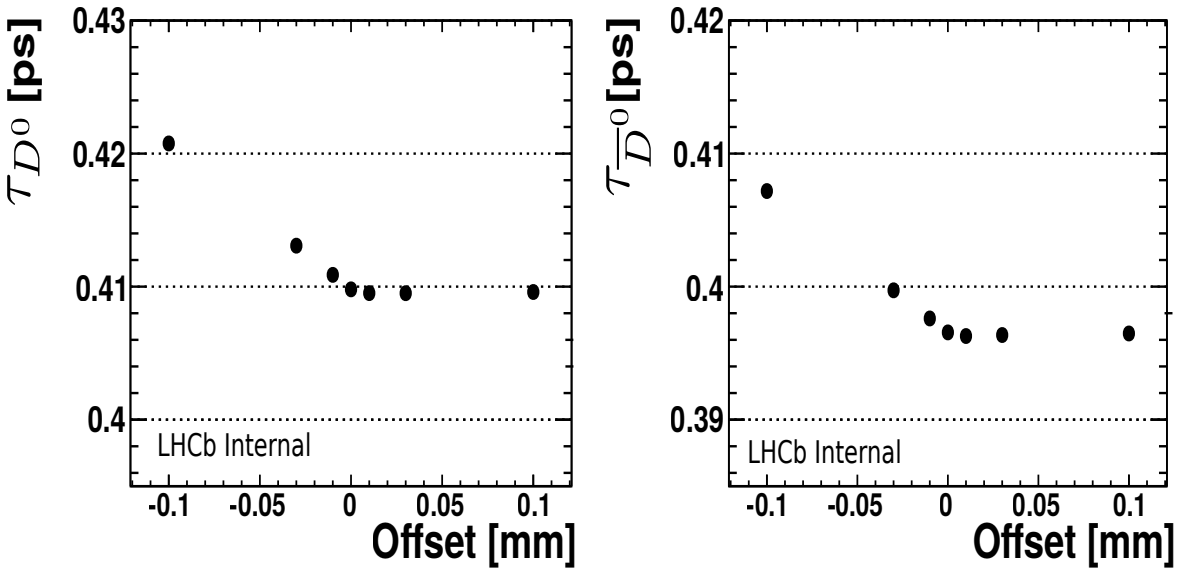


Figure 5.17: The variation of the D^0 (left) and $\overline{D}^0$ (right) lifetimes for the K^+K^- final state due to the TP offset test.

Source	$A_\Gamma(KK)$ $\times 10^{-3}$	$A_\Gamma(\pi\pi)$ $\times 10^{-3}$
Mis-reconstructed bkg.	± 0.09	-
Secondary bkg.	± 0.07	± 0.07
Combinatoric bkg.	± 0.02	± 0.04
π_s^+ bkg	± 0.01	± 0.01
Turning points	± 0.09	± 0.11
Time resolution	± 0.00	± 0.00
Reconstruction eff.	± 0.00	± 0.00
Total systematic uncertainty	± 0.12	± 0.14
Total statistical uncertainty	± 0.62	± 1.06

Table 5.14: The assigned systematic uncertainties for each potential source along with the total systematic and statistical uncertainties.

Variable	Ranges
D^0 p	$<45\,\text{GeV}/c$, 45–65 GeV/c , 65–95 GeV/c , $>95\,\text{GeV}/c$
D^0 p_T	$<3.3\,\text{GeV}/c$, 3.3–4.3 GeV/c , 4.3–6 GeV/c , $>6\,\text{GeV}/c$
D^0 η	0–3 , 3–3.5 , 3.5–4 , 4–6
D^0 ϕ	detector halves and quadrants
π_s^+ χ^2_{IP}	< 1.5 , 1.5–7.5 , >7.5
No. of PVs	1 , 2 , >2
No. of tracks	< 50 , 50–100 , >100

Table 5.15: A summary of the ranges of the kinematic checks.

the statistical uncertainty of the toy test is usually larger than the resulting shift in bias so one should be conservative.

5.4.3 Checks

Several checks were carried out in addition to the systematics checks detailed above. The first of these were to ensure that the result was not dependent upon the kinematics of the event. The data was split by kinematic variables and for each split the fits were carried out and A_Γ computed. These kinematic checks and splits are detailed in table 5.15. The split by ϕ refers to the azimuthal angle of the D^0 flight direction. The cuts are therefore designed to check that the results is not dependent on which part of the detector the D^0 decay components are found in; they pick out the top, bottom, left and right halves and the individual quadrants (top left, top right *etc.*.). Additional splits were made by the π_s^+ PID with respect to leptons and the level 0 trigger decision. For all these checks no systematic trend was seen. As an example the results of the D^0 transverse momentum split is shown in Fig. 5.18.

The effect of selecting events with only two turning points was checked by

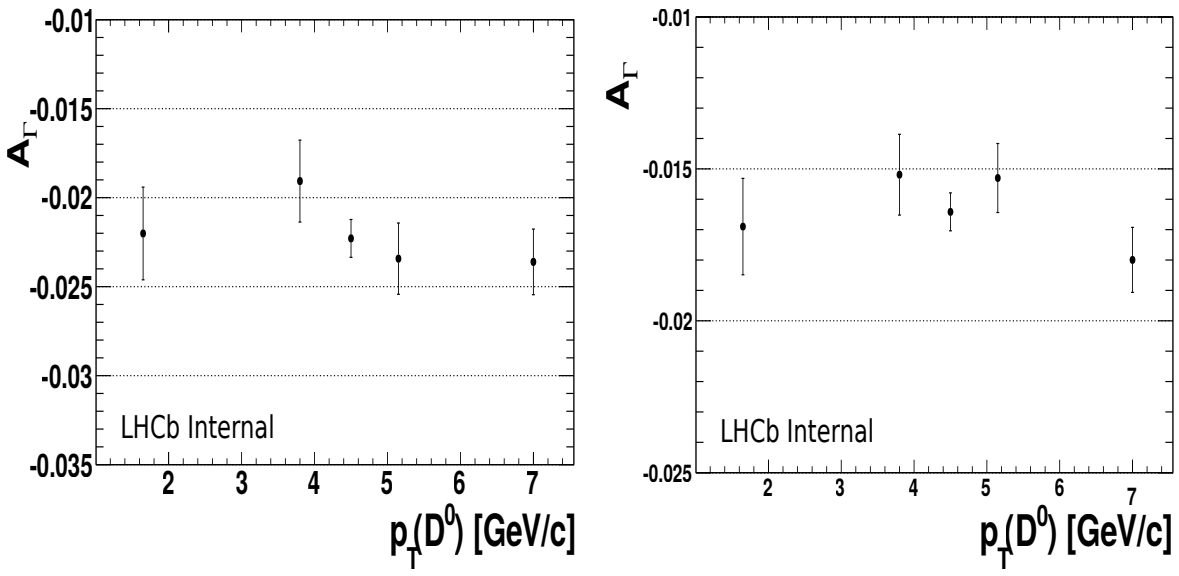


Figure 5.18: The value of A_Γ determined for different ranges of D^0 candidate transverse momentum. The $\pi^+\pi^-$ final state is on the left, K^+K^- is on the right.

doing the fits with the maximum number changed to four. The fits were also repeated with a different candidate selected in those events with multiple candidates to ensure the candidate selection did not bias the result. Different selections in the hardware trigger were tried to ensure the trigger selections did not affect the result. Finally cuts on the likelihood of the π_s^+ being a pion compared to a lepton were imposed. For all of these checks no significant deviation in A_Γ was observed beyond statistical variations.

5.5 Results

The results obtained for A_Γ measured using the 2011 data set are

$$A_\Gamma(KK) = (-0.35 \pm 0.62 \pm 0.12) \times 10^{-3}$$
$$A_\Gamma(\pi\pi) = (0.33 \pm 1.06 \pm 0.14) \times 10^{-3},$$

where the first error is statistical and the second systematic. The weighted average of the two modes is

$$A_\Gamma = (-0.17 \pm 0.54) \times 10^{-3}.$$

The two modes are consistent with each other and with there being no *CP* violation at the quoted precision. The ratios of D^0 and $\overline{D}^0$ candidates as a function of decay time are plotted in Fig. 5.19 for the two final states. Overlaid

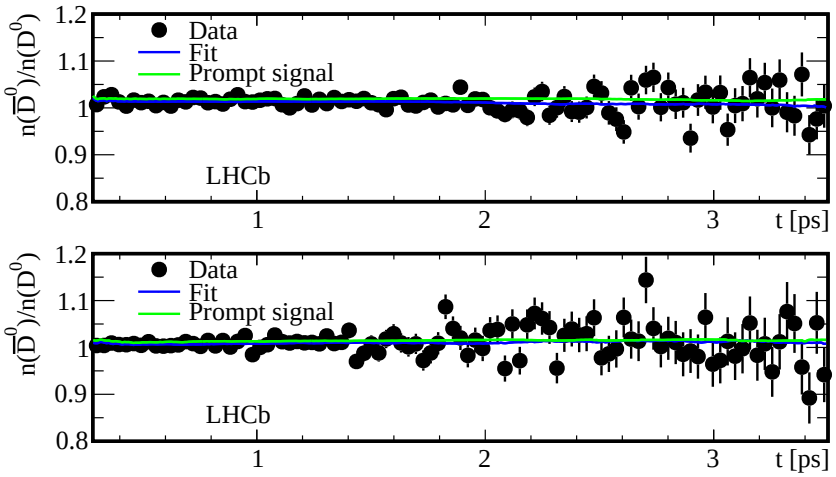


Figure 5.19: The ratios of the extracted D^0 and $\overline{D}^0$ yields as a function of decay time as they decay to the K^+K^- (top) and $\pi^+\pi^-$ (bottom) final states. The green curve is the ratio of the fitted signal PDFs, the blue the total PDFs.

are the ratios of the fitted signal PDF (green) and total PDF (blue). The data and fits are flat, showing no time dependence which indicates that there is no time-dependent *CP* violation observed.

The resulting HFAG [41] average is shown in Fig. 5.20. As can be seen the measurement with the K^+K^- final state represents a considerable improvement of a factor of four in precision over previous efforts. This was also the first time that A_Γ had been presented separately for the two modes.

5.6 Further checks

Following the publication of the A_Γ measurement [119], new sources of bias have been identified. These are the dependence of the $ln(\chi^2_{IP})$ variable on the D^0 mass and a potential mistreatment of the acceptance functions. By its nature, the observable A_Γ should be robust against such biases. A full description of these effects is given in sections 6.4.2 and 6.1.1 respectively. Here only the effect on A_Γ is described.

5.6.1 Acceptance function

The realisation that the acceptance that had been used in the fits is false has implications for the measurement of A_Γ. Whilst biases due to the acceptance should mostly cancel in the difference, the acceptance potentially skews the fit results in a non-trivial manner (as well as contributing to the poor fit quality). This should to a certain extent be described by the systematic tests regarding the

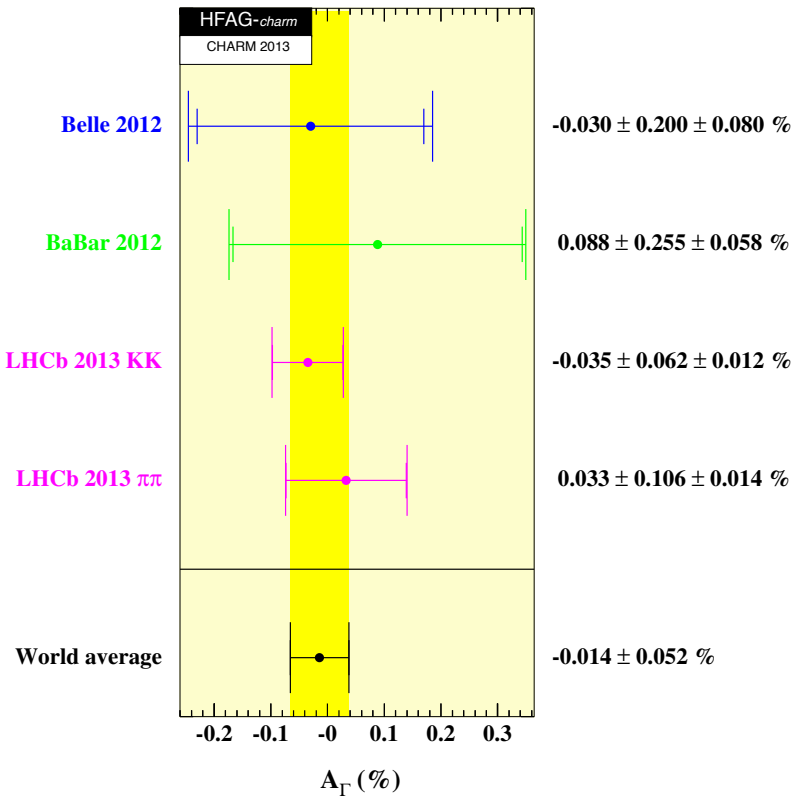


Figure 5.20: The HFAG average for A_Γ in September 2013 featuring the new $A_\Gamma(K^+K^-)$ and $A_\Gamma(\pi^+\pi^-)$ results. Reproduced from [41].

	$\pi^+\pi^-$	K^+K^-
Standard	1700255	4882719
Both tracks reco.	1584511	4580088

Table 5.16: The data set sizes for the standard reconstruction and that requiring both tracks to be reconstructed in the high level trigger. The fractional change is consistent between the final states, with about 93% of candidates returned in the latter sample.

turning point accuracies.

Another test has been carried out with fits to data. As well as the trigger decision, the swimming algorithm also contains information as to whether the track was reconstructed in the trigger. Therefore for this test (and for the future y_{CP} analysis) the requirement has been made that both tracks that have been reconstructed offline were reconstructed in HLT1. This should imply that for the dataset being analysed the online-offline efficiency is one so there is no step in the acceptance should a second track fire the trigger. There is a slight loss in statistics as shown in table 5.16 but it is far less severe than the selection on the trigger decisions.

As the precise fitter configuration has changed somewhat since the original analysis this test determines the change in the A_Γ central value between the full dataset and the dataset with both tracks reconstructed in HLT1. The full dataset and the dataset with restriction on the tracks were both fitted under the same conditions. For the $\pi^+\pi^-$ state the difference between the two is $(+0.24 \pm 0.22) \times 10^{-3}$ and for K^+K^- it is $(-0.12 \pm 0.16) \times 10^{-3}$ where the uncertainties are statistical due to the changes in sample size. These represent deviations of $1.1\,\sigma$ and $0.74\,\sigma$ relative to the change in statistical uncertainty. They are not in the same direction. One can therefore conclude that the shifts seen are statistical in nature and not systematic.

