THE UNIVERSITY *of* EDINBURGH

A measurement of the $\mathcal{CP}$-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\,\phi$

Conor Fitzpatrick
University of Edinburgh

A dissertation submitted to the University of Edinburgh
for the degree of Doctor of Philosophy
April 24, 2012

Abstract

The LHCb experiment is dedicated to making precision measurements involving beauty and charm hadrons at the CERN Large Hadron Collider. The LHCb RICH detectors provide charged particle identification required to distinguish final states in many decays important to the LHCb physics programme. Time alignment of the RICH photon detectors is necessary in order to ensure a high photon collection efficiency. Using both a pulsed laser and proton-proton collision data the photon detectors are aligned to within 1 ns. The LHCb detector is uniquely positioned to measure production cross-sections at energies and rapidities inaccessible to other experiments. With 1.81 nb^{-1} of proton-proton collisions collected by the LHCb experiment in 2010 at center-of-mass energy $\sqrt{s} = 7$ TeV the production cross-section of $\text{D}_\text{s}^\pm$ and $\text{D}^\pm$ mesons decaying to the $\phi\{\text{K}^+\text{K}^-\}\pi^\pm$ final state have been determined in bins of transverse momentum and rapidity. These measurements use a data-driven recursive optimisation technique to improve signal significance. The cross-section ratio is measured to be $\frac{\sigma(\text{D}^\pm)}{\sigma(\text{D}_\text{s}^\pm)} = 2.32 \pm 0.27(\text{stat}) \pm 0.26(\text{syst})$, consistent with the ratio of charm-quark hadronisation fractions to $\text{D}^\pm$ and $\text{D}_\text{s}^\pm$ mesons. Time-dependent interference between mixing of B_s^0-$\overline{\text{B}}_\text{s}^0$ mesons and decay to the final state $\text{J}/\psi\phi$ gives rise to a $\mathcal{CP}$ violating phase ϕ_s. This phase is constrained to be small within the Standard Model, a significant deviation from which would be a signal of new physics. ϕ_s has been measured with 0.37 fb^{-1} of proton-proton collision data recorded during 2011 by the LHCb experiment. Isolation of the signal distribution is achieved using the $\mathcal{S}$-plot technique, and the analysis accounts for inclusive $\text{B}_\text{s}^0 \to \text{J}/\psi\text{K}^+\text{K}^-$ s-wave contributions. The measured value of $\phi_s = 0.16 \pm 0.18(\text{stat}) \pm 0.06(\text{syst})$ rad is the most precise measurement to date, and is consistent with Standard Model predictions.

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This thesis is written in accordance with the postgraduate assessment regulations for research degrees.

Conor Fitzpatrick

Acknowledgements

I owe a debt of gratitude to the staff and students of the Edinburgh Particle Physics (Experiments) Group, in particular my supervisor, Franz, who has put up with me for longer than anyone deserves, and whose guidance has always been the perfect balance of carrot and stick. Pete, Greig, Yuehong, Matt and the other members of the LHCb β_s working group have been a constant source of advice and encouragement throughout the $B_s^0 \to J/\psi\phi$ analysis while Thierry, Carmelo, Olav, Antonis and Tom along with the rest of the RICH operations group past and present have made the many sleepless nights in the control room eagerly awaiting beam enjoyable. I am thankful to my Mother, Olivia, for forgiving all the forgotten phone calls and infrequent trips home by "her son, the rocket scientist". I am also indebted to Alan Walker for his encouragement, without which I would never have considered the path I have taken. Ben, Ailsa, Ross, Gemma and the rest of the Edinburgh postgraduates in the physics faculty deserve a thank you as they have shared in the tireless pursuit of intellectual [1] and social [2] stimulation. Lastly, I dedicate this thesis to Caterina, whose successes inspire what little there are of my own.