5.6.2 Mass–$ln(\chi^2_{IP})$ correlation

The potential bias due to the dependence of $ln(\chi^2_{IP})$ on mass for the prompt components of the data have been assessed using simplified simulations ('toys'). As one is comparing identical final states for A_Γ there is limited scope for this correlation to introduce a bias. If the size of the bias on lifetimes is somehow dependent on sample size then that would propagate to A_Γ as D^0 and $\overline{D}^0$ comprise different fractions of the data due to the charm production asymmetry [121] (the difference in data is close to 1% in favour of $\overline{D}^0$). Alternatively a significant true value of A_Γ could lead to the biases on the lifetimes of the D^0 and $\overline{D}^0$

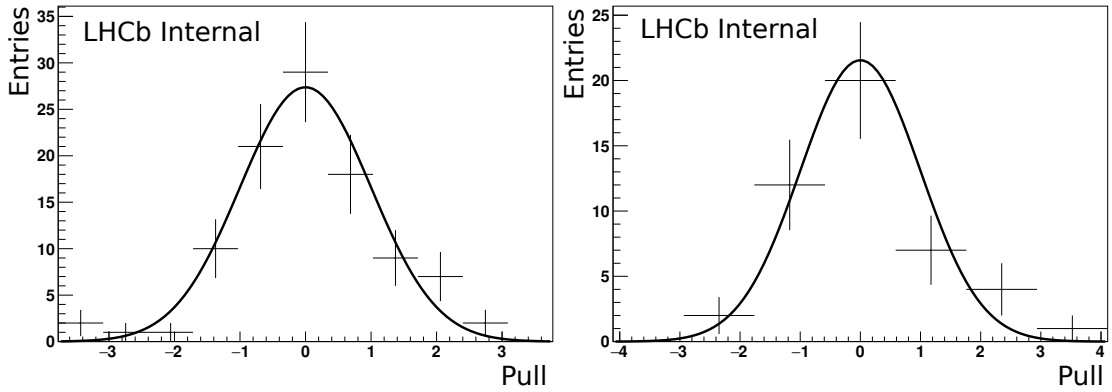


Figure 5.21: The result of the tests with simplified simulations for the $ln(\chi^2_{IP})$- mass correlations. On the left is the test with $A_\Gamma = 1 \times 10^{-3}$, on the right the test with the data generated with D^0 fraction of 40%. Both plots show the pulls of the fitted datasets. The means are 0.05 ± 0.11 and -0.01 ± 0.18 respectively.

measurements being different, again leading to a bias on the measured A_Γ.

Both hypotheses have been tested with toy simulations for the K^+K^- final state which has a more complex background description and so more scope for adjustments to creep in. Data was generated with the correlation included for the prompt components of the signal and random slow-pion background classes and then fitted with the standard fit configuration. The parametrisation for the correlation in the generation was taken from a fit to the simulated data in Fig. 6.3.

To test the effect of the difference in D^0 and $\overline{D}^0$ sample sizes toys were generated with the standard 1% larger $\overline{D}^0$ fraction as well as exaggerated samples of 60% and 40% D^0. In all three instances although the individual lifetimes were significantly skewed, the resulting A_Γ was unbiased. The results for the 40% D^0 test are shown in the right of Fig. 5.21.

For the second test the toys were generated with an A_Γ value of 1×10^{-3} which is indicative of the magnitude of the current HFAG average of $(-0.59 \pm 0.4) \times 10^{-3}$ [41] plus its uncertainty (the toys are agnostic to the sign of A_Γ). This is the largest value of A_Γ that the analysis would reasonably be expected to measure given previous results. Again, although the individual lifetimes are shifted, the resulting A_Γ measurement is unbiased. The pull plot for this test is the left plot of Fig. 5.21.

Chapter 6

2011 y_{CP} analysis

The measurement of y_{CP} using the 2011 data set was separated from the A_Γ results. Initially this was due to the greatly increased computing time required to process the systematic checks on the larger $K^-\pi^+$ data set and the desire to publish A_Γ in a timely fashion. Additionally y_{CP} is a rather more tricky measurement due to the different final states so any biases or systematic effects do not cancel to such a degree as they did for A_Γ. It was also hoped that a better fit configuration could be produced to improve the quality and stability of the fits. In this section the changes from the A_Γ analysis are detailed, the systematic studies and finally the result for the measurement of y_{CP} using the 2011 data set are presented.

At the time of writing this analysis is under review within the collaboration. The intention is to publish the final result in the first half of 2016.

6.1 Fitter updates

Several updates were implemented for the measurement of y_{CP} and they are discussed in the following sections. The measurement of A_Γ and the LHCb measurement of ΔA_{CP} [56, 55] showed that at this precision the K^+K^- and $\pi^+\pi^-$ final states should have the same effective lifetimes. Therefore the difference between them was unblinded (in addition to the difference between D^0 and $\overline{D}^0$ from the A_Γ analysis) in order to provide a further powerful check on the results.

6.1.1 Turning point treatment

During the course of updating the analysis several inconsistencies were found amongst the results. In particular there was a difference in lifetime between the data samples either side of the July technical stop of almost 1% ($\sim$4 fs), first noted in the 2011 A_Γ analysis. Although the resulting y_{CP} measurements were consistent

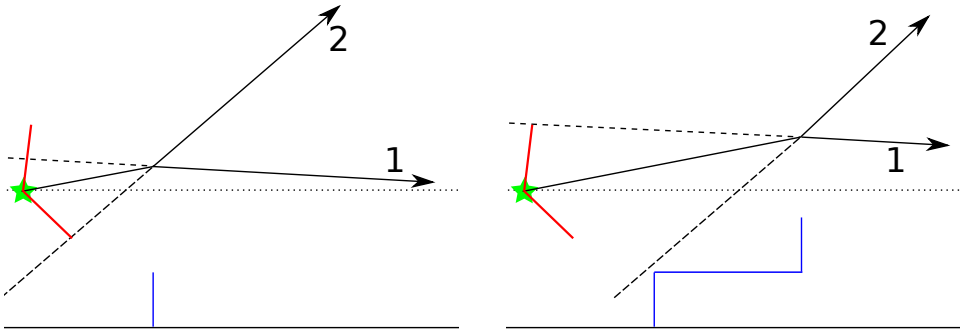


Figure 6.1: The acceptance function for an event with two tracks reconstructed in the trigger. On the left the D^0 decay time is such that the trigger impact parameter requirement on track two has been satisfied and there is a step in the acceptance. The other track in the event (one) does not pass the trigger selection. On the right the primary vertex has been swum such that the second track now passed the trigger, as well as the first track leading to a second step in the acceptance. The heights of the two steps depends on the efficiency for a track to be reconstructed in the trigger.

for the samples split by time period it was noted that such a large discrepancy of the lifetimes, ten times larger than the statistical uncertainty for the $K^-\pi^+$ mode, was undesirable if one were to claim that the acceptances were being dealt with properly. More concerning was a consistent difference in the measured lifetimes of the *CP* eigenstates of roughly two standard deviations significance (see the left plot of Fig. 6.2). This remained regardless of any alterations to the fit model or cuts on the backgrounds.

It was eventually realised that the previous treatment of the turning points to extract the per-event acceptance was flawed. The selection required one of the reconstructed tracks to be selected in the first level of the trigger. One therefore has an acceptance due to the trigger for each of the two tracks. If only one track from the event can pass the trigger over the whole swimming range the acceptance would be a simple top hat as before. However in the event that both tracks could pass the trigger, possibly at different lifetime ranges then one arrives at a more complex acceptance for the event as there would be two possibilities that the event would be selected. The problem is shown in Fig. 6.1. The correct implementation of such an acceptance is described in [73]. The acceptance has to be normalised accounting for both tracks in the trigger, with the single track efficiency (the number of the decay product tracks that are reconstructed by the trigger compared to the number selected offline) fitted for.

To facilitate the more complex acceptance would require extensive modification of the analysis method. As an alternative one could require both tracks to fire the first level trigger, obtaining only a single top-hat acceptance for each event. The caveat is that a large quantity of the recorded data is now rejected

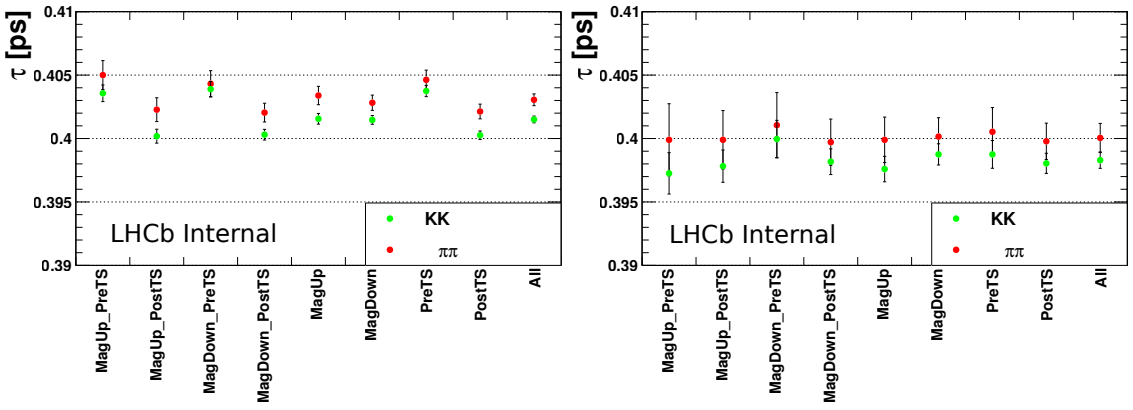


Figure 6.2: The effect of the acceptance treatment on the measured lifetimes of the K^+K^- (green) and $\pi^+\pi^-$ (red) modes. On the left are the results of fits to the entire dataset and on the right when selecting that both daughter tracks are selected by HLT1. The agreement between the two modes is much improved when applying the additional selection and the split in lifetimes around the technical stop is much reduced for both K^+K^- and $\pi^+\pi^-$.

with only around 30% of the data remaining. This leads to a considerable loss of statistical precision.

The effect of the acceptance correction can be seen in Fig. 6.2 which compares the lifetimes of the *CP* eigenstates when fitted with the entire data set or when selecting those candidates with both daughters firing HLT1. For the entire dataset there is a significant difference between the fitted K^+K^- and $\pi^+\pi^-$ lifetimes, although they should be the same. Both modes also show a large difference in the lifetimes either side of the July technical stop. For the data with both daughter tracks passing HLT1 on the right, the difference between the modes is reduced, as is the split around the technical stop.

6.1.2 Mass - $ln(\chi^2_{IP})$ correlation

The second major change to the analysis is again due to an assumption proving to be false. One of the most basic requirements of the analysis is that the variables in the first stage of the fit (D^0 mass and Δm) are independent of the second stage (decay time and $ln(\chi^2_{IP})$), allowing for the total fit PDF to be factorised and the analysis to be performed in these two stages. Unfortunately an examination of simulated data showed this is not the case. Consider Fig. 6.3 which plots the arithmetic mean of the $ln(\chi^2_{IP})$ distribution in bins of mass for simulated $D^0 \to K^-\pi^+$ decays; the correlation between the two variables is clear to see. This correlation could potentially cause a large bias in the individual lifetimes and if it is final state dependent that would propagate into a large bias on y_{CP} (as A_Γ measures the same final states any bias would cancel).

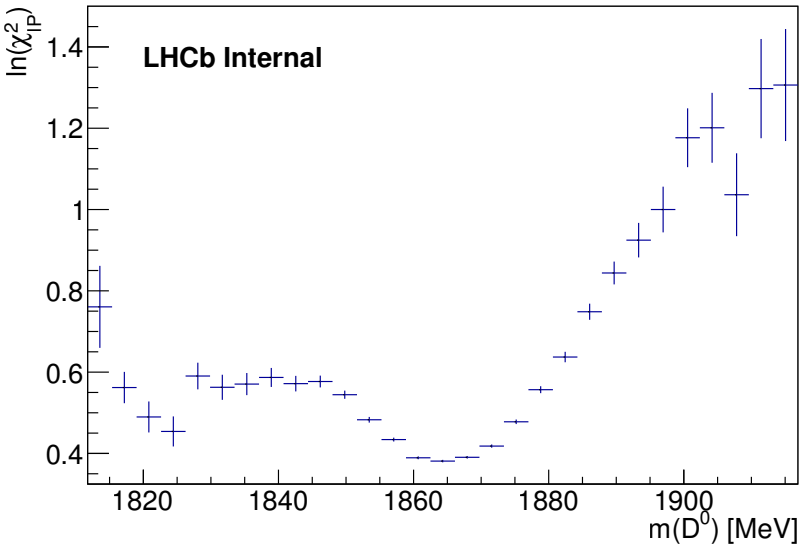


Figure 6.3: The correlation between D^0 mass and $ln(\chi^2_{IP})$. Plotted is the mean of the $ln(\chi^2_{IP})$ distribution in bins of D^0 mass ascertained from $D^0 \to K^-\pi^+$ simulation.

Quantity	Bias
$\tau_{\pi^+\pi^-}$	-0.33
$\tau_{K^+K^-}$	-0.99
$\tau_{K^-\pi^+}$	0.32
$y_{CP}(\pi\pi)$	0.461
$y_{CP}(KK)$	0.99

Table 6.1: The biases due to the correlation between the D^0 mass and $ln(\chi^2_{IP})$ variables. Show are the differences in the biases on the pulls between simulations with and without the correlation applied when generating the data.

To assess the size of the bias a test with simplified simulations was performed. Samples of simulated data for the three final states were obtained and the correlations between mass and $ln(\chi^2_{IP})$ plotted and fitted. Subsequently datasets were generated with these correlations and then fitted with the default fit. The subsequent bias on y_{CP} was compared to that obtained by generating and fitting data without the correlations (otherwise with the same configuration).

The results of the test for the three lifetimes is shown in table 6.1.

As the bias was found to potentially be very large (of the order of the statistical uncertainty in the K^+K^- mode) it was decided to perform the analysis in a reduced bin of D^0 candidate mass. The chosen bin is 1860–1870 MeV, centred around the D^0 mass.

To improve the statistics in each individual fit, D^0 and $\overline{D}^0$ candidates were fitted together. Additionally fits of mass and Δm over the full mass range were performed (1815–1915 MeV) in order to extract the sPlots of the background

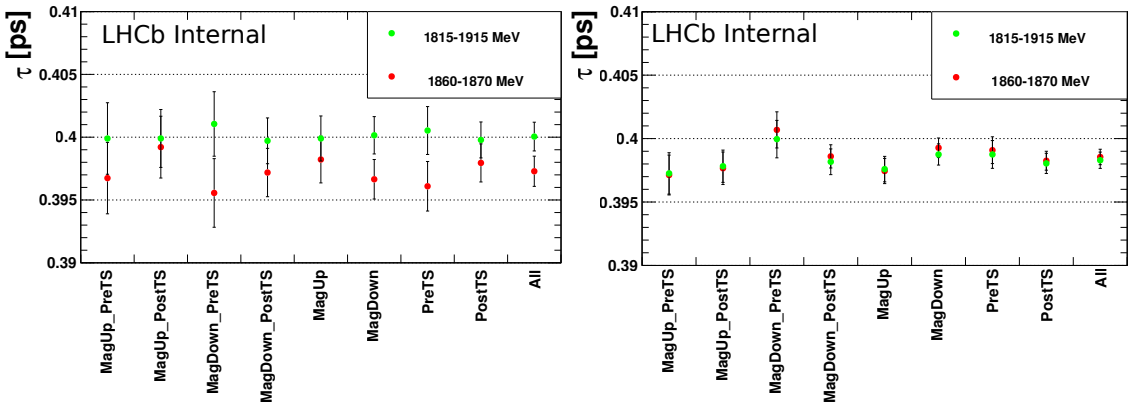


Figure 6.4: The measured blinded lifetimes when fitting to the whole mass range (red) and the mass range 1860–1870 MeV (green). Shown are the lifetimes of the $\pi^+\pi^-$ (left) and K^+K^- (right) final states. There is a significant change in the lifetimes of the $\pi^+\pi^-$ fits but for the K^+K^- mode there is little change.

components and the first estimates of the turning point PDFs. The turning point PDFs were subsequently unfolded individually for each sub-sample. These shapes were then used in the fits over the reduced range. For the individual fits in the mass bin the shape parameters of the PDFs for the mass variables were fixed. The Δm PDF shapes were allowed to float to account for the change in Δm resolution with D^0 mass. The fractions of each component in the fit in the bin in question were calculated and fixed in the fit as there is little discrimination between the components in such a small fit range.

The effect of the correlation on the lifetime fits with data can be seen in Fig. 6.4. The blinded results of the *CP* eigenstate lifetimes are shown, both with the fit performed in the small mass bin and for the single fit done over the full mass range. There is a clear shift in measured lifetimes of the $\pi^+\pi^-$ final state.

6.1.3 Backgounds

A bug was found in the fitter implementation such that in effect the acceptance function was being applied twice to the unbiased lifetime PDFs for these components. For the A_Γ analysis this was particularly noticeable in the $\pi^+\pi^-$ $ln(\chi^2_{IP})$ fits, where the combinatoric component was shifted to higher values than it should have been. It was also apparent in the toy studies although the cause of the poor fits was mis-attributed to a substandard secondary parametrisation.

Rectifying the bug fortunately produced little change in A_Γ for $\pi^+\pi^-$ $(0.80 \pm 1.08) \times 10^{-3}$ and none in K^+K^- $(-0.35 \pm 0.62) \times 10^{-3}$; $\pi^+\pi^-$ showed an alteration but it was within the statistical uncertainties and the physics conclusion was unaffected. However the fit quality was somewhat improved as a result,

Quantity	Bias
$\tau_{\pi^+\pi^-}$	-0.13
$\tau_{K^+K^-}$	-0.68
$\tau_{K^-\pi^+}$	-0.38
$y_{CP}(\pi^+\pi^-)$	-0.63
$y_{CP}(K^+K^-)$	+0.74

Table 6.2: The biases due to the correlation between Δm and decay time when using decay tree fitter to constrain the slow pion to the primary vertex. These numbers are the biases on the pull distributions.

particularly in the projection of the $ln(\chi^2_{IP})$ variable.

For the fits more freedom was given to the mis-reconstructed background components in the mass fit for y_{CP}. These are the backgrounds $D_s^+ \to K^+K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ present in the K^+K^- final state. For the former background the Δm shape was changed to be the same as the random slow pion, as its distribution should be comprised entirely of random pions. In mass the shape remained fixed and as before the fraction of the component in the fit was allowed to float.

For the $D^0 \to K^-\pi^+\pi^0$ background the shape of the mass component was again fixed. In Δm however the PDF was broken into a signal and random slow pion component. The shapes of each of these were fixed to the results of the fits to simulation as in A_Γ but the fraction between them could now float. Again the fraction of the background component in the data was allowed to float.

6.1.4 Decay Tree Fitter

As noted in section 5.1.5 the Decay Tree Fitter (DTF) [120] algorithm can be used to increase the resolution in the Δm variable. A correlation between decay time and Δm is however introduced by DTF for the secondary component of the data. In A_Γ this correlation was found to introduce no bias to the measurement. For y_{CP} as the comparison is now between different final states the potential bias was re-examined.

Simplified simulated data sets were generated with the secondary component having a time-dependent Δm description as shown in Fig. 5.2. Each final state was assessed separately with their own time-dependent parametrisations. The simulations were then fitted with the standard fit configuration (assuming t and Δm to be independent) and the bias compared to fits to simulations generated without a Δm time-dependence. The changes in the bias on the pulls are shown in table 6.2. As the biases on y_{CP} are so large it was decided to proceed without the use of decay tree fitter to constrain the slow pion to the primary vertex.

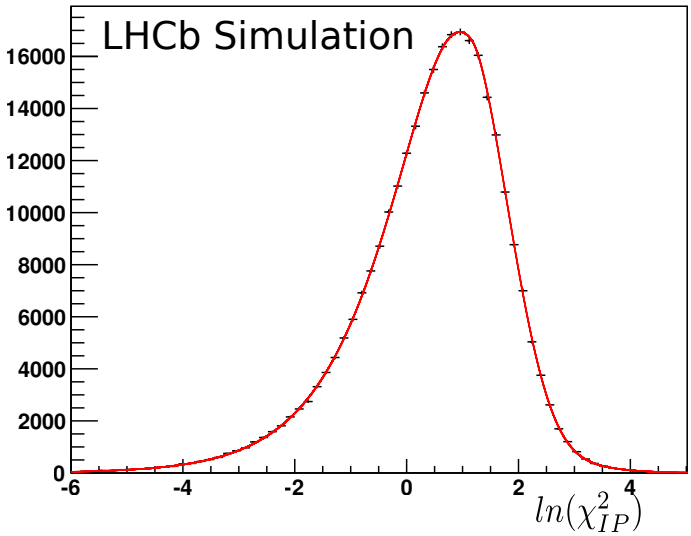


Figure 6.5: The distribution of $ln(\chi^2_{IP})$ for a simulated sample of prompt $D^0 \to K^-\pi^+$ decays. Overlaid is a fit of the function described in section 6.1.5.

6.1.5 The fit model

Some alterations were made to the fit model in order to produce better quality fits. Furthermore as the lifetimes of the different final states were now under direct comparison it was considered important that all three modes should have the same decay time fit model to avoid the choice of model biasing the result.

$ln(\chi^2_{IP})$ PDFs

The prompt $ln(\chi^2_{IP})$ PDF was updated to better describe the right hand side of the $ln(\chi^2_{IP})$ distribution seen in simulation in Fig. 6.5. The function used was the sum of a smeared $ln(\chi^2_{IP})$ PDF of equation 4.32 and a similar PDF but with a long tail on either side. This extra PDF has the form

$$P(y) = \begin{cases} \exp\{\alpha_L y - \exp\{\alpha_L(y-\mu)\}\} & y \leq \mu \\ A \cdot \exp\{-\alpha_R(y-\mu) - \exp\{\alpha_R(y-\mu)\}\} & y > \mu \end{cases}.$$

The fit to simulated prompt $K^-\pi^+$ decays is in Fig. 6.5. The two functions have the same left hand smearing parameters α_L and means μ, with independent right hand parameters, α_{R1} and α_{R2}. The smearing parameters are constant with decay time.

The secondary $ln(\chi^2_{IP})$ PDF was a single bifurcated Gaussian distribution. The time evolution of the width parameters σ_L and σ_R were studied again, this time with a larger sample of simulated secondary decays that had the analysis selection applied. Again the data were split into bins of decay time and the

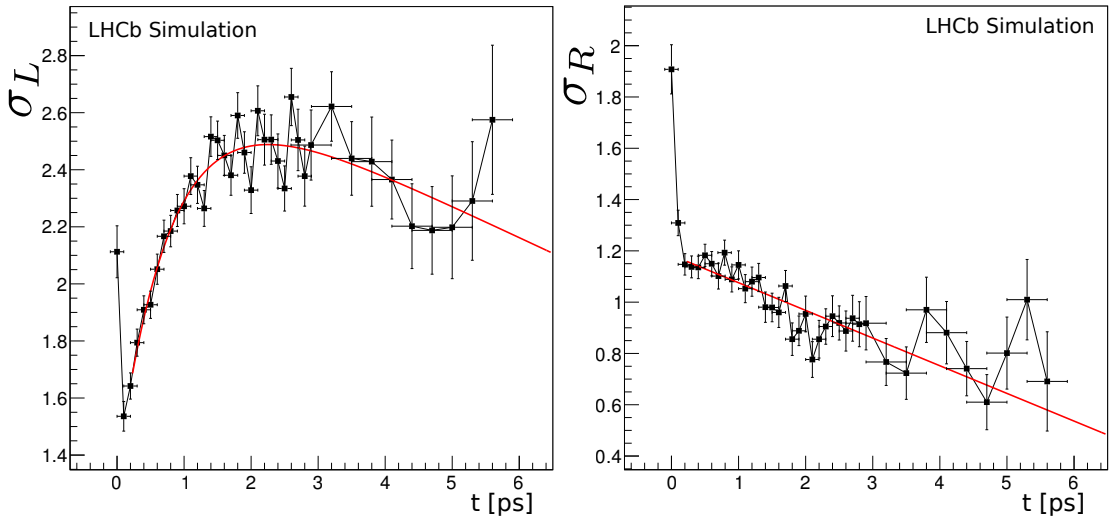


Figure 6.6: The time evolution of the secondary $ln(\chi^2_{IP})$ width parameters, extracted from simulation. Plotted are the width parameters extracted from fits to $ln(\chi^2_{IP})$ in bins of decay time. On the left is the σ_L parameter, on the right σ_R. Overlaid are the results of the fits. σ_L has been fitted with the function in equation 6.1, whilst σ_R has been fitted with a linear funciton.

$ln(\chi^2_{IP})$ distribution in each fitted. The fit parameters in each bin of time were extracted in order to assess their time evolutions. The two width parameters are shown in Fig. 6.6. The σ_L parameter is fitted with a function of the form

$$\sigma_L(t) = \sigma_L - A \cdot e^{-B \cdot t}. \tag{6.1}$$

σ_R is fitted with a linear function.

In the fit to collision data the exponent parameter B of the σ_L time evolution has been fixed to the result of the fit to simulation, $-1.49\,\mathrm{ps}^{-1}$. The other two parameters are left to float, as are both of the parameters of the linear function describing the time evolution of σ_R.

The time dependency of the prompt mean parameter was also revised. There is a cut on the angle between the momentum vector of the D^0 candidate and the displacement vector between the decay vertex and primary vertex (this quantity is referred to as the DIRA, for direction angle). It is designed to suppress secondary candidates which in general will have a large angle between the two vectors. The cut also introduces a time dependence on the mean of the prompt $ln(\chi^2_{IP})$ decays. At low decay times a D^0 candidate with large impact parameter (hence large $ln(\chi^2_{IP})$) would have a large angle between the momentum vector and displacement vector as shown in Fig. 6.7. At large decay times the angle can be small and still give a large $ln(\chi^2_{IP})$.

The time evolution for a sample of simulated prompt decays with the DIRA

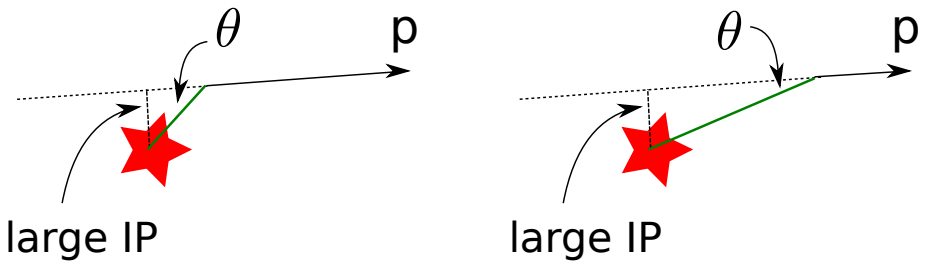


Figure 6.7: The time dependent effect of the DIRA cut. At low decay times (left) the candidate would fail the DIRA cut due to the large angle between the momentum vector, p, and the flight distance vector. At higher decay times this angle is smaller for the same impact parameter and the candidate would pass the cut.

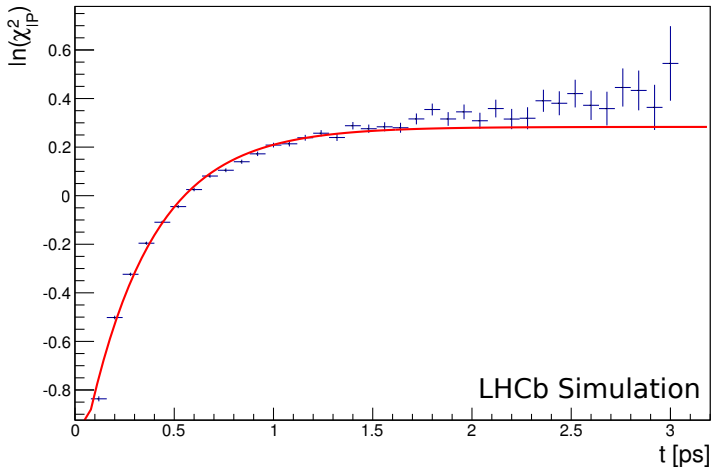


Figure 6.8: The effect of the DIRA cut on the $ln(\chi^2_{IP})$ mean. Plotted is the arithmetic mean of the distribution of the $ln(\chi^2_{IP})$ variable in bins of proper time. Overlaid is a fit of the form 6.1. The data is simulated prompt $D^0 \to K^-\pi^+$ decays.

cut alone applied (the sample is otherwise unbiased) is shown in Fig. 6.8 overlaid with a fit of a function of similar form to equation 6.1. In the full fits to data the B parameters have been fixed to the values obtained from the fits to the simulation in Fig. 6.8, although separate values were obtained for each of the three final states.

The secondary time evolution remained the arctan function as before and again the assumption was made that at $t = 0$ the prompt and secondary $ln(\chi^2_{IP})$ distributions should have the same means. This is important as it allows one to constrain the low lifetime behaviour of the secondary component.

Secondary lifetime PDF

In order to improve the fit quality the secondary description was examined more carefully. A sample of simulated $D^0 \to K^-\pi^+$ decays, with the D^0 originating from

a B decay was obtained. The complete trigger and offline selection was applied to produce a biased sample for comparison with the collision data. The swimming was subsequently run on the simulation to examine the acceptance function. Furthermore a comparison was made with unbiased simulation to ensure the swimming was functioning as expected for the secondary decays.

The unbiased and biased data can be seen in Fig. 6.9. Both have been overlaid with the results of the fit of the PDF

$$f(t) = \frac{\tau_1\tau_2}{\tau_1 - \tau_2}\left(e^{-t/\tau_1} - e^{-t/\tau_2}\right). \tag{6.2}$$

The fit to the biased sample has been done with the per-event acceptance accounted for. The two samples give very different fit results, as shown in table 6.3. It appears as if applying the swimming to the secondary decays does not reproduce the unbiased distribution. Furthermore, combining the unbiased PDF with the per-event acceptance given by the turning points produces a distribution rather different to the biased simulation.

Such a result was somewhat surprising as it suggests the swimming method does not work for the secondary component. This is perhaps to be expected as the momentum vector and the flight distance vector are not parallel for such decays. The swimming is performed along the momentum direction and the projection of the momentum along the flight distance is subsequently used to calculate the lifetime.

As well as the lifetimes the poor fit quality obtained using the PDF in equation 6.2 was noted. It was hypothesised that the poor description of the secondary lifetime was responsible for the fit quality. Adding a further exponential to give a PDF of the form

$$f(t) = f \cdot \frac{\tau_1\tau_2}{\tau_1 - \tau_2}\left(e^{-t/\tau_1} - e^{-t/\tau_2}\right) + (1-f)\cdot\frac{1}{\tau_3}e^{-t/\tau_3}, \tag{6.3}$$

markedly improves the fits to the simulations (again with a large discrepancy in the lifetimes between the biased and unbiased samples) as can be seen in Fig. 6.10. The fit results are shown in table 6.4.

This more complex PDF was used to fit the collision data. The secondary lifetime parameters tend to very small values when left to float, suggesting the fit is not sensitive the to the secondary component at low lifetimes. Therefore the parameters τ_1 and τ_2 were fixed to those obtained from the fit to simulation. A systematic uncertainty for the description of the secondary component was ascertained as described in section 6.4.3.

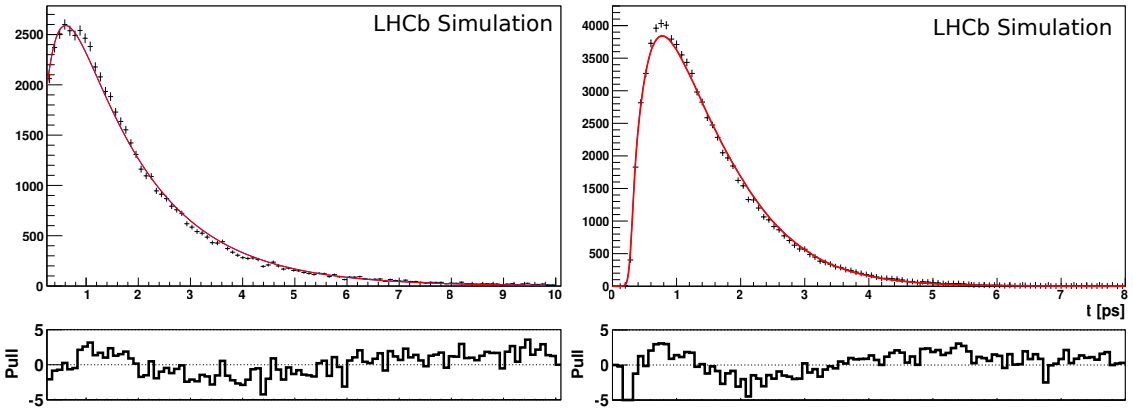


Figure 6.9: The decay time distribution of simulated $D^0 \to K^-\pi^+$ decays where the D^0 has originated from a B decay. On the left is the unbiased sample, on the right the sample with selections applied. Overlaid are fits of the secondary lifetime PDF in equation 6.2.