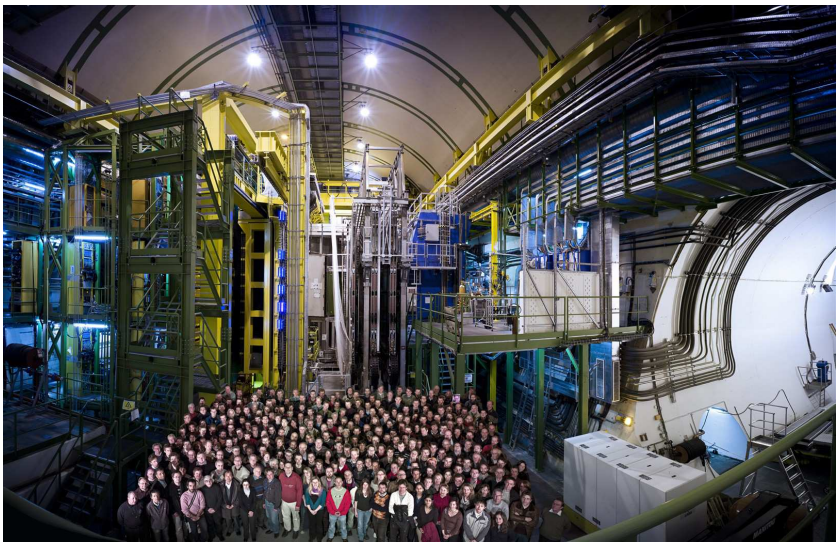


Figure 1.: The LHCb collaboration, pictured in front of the LHCb detector.

Preface

This thesis describes my research as a member of the LHCb collaboration, one of the experiments on the Large Hadron Collider at CERN in Geneva. The LHCb detector, shown in Figure 1, has been designed and built in order to probe conditions in the universe shortly after the Big Bang, when matter and antimatter were created in equal amounts. This equality was short lived; matter and antimatter failed to completely annihilate due to subtle differences or *asymmetries* in their behavior. These asymmetries led to the universe we observe today, in which less than one antimatter particle exists for every billion matter ones. LHCb studies one possible way in which these asymmetries could occur using particles produced only at high energies called *beauty* and *charm* quarks. The main body of this thesis focusses on results obtained with the first two years running of the LHCb experiment. One measurement relates to the rate of production of particles containing charm quarks at LHC energies, and another more detailed measurement is made of beauty-containing particles in which a specific case of the antimatter-matter asymmetry is measured.

Contents

"To study and not think is a waste. To think and not study is dangerous."

— Confucius

Chapter 1.

$\mathcal{CP}$-violation in the Standard Model and beyond

"Nature is wont to hide herself"
— Heraclitus

The Standard Model attempts to describe nature at the most fundamental level. It has been overwhelmingly successful in both predicting and incorporating measurements of the properties of the universe on both the large (relativistic) and small (quantum) scale. It is the role of experimentalists to overconstrain the parameters of the Standard Model through direct and indirect measurements in order to confirm the Standard Model, or herald its replacement with new theoretical models. This chapter will briefly describe the Standard Model and culminate in a description of one such Standard Model parameter and the motivation for its measurement, $\phi_s^{SM} = -2\beta_s$.

1.1. A Brief Introduction to the Standard Model

The Standard Model (SM) of particle physics is an $\mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ gauge theory that unites three of the fundamental forces of nature in order to explain the structure of the fundamental particles and their interactions.