Parameter	Unbiased value	Biased value
τ_1 (fs)	317 ± 7	178 ± 5
τ_2 (fs)	1447 ± 6	771 ± 4

Table 6.3: The results of fits to simulated secondary decays using the PDF in equation 6.2. The fits to the unbiased and biased samples return very different results.

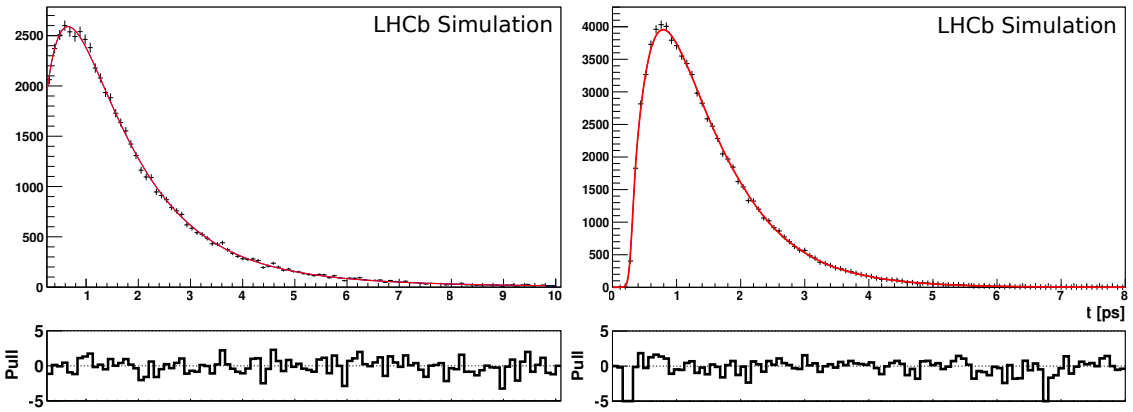


Figure 6.10: The decay time distribution of simulated $D^0 \to K^-\pi^+$ decays where the D^0 has originated from a B decay. On the left is the unbiased sample, on the right the sample with selections applied. Overlaid are fits of the secondary lifetime PDF in equation 6.3.

Parameter	Fit value
f	0.19 ± 0.04
τ_1 (fs)	265 ± 15
τ_2 (fs)	605 ± 25
τ_3 (fs)	1348 ± 89

Table 6.4: The results of a fit to simulated secondary decays using the PDF in equation 6.3.

Data split	$\tau_{K^-\pi^+}$ (fs)
MagUp pre-TS	407.49 ± 0.55
MagDown pre-TS	405.93 ± 0.44
MagUp pre-TS	406.93 ± 0.52
MagDown post-TS	405.80 ± 0.35
MagUp	406.55 ± 0.34
MagDown	406.13 ± 0.29
pre-TS	407.21 ± 0.38
post-TS	405.85 ± 0.27
Total	406.31 ± 0.22

Table 6.5: The measured $K\pi$ lifetimes for y_{CP}. The top four rows are averaged results for both flavours in each split by polarity and run period. The next four rows are the weighted averages for each run period or polarity. The final column is the total weighted average.

6.2 Fitter validation

The veracity of the fit results can be checked with the fits to data. In particular as the K^+K^- and $\pi^+\pi^-$ final states have the same blinding applied, agreement between the two is a powerful test of the fitter stability. Additionally the fits to the four individual samples within each final state (the two magnet polarities and two run periods) should agree with each other within their statistical uncertainties.

6.2.1 $K^-\pi^+$ lifetime

The measured $K^-\pi^+$ lifetimes are shown in table 6.5. The fits to the flavours have been averaged and are not shown here. As can be seen there is a difference of almost 1% between the period before the July technical stop (pre-TS) and that after it (post-TS) as was noted when measuring A_Γ (see section 5.4.1). The total averaged lifetime is in reasonable agreement with the world average with the post-TS being within the quoted uncertainty on the world average.

Data split	$y_{CP}(KK) \times 10^{-3}$	$y_{CP}(\pi\pi) \times 10^{-3}$
MagUp pre-TS	25.93 ± 4.28	27.19 ± 7.96
MagDown pre-TS	20.58 ± 3.52	16.43 ± 6.39
MagUp post-TS	15.40 ± 3.84	27.54 ± 7.23
MagDown post-TS	17.88 ± 2.66	21.55 ± 5.03
MagUp	22.70 ± 2.72	20.69 ± 4.98
MagDown	16.97 ± 2.19	23.41 ± 4.13
pre-TS	20.17 ± 2.86	27.43 ± 5.35
post-TS	16.97 ± 2.19	19.61 ± 3.95
Total	19.26 ± 1.70	22.32 ± 3.18

Table 6.6: The blinded values of y_{CP} for the two final states.

6.2.2 Final state agreement

The blinded values for y_{CP} are shown in table 6.6 and plotted in Fig. 6.11. As can be seen the results for the two modes agree with each other within the quoted statistical uncertainties.

6.3 Fit results

The projections of the data of the three decay modes in the four fitted variables (m_{D^0}, Δm, $ln(\chi^2_{IP})$ and t) and the turning point functions (TP_1 and TP_{diff}) with the fit results overlaid are shown in Fig. 6.12 ($\pi^+\pi^-$), Fig. 6.13 (K^+K^-) and Fig. 6.14. The data sample shown is that with the magnet polarity down, before the July technical stop. The two-dimensional pull distributions for the three final states are shown in Fig. 6.15.

6.4 Systematic uncertainties

For the most part the assessment of the systematic uncertainties for y_{CP} followed that of the A_Γ analysis. The reader may refer back to the relevant sections (5.4 and 4.2.11) for the details. Here only those systematics studies that changed compared to the A_Γ analysis will be described in detail. All the systematic uncertainties are summarised in table 6.7.

6.4.1 Turning point accuracy

The systematic uncertainty due to the accuracy of the turning points obtained by the swimming method had previously been assessed by applying either a scale to the turning points of up to 1% or an offset of up to 0.1 mm. The test that scaled

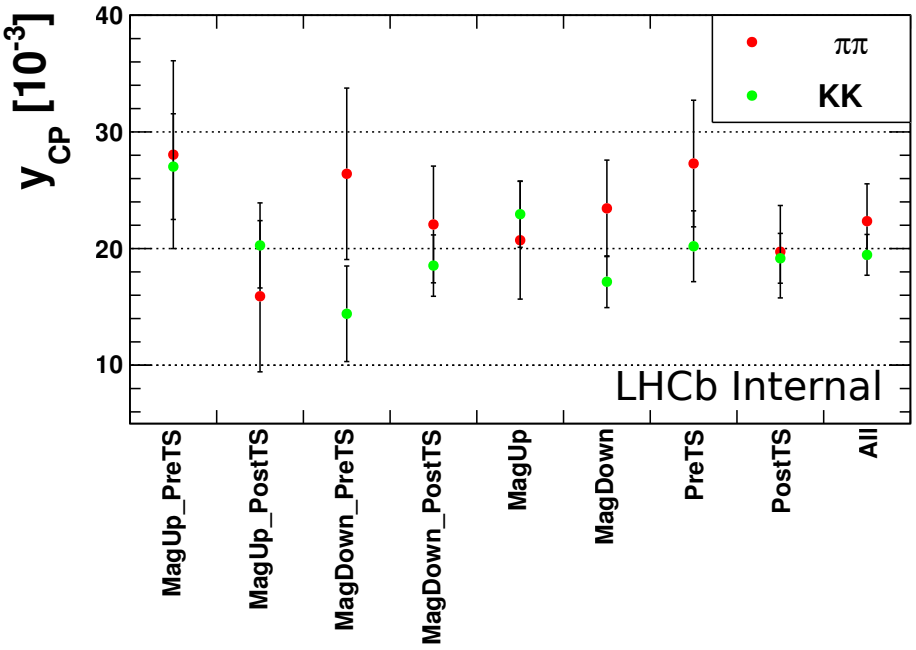


Figure 6.11: The blinded values of y_{CP} for the two final states, K^+K^- in green and $\pi^+\pi^-$ in red.

the turning points by up to 1% introduced a shift in TP_1 of between about 2.5 fs and 30 fs. The size of the offset was roughly four times the smallest search increment of the swimming algorithm and could cause a shift of well over 100 fs in the turning point.

The true turning point resolution can be estimated from data by considering the distribution of the difference between the first turning point and the measured decay time. This should be a step function at $t = 0$, with some spreading out due to the turning point resolution. The distribution of $t - TP_1$ is shown in Fig. 6.16 with the fit of a frequency function overlaid (the convolution of the step function and a Gaussian resolution function). The fitted width of the resolution is (0.17 ± 0.02) fs with a mean of (0.42 ± 0.01) fs.

The sizes of the variations of the scale and bias in the systematics tests had therefore been greatly overestimated. To make the assessment more commensurate with the resolution as measured with the data the ranges were altered. The scale was changed to be up to 0.01% and the bias to 5 µm, the expected resolution of the smallest increment in the search. Furthermore the changes were only applied in the positive direction (previously both positive and negative alterations had been used). The negative shift in the turning points increases the acceptance at the lowest lifetimes to be below where the cut was applied. As the data has already been taken this is unreasonable as the events cannot be regained so a bias would be introduced.

For both $\pi^+\pi^-$ and K^+K^- the test which introduced an offset to the turning

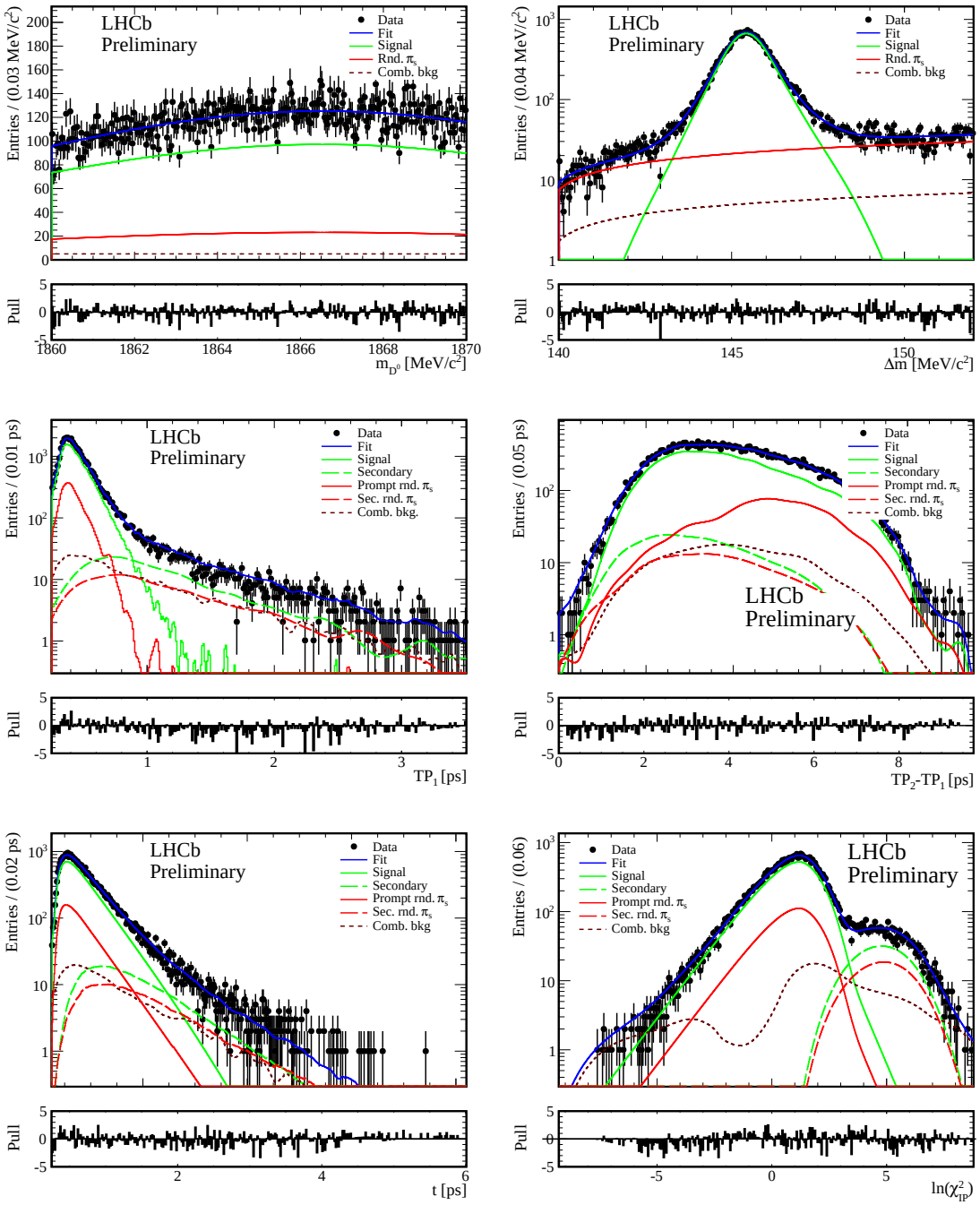


Figure 6.12: The projections of the data in m_{D^0} and Δm (top), TP_1 and TP_{diff} (middle) and $ln(\chi^2_{IP})$ and t (bottom), overlaid with the results of the fits, for the $\pi^+\pi^-$ final state.

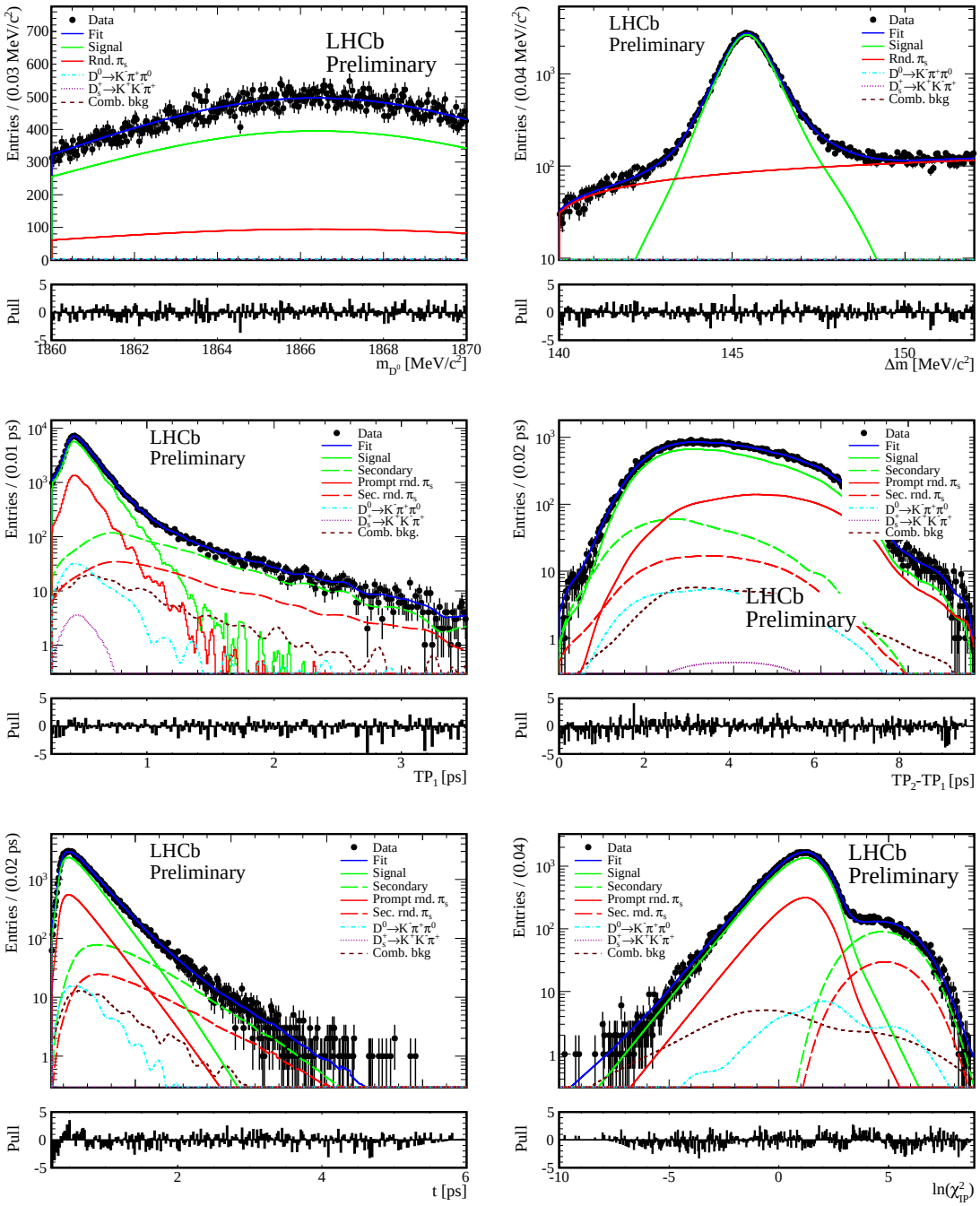


Figure 6.13: The projections of the data in m_{D^0} and Δm (top), TP_1 and TP_{diff} (middle) and $ln(\chi^2_{IP})$ and t (bottom), overlaid with the results of the fits, for the K^+K^- final state.

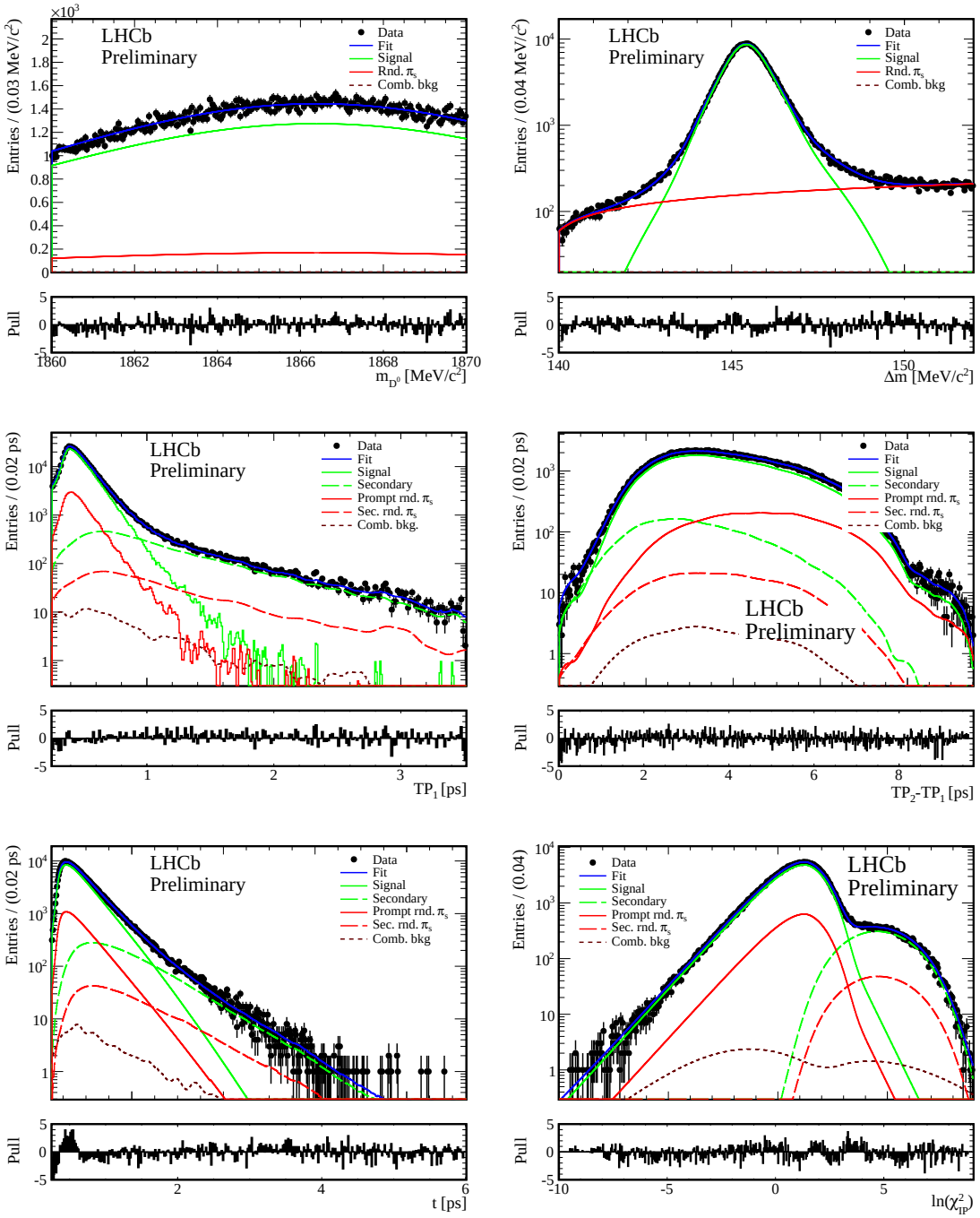


Figure 6.14: The projections of the data in m_{D^0} and Δm (top), TP_1 and TP_{diff} (middle) and $ln(\chi^2_{IP})$ and t (bottom), overlaid with the results of the fits, for the $K^-\pi^+$ final state.

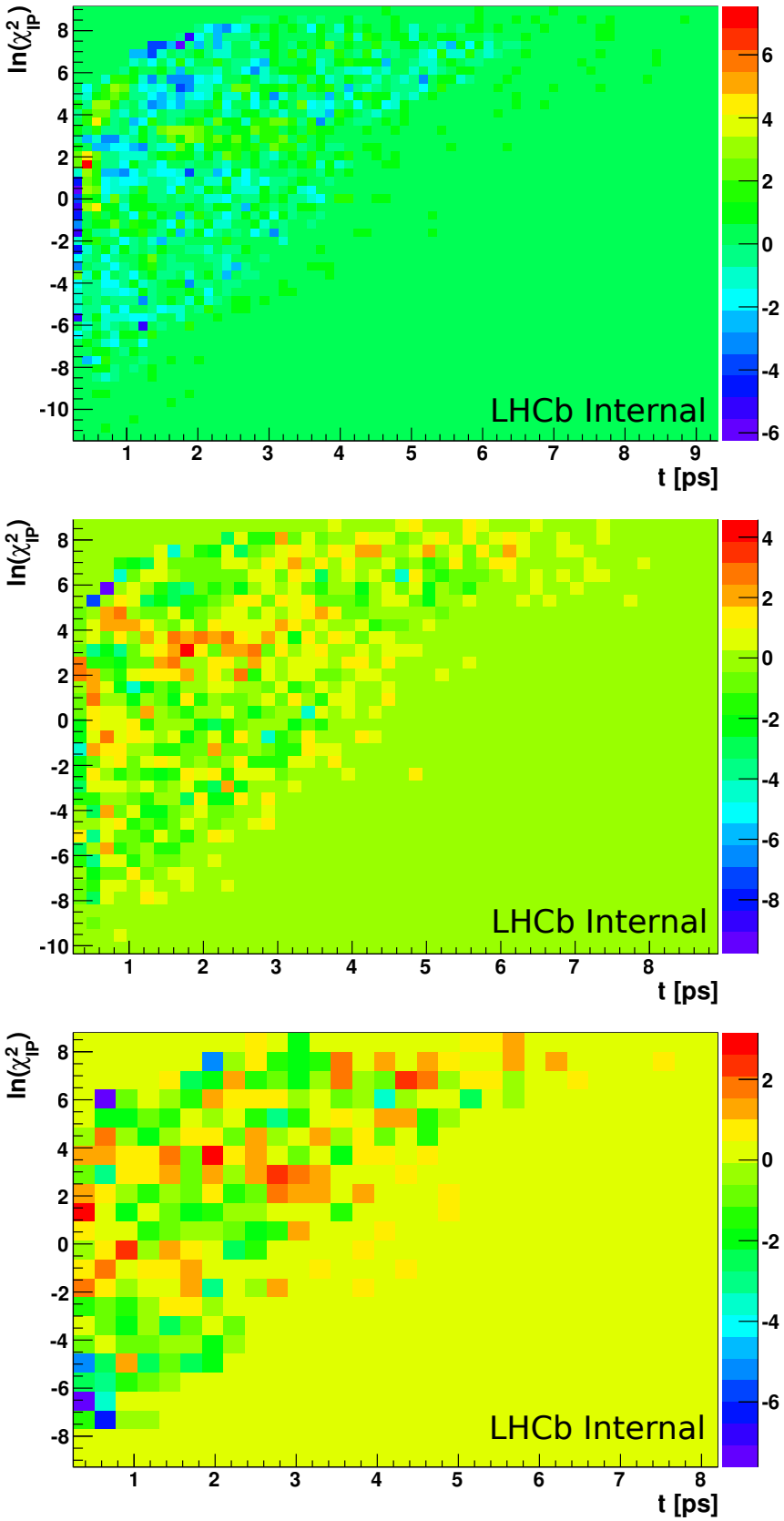


Figure 6.15: The two-dimensional pull distributions in the $ln(\chi^2_{IP})$ and t variables of the fits to the three decay modes: $K^-\pi^+$ (top), K^+K^- (middle) and $\pi^+\pi^-$ (bottom). The sample is the magnet down polarity, before the July technical stop.

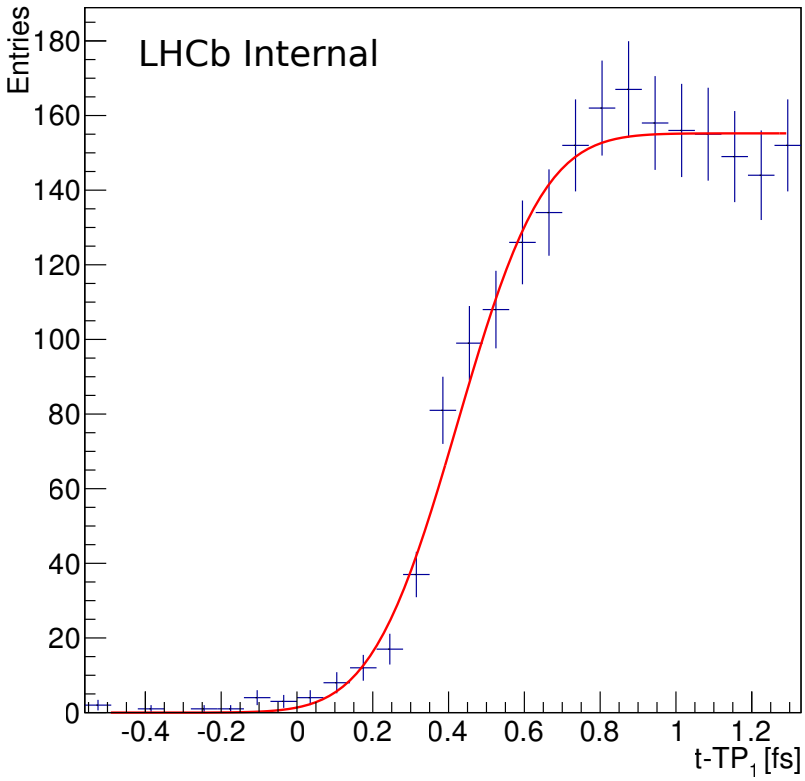


Figure 6.16: The distribution of $t - TP_1$, where t is the measured decay time and TP_1 is the first turning point. Overlaid is the fit of a frequency function to extract the turning point resolution.

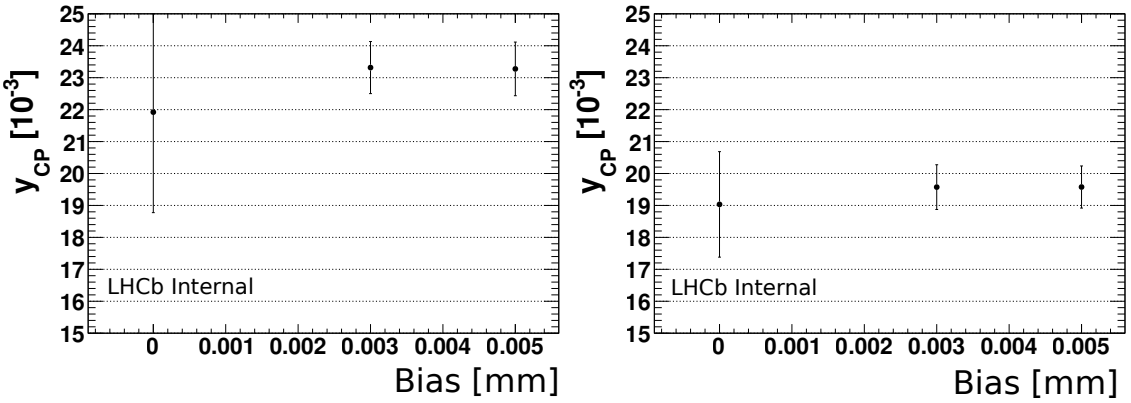


Figure 6.17: The variation of y_{CP} when offsetting the turning point distributions. The $\pi^+\pi^-$ mode is on the left, K^+K^- on the right.

points produced the biggest variation in the measured value of y_{CP}: 0.70×10^{-3} and 0.27×10^{-3} respectively. These values were therefore assigned as the uncertainties due to the accuracy of the turning points. Although the deviations in y_{CP} were all positive, so were the applied biases. It is not know what the result would be for negative biases so a symmetric uncertainty has been assumed. The variations of y_{CP} for the offset test are shown in Fig. 6.17. As the test allows only for a positive bias or scale, the resulting systematic uncertainty is assumer to be symmetric.

6.4.2 Mass-$ln(\chi^2_{IP})$ correlation

The uncertainty due to the remaining correlation between the D^0 mass and $ln(\chi^2_{IP})$ fit variables in the fitted mass bin is assessed by splitting the bin in two and comparing the standard y_{CP} result with the average from the two bins, the difference being the assigned error. For K^+K^- the difference is 0.21×10^{-3}, for $\pi^+\pi^-$ it is 0.28×10^{-3}.

It should be noted that this correlation will not be relevant in the future. The full analysis will account for correlations between the fit variables in a more comprehensive manner. An alternative option for assessing this systematic uncertainty would be a study with simplified simulations that closely represents the analysis method.

6.4.3 Secondary lifetime PDF

As noted in section 6.1.5 the secondary PDF is more complex than had previously been used in the A_Γ analysis. There are two fixed parameters in this PDF: τ_1 and τ_2 that have been ascertained from fits to simulation. To assess a systematic

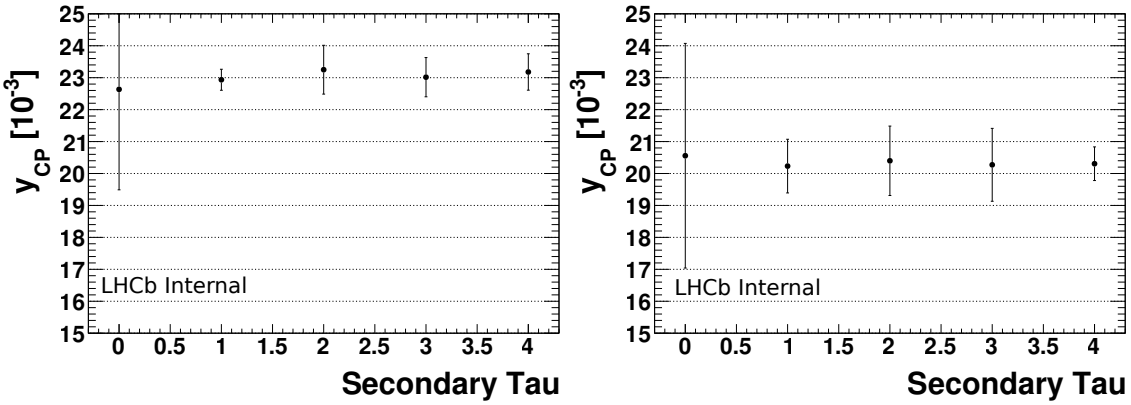


Figure 6.18: The variation of y_{CP} when varying the secondary lifetime PDF. The $\pi^+\pi^-$ mode is on the left, K^+K^- on the right. For K^+K^- the systematic was assessed with the magnet polarity up, post technical stop sample, shown here.

uncertainty due to the secondary component of the data the two fit variables are varied by twice the uncertainty of the result of the fit to simulation. Four tests are done; two where the variation is in the same direction and two where the parameters are varied in opposite direction.

The resulting uncertainties are 0.31×10^{-3} and 0.27×10^{-3} for $\pi^+\pi^-$ and K^+K^- respectively. The variations of y_{CP} are displayed in Fig. 6.18. The K^+K^- error was ascertained from a single sub-sample (magnet polarity up, post technical stop) due to the fits failing for the magnet down samples.

6.4.4 Combinatoric background

The uncertainty due to the combinatoric background has been assessed with two tests using simplified simulations. Both involved generating 330 toy datasets for each final state. The K^+K^- and $K^-\pi^+$ datasets had 4.8 million candidates each, with 1.5 million in the $\pi^+\pi^-$. The two tests are compared to a control sample whereby the generated data are fitted with the same fit model to ascertain the change in bias caused by the test.

The data are split into eight sub-samples (D^0 and $\overline{D}^0$, the two magnet polarities and the two run periods) which were fitted across the entire mass range (1815–1915 MeV). The results are combined and the pulls for the combination of fits for each dataset ascertained to gain the bias of the fits as a fraction of the statistical errors.

The first test changes the combinatoric mass model used in the generation from an exponential to a linear function. The data are subsequently fitted with the standard exponential function. For the K^+K^- mode the change in the bias from the standard fits is 5.4% of the statistical uncertainty, giving an absolute

systematic uncertainty of 0.09×10^{-3}. In $\pi^+\pi^-$ the change in bias is 4.3%, yielding a systematic uncertainty of 0.14×10^{-3}.

The second test varies the generated combinatoric time fit model from the standard lifetime extracted from the mean of the sWeights in the data. For K^+K^- the resulting change in the bias is 6.4%, with an absolute uncertainty of 0.11×10^{-3}. For $\pi^+\pi^-$ the bias change is 1.4% with the absolute uncertainty being 0.04×10^{-3}.

As the fits are done to the entire mass range these systematic uncertainties are likely an overestimate. When cutting tightly on mass around the signal peak one significantly increases the signal to background ratio, limiting the effect of the combinatoric backgrounds on the fit. For each mode the assigned systematic uncertainty is taken from the test that yields the larger result.

6.4.5 Reconstruction efficiency

The standard fits were done again with the correction of the signal lifetime PDF due to the time dependent reconstruction efficiency applied. These results were compared to the standard fits without the efficiency correction and half the difference taken as the systematic error. For K^+K^- the assigned uncertainty is $\pm 0.17 \times 10^{-3}$ and for $\pi^+\pi^-$ it is $\pm 0.09 \times 10^{-3}$.

6.4.6 Time resolution

The uncertainty due to the time resolution model was assessed using the test whereby the average error is varied between 0.03 ps and 0.1 ps. This test yielded errors of $\pm 0.22 \times 10^{-3}$ for both the K^+K^- and $\pi^+\pi^-$ modes.

6.4.7 Minimum time cut

The effect of the minimum decay time cut ($t = 0.25$ ps) is examined in the same fashion as was done for the A_Γ analysis (section 5.4.2), whereby the cut value is varied between 0.15 ps and 0.35 ps; four values were assessed. Again the variation of the measured y_{CP} should be statistical in nature and a χ^2 probability has been calculated to assess that hypothesis. In each case the two variations that had the largest difference to the standard fit were used to calculate the total χ^2 (giving two degrees of freedom).

In the K^+K^- mode the two differences between the tests and the standard fit were $(0.24 \pm 0.25) \times 10^{-3}$ and $(0.17 \pm 0.29) \times 10^{-3}$, leading to a χ^2 probability of agreement with the statistical variation of 53%. For $\pi^+\pi^-$ the the differences

Source	$y_{CP}(KK)$ $\times 10^{-3}$	$y_{CP}(\pi\pi)$ $\times 10^{-3}$
Mis-reconstructed bkg.	± 0.08	-
Secondary bkg.	± 0.23	± 0.31
Combinatoric bkg.	± 0.11	± 0.14
Turning points	± 0.27	± 0.70
Time resolution	± 0.22	± 0.22
Reconstruction eff.	± 0.17	± 0.09
Mass-$ln(\chi^2_{IP})$	± 0.21	± 0.28
Total systematic uncertainty	± 0.52	± 0.86
Total statistical uncertainty	± 1.65	± 3.15

Table 6.7: The assigned systematic uncertainties for each potential source along with the total systematic and statistical uncertainties.

were $(0.10 \pm 0.09) \times 10^{-3}$ and (0.33 ± 0.19) leading to a χ^2 probability of 12%. The variation for both final states is therefore consistent with being of statistical origin.

6.4.8 Summary of uncertainties

The assigned systematic uncertainties are summarised in table 6.7. The total systematic error is smaller than the statistical uncertainty for both modes so one should expect to obtain greater precision with higher statistics once the acceptance has been correctly accounted for in the remaining fraction of the data sample.

6.5 Results

The unblinded values of y_{CP} measured using the 2011 data set are

$$y_{CP}(KK) = (5.06 \pm 1.70 \pm 0.52) \times 10^{-3}$$
$$y_{CP}(\pi\pi) = (8.08 \pm 3.18 \pm 0.86) \times 10^{-3},$$

where the first error is statistical and the second systematic. The weighted average of the two modes is

$$y_{CP} = (5.61 \pm 1.56) \times 10^{-3}.$$

The two modes are consistent with each other within their statistical uncertainties. This represents a measurement of y_{CP} with significance of $3.6\,\sigma$ relative to zero. The precision of this measurement matches the current world average. The central value is consistent with the HFAG average [41] within 1.25 standard deviations.

Chapter 7

Conclusions

Measurements of *CP* violation and mixing in charm mesons have been described in this thesis, along with a study of the resolution of the LHCb vertex locator. The VELO work has been commented on extensively in section 3.6. The error parametrisation of the VELO has been extracted from data for the first time and a surprisingly large change in the resolution during Run 1 was observed. In this chapter the impact of the results given in chapters 5 and 6 on the overall picture of charm flavour physics will be described. Finally the prospects for these measurements in the future will be investigated.

7.1 Status of y_{CP}

At the time of writing, the analysis of y_{CP} described in chapter 6 is ongoing. The result presented in this document represent a fraction of the available data and so suffers from a significant statistical uncertainty. Much work is being carried out to overcome the complications outlined in the previous chapter.

In particular the correlation between the $ln(\chi^2_{IP})$ and D^0 mass variables is being much more rigorously treated. The signal and random slow-pion backgrounds now include a dependence on the mass that is modelled on data. Of course the sWeights that are used for the description of the mis-reconstructed and combinatoric backgrounds will be biased due to the correlation. Unfortunately there is no real alternative. One could take the combinatoric shape from the mass and Δm sidebands as this component of the fit does not exhibit the correlation. This is shown in Fig. 7.1 which shows the dependence of $ln(\chi^2_{IP})$ on mass for same-sign $K^+\pi^+$ data. This is data that has the same selections as the regular $D^0 \to K^-\pi^+$ mode but with the requirement that the decay products have the same charge.The statistics in the sidebands are low and there will be varying amounts of mis-reconstructed backgrounds that would not be properly represented in them. Furthermore there is a background that peaks in Δm in the K^+K^- final

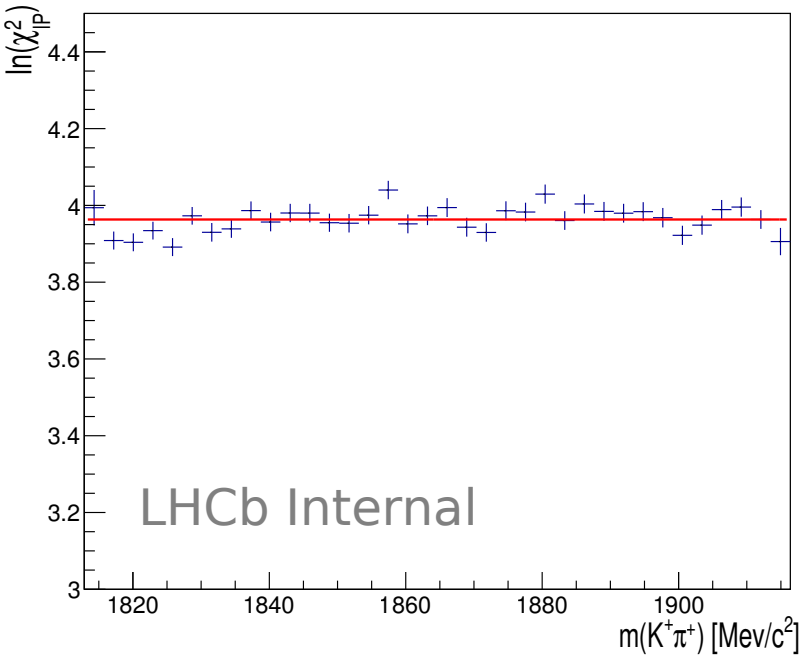


Figure 7.1: The dependence of $ln(\chi^2_{IP})$ on invariant mass for same-sign $K^+\pi^+$ data. Shown is the mean of the $ln(\chi^2_{IP})$ of the data in bins of mass, overlaid with a fit of a constant in red.

state that still needs to be dealt with. The approach is therefore to persist with using the sWeights to describe the background shapes and assign a corresponding systematic uncertainty using simplified simulations. As the backgrounds form a relatively small fraction of the data then this uncertainty should be small.

The effect of the step in the acceptance function, described in section 6.1.1 has also been more closely investigated. The tracking efficiency can be evaluated using the swimming information. At each trialled swum point information is gained as to whether the track was reconstructed in the trigger as well as whether it subsequently passed. Therefore one can impose the requirement that both tracks were reconstructed in the trigger instead of requiring them to fire it. This essentially ensures that the reconstruction efficiency of the selected sample is always one with respect to the offline reconstruction. This approach is now being used for the y_{CP} analysis.

The information from the swimming can also be used to assess the online reconstruction efficiency of the sample with respect to the offline reconstruction. Fig. 7.2 shows the efficiency for reconstructing the tracks of the decay products for the K^+K^- and $\pi^+\pi^-$ final states. As can be seen it is very high at LHCb, being over 99% for both modes. The efficiency relative to the offline reconstruction also appears to be independent of decay time (although the absolute efficiency is still time-dependent as described in section 4.2.6).

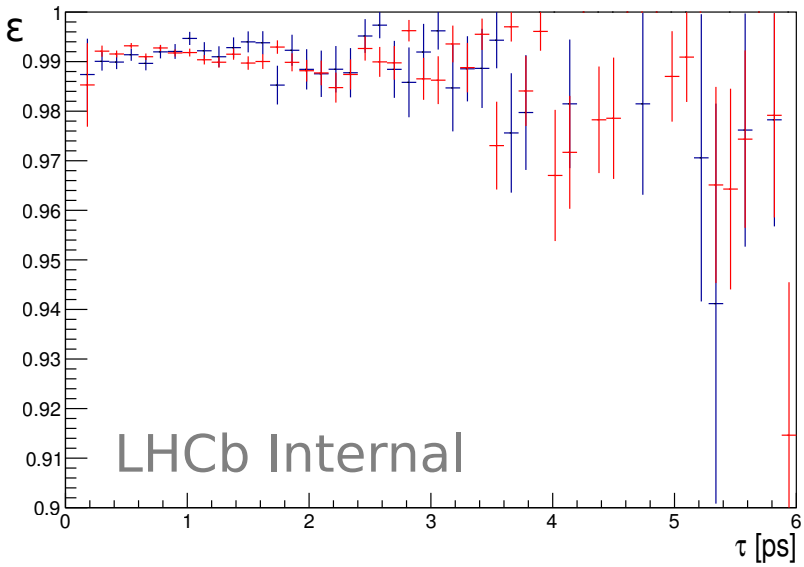


Figure 7.2: The efficiency of reconstructing tracks in the HLT relative to the offline reconstruction for the K^+K^- (red) and $\pi^+\pi^-$ (blue) final states. The efficiency is plotted in bins of decay time.

7.2 Impact of the measurements

The result presented in sections 5.5 showed a considerable improvement in precision over the previous world average for A_Γ. The y_{CP} measurement in section 6.5 is of similar precision to the current world average. Together with several other measurements made at LHCb during Run 1 (particularly the wrong-sign mixing analysis [38, 61], the measurements of ΔA_{CP} [56, 55] and the semi-leptonic A_Γ analysis [122]) these results have considerably tightened the constraints on the sizes of the indirect and direct charm *CP* violation parameters.

The change in the constraint on $\left|\frac{q}{p}\right|$ and ϕ (the argument of q and p) can be seen in Fig. 7.3. The point for no *CP* violation in mixing is within the $1\,\sigma$ contour of the fit. The current average of direct and indirect *CP* violation is in Fig. 7.4. The no *CP* violation point is now a little over $2\,\sigma$ from the world average, consistent with *CP* symmetry. The measurement of A_Γ with prompt $D^0 \to K^+K^-$ decays, presented in this thesis is the most precise result thus far. Together these LHCb A_Γ measurements have tightly constrained the possible value of indirect *CP* violation.

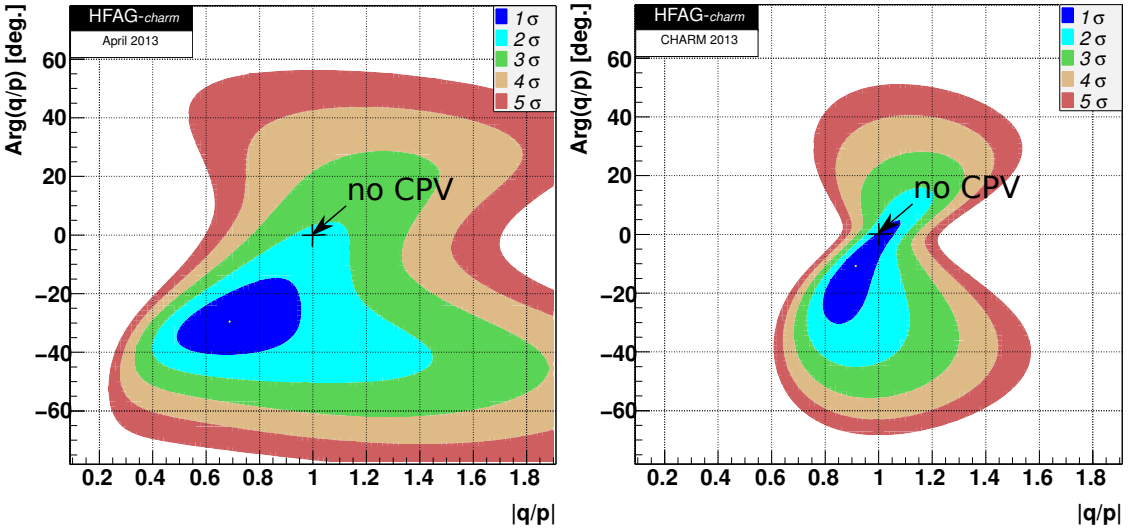


Figure 7.3: The change in the constraints on the magnitude and phase of $\frac{q}{p}$. On the left is the HFAG fit from April 2013, on the right the same fit in September 2013. The update in the latter fit is the inclusion of the the A_Γ result [119] presented in this thesis and the wrong-sign analysis [38]. From [41].

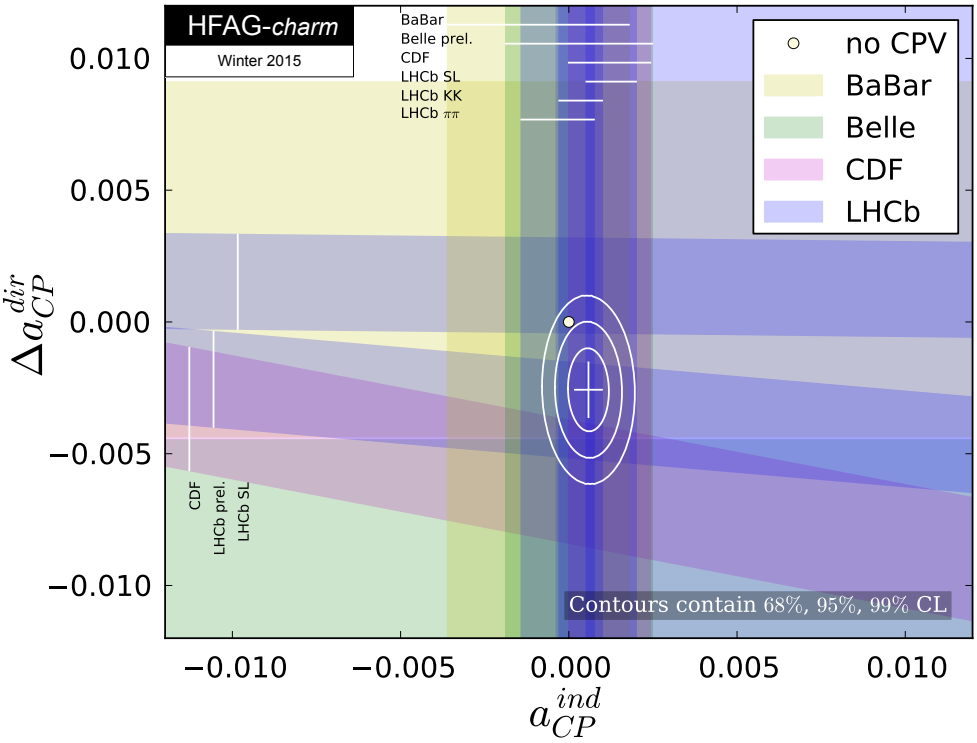


Figure 7.4: The HGAF fit of direct and indirect charm *CP* violation in winter 2015. From [41].

7.3 Future prospects

7.3.1 A_Γ and y_{CP}

In the immediate future the acceptance effect discussed in section 6.1.1 will be properly accounted for in the fit method. This should triple the statistics analysed for y_{CP} in the 2011 data set. Subsequently both A_Γ and y_{CP} will be measured with the data taken by LHCb in 2012 [123]. That data set is roughly double the size of that used in the analyses with the 2011 data and as such the results should reflect that increased precision. The method for these updated measurements will be the same as for the analyses described in this document.

A considerable challenge of the current analysis technique has been to effectively deal with the time-dependent acceptance introduced by the selections. The swimming method is effective but presents many challenges, in particular unfolding the turning point distributions for each component of the data. The acceptance has brought about extra complexities when dealing with the secondary background component. There is so little data at low lifetimes that it is very difficult to effectively discriminate the prompt and secondary candidates in this region. The large systematic uncertainty in y_{CP} for the secondary component reflects this.

The method is being reconsidered for Run 2 to simplify this, profiting from improvements made to the trigger. Whereas previously only individual tracks were reconstructed in the first stage of the trigger, leading to a complex acceptance, for the next period of data taking the full reconstruction is being run in HLT1 [124]. This means that vertices will be fitted and final state particles combined to reconstruct decay candidates before any lifetime biasing selections need to be applied. It is therefore expected that the sample of charm candidates produced by this will have a much simpler acceptance. The effects due to the physical dimensions of the VELO, tracking efficiency, clone killing etc. still exist but they are much more readily dealt with than those of the trigger selection.

If the result of the trigger improvements is a trivially biased data set the lifetime analyses can be greatly simplified and a potentially large systematic uncertainty removed. The swimming algorithm (a computationally intensive process) will not be needed as there will be no per-event acceptance to account for. As well as saving computing time this will allow the analysts to make changes to the reconstruction of the data (for example refitting the primary vertex without the signal candidate tracks) quickly. Furthermore the fit to the lifetime distribution will be drastically simplified and sped up; one may not need to calculate an integral for the event without the acceptance if the efficiencies are dealt with as an average over the whole dataset.

Furthermore a simpler acceptance function will allow alternative fit methods to be used. In the future with the greatly increased delivered luminosities a binned fit to the decay time could be performed with equivalent precision to the current unbinned maximum likelihood fit, again increasing the speed of the analysis. The ratio method, introduced in section 5.4.1, which looks at the yields of signal candidates in bins of decay time could hitherto only be used for A_Γ where the identical final states means that the acceptance cancels. If the biasing effects of the acceptance are readily dealt with and similar between the final states the ratio method can be readily adapted for a measurement of y_{CP}, providing a useful cross-check to an analysis measuring the effective lifetimes.

7.3.2 Charm physics

At the end of Run 1 of the LHC *CP* violation in charm has yet to be unambiguously observed. Mixing has now been observed in individual measurements although some work remains, such as a $5\,\sigma$ determination of the parameter x to assess which of the two mass states is heavier. One may ask at what point one would expect to make a discovery of *CP* violation in charm and how that might come about.

For LHCb the trigger improvements mentioned in the previous section should help to greatly increase the size of the sample of charm decays collected in Run 2. During that time it is hoped that LHCb will have recorded an integrated luminosity of around $5\,\text{fb}^{-1}$ [125]. During the next long shutdown (LS2) of the LHC the detector will be upgraded and by the end of 2030 have recorded in total $50\,\text{fb}^{-1}$. The estimated precision for the measurement of A_Γ by that point is 0.07×10^{-3}[126]. Assuming a value of the order of 10^{-4} for the Standard Model A_Γ expectation then one would expect to see signs of *CP* violation by the end of the upgrade, assuming the systematic uncertainties are likewise reduced. Benefits to the precision come from several sources beyond simply the delivered luminosity. The charm production cross-section in the LHCb acceptance will increase with collision energy augmenting the yield. One would also expect an increase in the performance of the trigger in Run 2 as noted previously. For the upgrade further improvements are planned, in particular the removal of the level-0 hardware trigger and the reading out of the whole detector at the nominal bunch-crossing frequency of 40 MHz. The level-0 trigger is inefficient for hadronic charm due to the selection of tracks with high transverse momentum. Therefore doing away with it should improve the efficiency considerably.

As noted in section 1.3.4 the *CP* violation parameters are obtained by several complementary measurements. An estimate of the sensitivity of LHCb

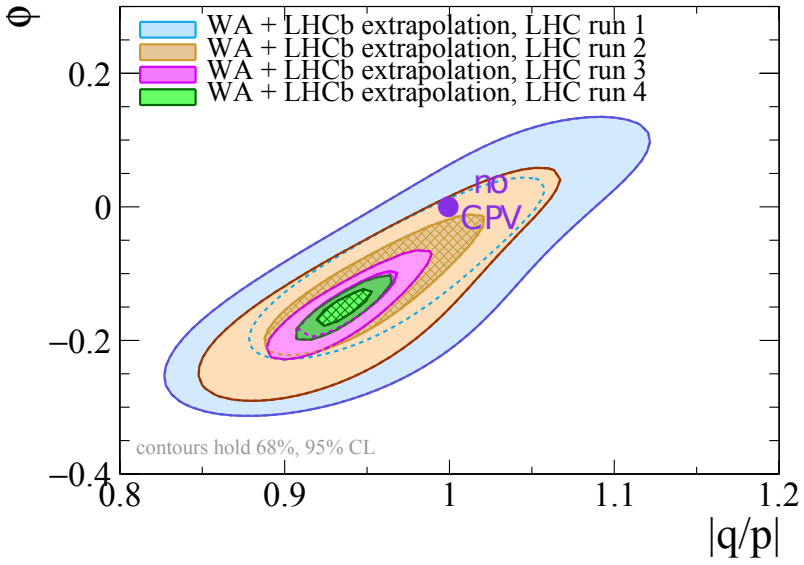


Figure 7.5: The expected sensitivity of measurements of the *CP* violation parameters q and p by the end of the LHCb upgrade. The estimate includes contributions from y_{CP}, A_Γ the wrong-sign mixing analysis (see section 1.3.4 and an analysis of $D^0 \to K_s\pi^+\pi^-$). From [125].

measurements to *CP* violation in mixing (the parameters ϕ and $\left|\frac{q}{p}\right|$) is shown in Fig. 7.5, taken from [125]. The projection assumes the current world averages for the central point and estimates an increase in production cross section with collision energy and selection efficiency with the upgrade. As can be seen by the end of Run 4 the combination of the measurements should be sensitive to *CP* violation in mixing.

As well as LHCb other experiments are expected to make a contribution to charm measurements. In particular Belle II [127] is expected to start taking physics data in 2018 with an integrated luminosity of $50\,\text{ab}^{-1}$ collected by 2024. With that data set the expected uncertainties on y_{CP} and A_Γ measured at Belle II would be 0.05 % and 0.04 % respectively [128], roughly equivalent to that achieved by LHCb at the end of Run 1. Having another experiment with comparable statistical reach to LHCb is important to verify measurements, particularly as the required precision will be very high.

With the start of Run 2 LHCb is well placed to drive down the statistical uncertainties of several measurements. In particular the new data will help to clear up the situation with regards ΔA_{CP} and the interplay of direct and indirect *CP* violation in charm. Eventually *CP* violation may finally be observed in the charm sector.

The implications for finding a significant amount of *CP* violation in the charm sector are hard to discern. As explained in section 1.3.3 the expectations of

what to expect in charm are somewhat unclear. Certainly if *CP* violation were discovered at the current experimental precision ($\sim 10^{-3}$) it would be a strong hint of new physics at play. It would however need to be seen in more than one observable to be able to characterise its nature.

Bibliography

[1] 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) 1–29, `arXiv:1207.7214`.

[2] 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) 30–61, `arXiv:1207.7235`.

[3] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (Nov, 1964) 585–587.

[5] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509.

[6] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (Oct, 1974) 2445–2459.

[7] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys.Rev.Lett. **30** (1973) 1343–1346.

[8] D. Gross and F. Wilczek, *Asymptotically Free Gauge Theories. 1*, Phys.Rev. **D8** (1973) 3633–3652.

[9] D. Gross and F. Wilczek, *Asymptotically Free Gauge Theories. 2.*, Phys.Rev. **D9** (1974) 980–993.

[10] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental test of parity conservation in Beta decay*, Phys. Rev. **105** (Feb, 1957) 1413–1415.

[11] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (Jul, 1964) 138–140.

[12] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652–657.

[13] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (Oct, 1970) 1285–1292.

[14] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (Nov, 1983) 1945–1947.

[15] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (Jun, 1963) 531–533.

[16] A. Abashian et al., *The Belle Detector*, Nucl. Instrum. Meth. **A479** (2002) 117–232.

[17] BaBar Collaboration, D. Boutigny et al., *The BABAR physics book: Physics at an asymmetric B factory*, in *Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting) Pasadena, California, September 22-24, 1997*, 1998.

[18] 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) 1–131, arXiv:hep-ph/0406184.

[19] R. P. Feynman and M. Gell-Mann, *Theory of the Fermi interaction*, Phys. Rev. **109** (Jan, 1958) 193–198.

[20] R. Garwin, L. Lederman, and M. Weinrich, *Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: The Magnetic Moment of the Free Muon*, Phys.Rev. **105** (1957) 1415–1417.

[21] A. Alavi-Harati et al., *Observation of direct CP violation in $K_{S,L} \to \pi\pi$ decays*, Phys. Rev. Lett. **83** (Jul, 1999) 22–27.

[22] V. Fanti et al., *A new measurement of direct CP violation in two pion decays of the neutral kaon*, Physics Letters B **465** (1999), no. 14 335 – 348.

[23] Belle Collaboration, K. Abe et al., *Observation of large CP violation in the neutral B meson system*, Phys.Rev.Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.

[24] BaBar Collaboration, B. Aubert et al., *Observation of CP violation in the B^0 meson system*, Phys.Rev.Lett. **87** (2001) 091801, arXiv:hep-ex/0107013.

[25] A. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh.Eksp.Teor.Fiz. **5** (1967) 32–35.

[26] C. A. Baker et al., *Improved experimental limit on the electric dipole moment of the neutron*, Phys. Rev. Lett. **97** (Sep, 2006) 131801.

[27] R. D. Peccei and H. R. Quinn, *cp conservation in the presence of pseudoparticles*, Phys. Rev. Lett. **38** (Jun, 1977) 1440–1443.

[28] R. Ballou et al., *Latest Results of the OSQAR Photon Regeneration Experiment for Axion-Like Particle Search*, `arXiv:1410.2566`.

[29] K. Barth et al., *CAST constraints on the axion-electron coupling*, Journal of Cosmology and Astroparticle Physics **2013** (2013), no. 05 010.

[30] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (Aug, 1998) 1562–1567.

[31] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog.Theor.Phys. **28** (1962) 870–880.

[32] J. E. Augustin et al., *Discovery of a narrow resonance in e^+e^- annihilation*, Phys. Rev. Lett. **33** (Dec, 1974) 1406–1408.

[33] J. J. Aubert et al., *Experimental observation of a heavy particle J*, Phys. Rev. Lett. **33** (Dec, 1974) 1404–1406.

[34] BaBar Collaboration, B. Aubert et al., *Observation of a narrow meson decaying to $D_s^+\pi^0$ at a mass of 2.32-GeV/c^2*, Phys.Rev.Lett. **90** (2003) 242001, `arXiv:hep-ex/0304021`.

[35] CLEO Collaboration, D. Besson et al., *Observation of a narrow resonance of mass 2.46-GeV/c^2 decaying to $D_s^{*+}\pi^0$ and confirmation of the $D_{sJ}^*(2317)$ state*, Phys.Rev. **D68** (2003) 032002, `arXiv:hep-ex/0305100`.

[36] LHCb Collaboration, R. Aaij et al., *Observation of the resonant character of the $Z(4430)^-$ state*, Phys.Rev.Lett. **112** (2014), no. 22 222002, `arXiv:1404.1903`.

[37] LHCb Collaboration, R. Aaij et al., *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.0341`.

[38] LHCb Collaboration, R. Aaij et al., *Observation of $D^0-\bar{D}^0$ oscillations*, Phys. Rev. Lett. **110** (Mar, 2013) 101802.

[39] M. Gersabeck, *Brief Review of Charm Physics*, Mod.Phys.Lett. **A27** (2012) 1230026, arXiv:1207.2195.

[40] M. Gersabeck, M. Alexander, S. Borghi, V. Gligorov, and C. Parkes, *On the interplay of direct and indirect CP violation in the charm sector*, J.Phys. **G39** (2012) 045005, arXiv:1111.6515.

[41] Y. A. et al., *Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012*, online update at http://www.slac.stanford.edu/xorg/hfag arXiv:1207.1158.

[42] BaBar Collaboration, P. del Amo Sanchez et al., *Measurement of $D^0 - \bar{D}^0$ mixing parameters using $D^0 \to K^0_S\pi^+\pi^-$ and $D^0 \to K^0_S K^+ K^-$ decays*, Phys. Rev. Lett. **105** (2010) 081803, arXiv:1004.5053.

[43] Belle Collaboration, T. Peng et al., *Measurement of $D^0 - \bar{D}^0$ mixing and search for indirect CP violation using $D^0 \to K^0_S\pi^+\pi^-$ decays*, Phys. Rev. **D89** (2014), no. 9 091103, arXiv:1404.2412.

[44] E. Golowich, *Charm mixing: Theory*, eConf **C070805** (2007) 05, arXiv:0710.0151.

[45] A. A. Petrov, *Charm physics: Theoretical review*, eConf **C030603** (2003) MEC05, arXiv:hep-ph/0311371.

[46] T. A. Kaeding, *D meson mixing in broken SU(3)*, Phys.Lett. **B357** (1995) 151–155, arXiv:hep-ph/9505393.

[47] P. Colangelo, G. Nardulli, and N. Paver, *On $D^0 - \bar{D}^0$ Mixing in the Standard Model*, Phys.Lett. **B242** (1990) 71.

[48] M. Neubert, *B decays and the heavy quark expansion*, Adv.Ser.Direct.High Energy Phys. **15** (1998) 239–293, arXiv:hep-ph/9702375.

[49] A. Lenz and M. Bobrowski, *Standard Model Predictions for D^0-oscillations and CP-violation*, Int.J.Mod.Phys.Conf.Ser. **02** (2011) 117–121, arXiv:1011.5608.

[50] LHCb Collaboration, R. Aaij et al., *Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates*, Phys. Rev. Lett. **108** (Mar, 2012) 111602.

[51] CDF Collaboration, T. Aaltonen et al., *Measurement of the difference of CP–violating asymmetries in $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays at CDF*, Phys.Rev.Lett. **109** (2012) 111801, arXiv:1207.2158.

[52] Belle Collaboration, B. Ko, *Direct CP violation in charm at Belle*, PoS **ICHEP2012** (2013) 353, arXiv:1212.1975.

[53] E. Franco, S. Mishima, and L. Silvestrini, *The Standard Model confronts CP violation in $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$*, JHEP **1205** (2012) 140, arXiv:1203.3131.

[54] T. Feldmann, S. Nandi, and A. Soni, *Repercussions of Flavour Symmetry Breaking on CP Violation in D-Meson Decays*, JHEP **1206** (2012) 007, arXiv:1202.3795.

[55] LHCb Collaboration, *A search for time-integrated CP violation in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, . Linked to LHCb-ANA-2012-034.

[56] LHCb Collaboration, R. Aaij et al., *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, JHEP **1407** (2014) 041, arXiv:1405.2797.

[57] LHCb Collaboration, 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 $B^0 \to J/\psi K^{*0}$*, Physics Letters B **719** (2013), no. 45 318 – 325.

[58] Belle Collaboration, M. Starič et al., *Evidence for D^0-$\bar{D}^0$ mixing*, Phys. Rev. Lett. **98** (May, 2007) 211803.

[59] BaBar Collaboration, B. Aubert et al., *Evidence for D^0-$\bar{D}^0$ mixing*, Phys. Rev. Lett. **98** (May, 2007) 211802.

[60] CDF Collaboration, T. Aaltonen et al., *Evidence for $D^0 - \bar{D}^0$ mixing using the CDF II detector*, Phys. Rev. Lett. **100** (Mar, 2008) 121802.

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

[62] Belle Collaboration, M. Staric, *New Belle results on $D^0 - \bar{D}^0$ mixing*, arXiv:1212.3478.

[63] BaBar Collaboration, J. Lees et al., *Measurement of $D^0 - \bar{D}^0$ Mixing and CP Violation in Two-Body D^0 Decays*, Phys.Rev. **D87** (2013) 012004, arXiv:1209.3896.

[64] BaBar Collaboration, B. Aubert et al., *Search for $D^0 - \bar{D}^0$ mixing using doubly flavor tagged semileptonic decay modes*, Phys. Rev. D **76** (Jul, 2007) 014018.

[65] Belle Collaboration, U. Bitenc et al., *Improved search for D^0-$\bar{D}^0$ mixing using semileptonic decays at Belle*, Phys. Rev. D **77** (Jun, 2008) 112003.

[66] M. Gronau, Y. Grossman, and J. L. Rosner, *Measuring $D^0 - \bar{D}^0$ mixing and relative strong phases at a charm factory*, Phys.Lett. **B508** (2001) 37–43, arXiv:hep-ph/0103110.

[67] BESIII Collaboration, M. Ablikim et al., *Measurement of the $D \to K^-\pi^+$ strong phase difference in $\psi(3770) \to D^0\bar{D}^0$*, Phys.Lett. **B734** (2014) 227, arXiv:1404.4691.

[68] CLEO Collaboration, D. Asner et al., *Updated Measurement of the Strong Phase in $D^0 \to K^+\pi^-$ Decay Using Quantum Correlations in $e^+e^- \to D^0\bar{D}^0$ at CLEO*, Phys.Rev. **D86** (2012) 112001, arXiv:1210.0939.

[69] Y. Grossman, A. L. Kagan, and Y. Nir, *New physics and CP violation in singly cabibbo suppressed d decays*, Phys. Rev. D **75** (Feb, 2007) 036008.

[70] LHCb Collaboration, R. Aaij et al., *Model-independent search for CP violation in $D^0 \to K^-K^+\pi^-\pi^+$ and $D^0 \to \pi^-\pi^+\pi^+\pi^-$ decays*, Phys.Lett. **B726** (2013) 623–633, arXiv:1308.3189.

[71] LHCb Collaboration, R. Aaij et al., *Search for CP violation in $D^0 \to \pi^-\pi^+\pi^0$ decays with the energy test*, Phys.Lett. **B740** (2015) 158, arXiv:1410.4170.

[72] M. Alexander et al., *Measurement of the charm mixing and CP violation parameters y_{CP} and A_Γ with 2011 data*, LHCb-ANA-2012-039.

[73] S. Malde, C. Thomas, and G. Wilkinson, *Measuring CP violation and mixing in charm with inclusive self-conjugate multibody decay modes*, Phys. Rev. **D91** (2015), no. 9 094032, arXiv:1502.0456.

[74] M. Nayak et al., *First determination of the CP content of $D \to \pi^+\pi^-\pi^0$ and $D \to K^+K^-\pi^0$*, Phys. Lett. **B740** (2015) 1–7, arXiv:1410.3964.

[75] LHCb Collaboration, R. Aaij et al., *Measurement of mixing and CP violation parameters in two-body charm decays*, JHEP **1204** (2012) 129, arXiv:1112.4698.

[76] *Design Report Tevatron 1 project*, Tech. Rep. FERMILAB-DESIGN-1984-01, 1984.

[77] LHCb VELO Group, R. Aaij et al., *Performance of the LHCb Vertex Locator*, J. Instrum. **9** (May, 2014) P09007. 61 p. Comments: 61 pages, 33 figures.

[78] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, `arXiv:1410.3667`.

[79] LHCb Collaboration, A. A. Alves et al., *The LHCb Detector at the LHC*, J. Instrum. **3** (2008), no. LHCb-DP-2008-001. CERN-LHCb-DP-2008-001 S08005. Also published by CERN Geneva in 2010.

[80] CERN, *CERN FAQ: LHC - the guide*. 2009.

[81] LHCb Collaboration, B. Adeva et al., *Road map for selected key measurements from LHCb*, Tech. Rep. LHCb-PUB-2009-029. CERN-LHCb-PUB-2009-029, CERN, Geneva, May, 2010. Comments: 379 pages.

[82] LHCb Collaboration, *Improved constraints on γ: CKM2014 update*, in *8th International Workshop on the CKM Unitarity Triangle (CKM 2014) Vienna, Austria, September 8-12, 2014*, 2014.

[83] LHCb Collaboration, R. Aaij et al., *Search for direct CP violation in $D^0 \to h^- h^+$ modes using semileptonic B decays*, Phys. Lett. B **723** (Mar, 2013) 33–43. 21 p. Comments: 16 pages, 5 figures. Matches the published version in PLB.

[84] CMS, LHCb, V. Khachatryan et al., *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522** (2015) 68–72, `arXiv:1411.4413`.

[85] LHCb Collaboration, R. Aaij et al., *Updated measurements of exclusive J/ψ and $\psi(2S)$ production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV*, J. Phys. G **41** (Jan, 2014) 055002. 17 p. Comments: 17 pages, 7 figures.

[86] LHCb Collaboration, R. Aaij et al., *Inclusive W and Z production in the forward region at $\sqrt{s} = 7$ TeV*, J. High Energy Phys. **06** (Apr, 2012) 058. 33 p. Comments: 27 pages, 11 figures, 6 tables.

[87] C. Elsasser, $b\bar{b}$ *production angle plots*, LHCb Material for Presentations.

[88] LHCb Collaboration, *For LHCb Talks*, General Photo, July, 2012.

[89] T. Latham, *Velo material approved conference plots*, LHCb VELO group.

[90] M. Ferro-Luzzi, T. Latham, and C. Wallace, *Determination of the aperture of the LHCb VELO RF foil*, Tech. Rep. LHCb-PUB-2014-012. CERN-LHCb-PUB-2014-012, CERN, Geneva, Apr, 2014.

[91] Y. Amhis et al., *b-hadrons absolute lifetime measurements*, LHCb-ANA-2011-055 (Feb, 2014).

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

[93] LHCb Collaboration, R. Aaij et al., *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (Dec, 2014) 1530022. 82 p.

[94] R. W. Forty and O. Schneider, *RICH pattern recognition*, Tech. Rep. LHCb-98-040, CERN, Geneva, Apr, 1998.

[95] LHCb RICH group, M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. C **73** (Nov, 2012) 2431. 25 p.

[96] A. A. Alves Jr et al., *Performance of the LHCb muon system*, J. Instrum. **8** (Nov, 2012) P02022. 29 p.

[97] LHCb MuonID group, F. Archilli et al., *Performance of the Muon Identification at LHCb*, J. Instrum. **8** (Jun, 2013) P10020. 17 p. Comments: 17 pages, 10 figures.

[98] J. Prisciandaro, F. Blanc, and T. Nakada, *Improved magnetic field map with 2011 measurements for the LHCb dipole magnet*, Tech. Rep. LHCb-INT-2012-012. CERN-LHCb-INT-2012-012, CERN, Geneva, Mar, 2012.

[99] M. Frank, C. Gaspar, E. v. Herwijnen, B. Jost, and N. Neufeld, *Deferred High Level Trigger in LHCb: A Boost to CPU Resource Utilization*, J.Phys.Conf.Ser. **513** (2014) 012006.

[100] 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) 209–216, arXiv:1009.2731.

[101] LHCb Collaboration, R. Aaij et al., *Prompt charm production in pp collisions at* $\sqrt{s} = 7\ \ TeV$, Nucl.Phys. **B871** (2013) 1–20, arXiv:1302.2864.

[102] LHCb Collaboration, R. Aaij et al., *Search for the rare decay* $D^0 \to \mu^+\mu^-$, Phys. Lett. B **725** (May, 2013) 15–24. 22 p. Comments: 22 pages, 8 figures, Submitted to Physics Letters B.

[103] LHCb Collaboration, R. Aaij et al., *Search for* $D^+_{(s)} \to \pi^+\mu^+\mu^-$ *and* $D^+_{(s)} \to \pi^-\mu^+\mu^+$ *decays*, Phys. Lett. B **724** (Apr, 2013) 203–212. 21 p. Comments: 21 pages, 5 figures.

[104] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Journal of Basic Engineering **82** (1960) 35–45.

[105] R. Frhwirth, *Application of kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 23 444 – 450.

[106] P. Bartalini et al., *VELO telescope resolution and efficiency measurements*, Tech. Rep. LHCb-2000-099, CERN, Geneva, Oct, 2000. revised version number 1 submitted on 2000-10-27 16:06:45.

[107] C. Parkes, T. Ruf, and T. Szumlak, *Reconstruction of Cluster Positions in the LHCb VELO*, Tech. Rep. LHCb-2007-151. CERN-LHCb-2007-151, CERN, Geneva, Dec, 2007.

[108] W. D. Hulsbergen, *A study of track reconstruction and massive dielectron production in HERA-B.* PhD thesis, Amsterdam University, 2002.

[109] G. Gorfine, M. Hoeferkamp, G. Santistevan, and S. Seidel, *Capacitance of silicon pixels*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **460** (2001), no. 23 336 – 351.

[110] T. Brodbeck and A. Chilingarov, *Simulation of charge collection and sharing in microstrip detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **395** (1997), no. 1 29 – 34. Proceedings of the Fourth International Workshop on GaAs Detectors and Related Compounds.

[111] J. Harrison, S. Chen, and C. Parkes, *Determination of the operating voltages of the VELO and optimisation of the radiation damage scan data collection*, LHCb-ANA-2015-041 (Jul, 2015).

[112] M. Pivk and F. Le Diberder, *: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (Dec., 2005) 356–369, `arXiv:0402083`.

[113] D. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization.* John Wiley and Sons, Inc, 1992.

[114] M. Alexander et al., *Time-Dependent Analyses of Two-Body Charm Decays*, . LHCb-CONF-2011-046; LHCb-CONF-2011-054.

[115] LHCb Collaboration, R. Aaij et al., *Measurement of mixing and CP violation parameters in two-body charm decays*, JHEP **04** (2012) 129, arXiv:1112.4698.

[116] V. V. Gligorov et al., *Swimming : a data driven acceptance correction algorithm*, Journal of Physics: Conference Series **396** (2012), no. 2 022016.

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

[118] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986.

[119] LHCb Collaboration, R. Aaij et al., *Measurements of indirect CP asymmetries in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, Phys. Rev. Lett. **112** (2014), no. 4 041801, arXiv:1310.7201.

[120] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl.Instrum.Meth. **A552** (2005) 566–575, arXiv:physics/0503191.

[121] LHCb Collaboration, R. Aaij et al., *Measurement of the $D^{+/-}$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B718** (2013) 902–909, arXiv:1210.4112.

[122] LHCb Collaboration, R. Aaij et al., *Measurement of indirect CP asymmetries in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays using semileptonic B decays*, JHEP **04** (2015) 043, arXiv:1501.0677.

[123] M. Alexander et al., *A_Γ and y_{CP} in $D^0 \to hh$ with $3\,fb^{-1}$*, LHCb-ANA-2015-039 (Jul, 2015).

[124] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The LHCb Turbo Stream*, LHCb-PROC-2015-013 (May, 2015).

[125] M. Gersabeck, *Charm physics at LHCb - Status and prospects*, LHCb-TALK-2015-039 (Mar, 2015).

[126] LHCb Collaboration, I. Bediaga et al., *Framework TDR for the LHCb Upgrade: Technical Design Report*, Tech. Rep. CERN-LHCC-2012-007. LHCb-TDR-12, CERN, Geneva, Apr, 2012.

[127] Belle-II, T. Abe et al., *Belle II Technical Design Report*, arXiv:1011.0352.

[128] T. Aushev et al., *Physics at Super B Factory*, arXiv:1002.5012.