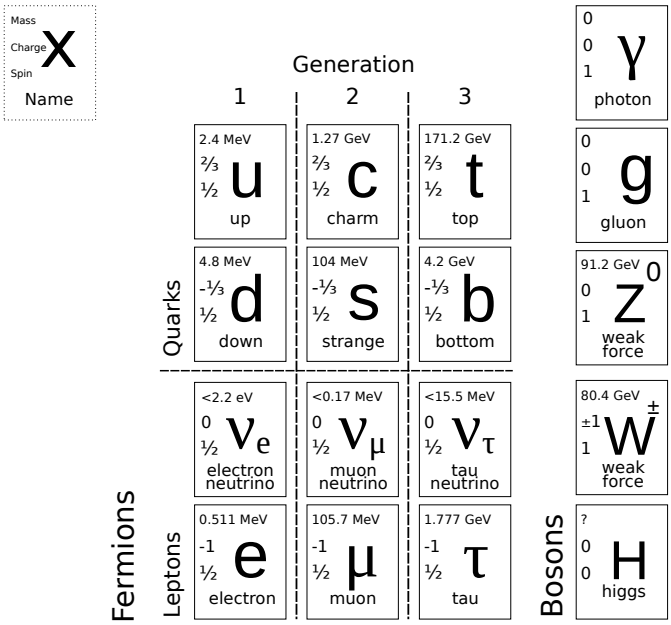


Figure 1.1.: The fundamental particles of the Standard Model, sorted according to family, generation and mass.

1.1.1. The fundamental particles

The SM consists of 12 spin 1/2 fermions and their antiparticles, 12 vector gauge bosons, and 1 as-yet-unmeasured scalar boson, summarised in Figure 1.1. The fermions are pointlike as to current resolution they show no sign of internal structure. The gauge bosons are mediators of the forces which act upon the fermions, leading to the interactions and rich structure of the SM which continues to accommodate the processes observed in HEP experiments.

1.1.2. The fundamental forces

The fundamental forces described by the Standard Model operate in different sectors, and act upon subgroups of the fermions by way of exchange of virtual (gauge) bosons. The SM consists of two sectors, electro-weak and strong (QCD), in which the different forces act. The following subsections summarise these sectors.

The electro-weak theory

As a quantum field theory, tests of the electro-weak (EW) sector of the SM agree to a precision better than 1% across many measurements. It is a triumph of the field theory approach [3]. The EW theory combines electromagnetism with the weak force to form a single relativistic quantum field theory. This is unusual as the electromagnetic force has infinite range, whilst the weak force has influence only over ranges of $O(10^{-18})$ m. The relative strengths of the weak and electromagnetic forces differ greatly as well. The EW sector is characterised by an $\mathrm{SU}(2)_L$ weak isospin symmetry with a $\mathrm{U}(1)_Y$ weak hypercharge phase symmetry, yielding quantum numbers for all of the fermions: The leptonic sector consists of the left-handed $\mathrm{e}^-, \mu^-, \tau^-$ and associated neutrinos $\nu_\mathrm{e}, \nu_\mu, \nu_\tau$ having weak isospin $I = \frac{1}{2}$ and weak hypercharge $Y(L_l) = -1$,

$$L_\mathrm{e} = \begin{pmatrix} \nu_\mathrm{e} \\ \mathrm{e}^- \end{pmatrix}_L \qquad L_\mu = \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L \qquad L_\tau = \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \tag{1.1}$$

and the right-handed weak isoscalar leptons with weak hypercharge $Y(R_l) = -2$:

$$R_\mathrm{e} = \mathrm{e}_R \qquad R_\mu = \mu_R \qquad R_\tau = \tau_R \tag{1.2}$$

where electric charge is determined as $Q = I_3 + \frac{1}{2}Y$, with I_3 being the third component of the weak isospin. The hadronic (quark) sector of the EW theory consists of the left-handed quark doublets:

$$L_q^1 = \begin{pmatrix} u \\ d' \end{pmatrix}_L \qquad L_q^2 = \begin{pmatrix} c \\ s' \end{pmatrix}_L \qquad L_q^3 = \begin{pmatrix} t \\ b' \end{pmatrix}_L \tag{1.3}$$

having weak isospin $I = \frac{1}{2}$ and weak hypercharge $Y(L_q) = \frac{1}{3}$, and right-handed weak isoscalar quarks:

$$R_u^{(1,2,3)} = u_R, c_R, t_R \qquad R_d^{(1,2,3)} = d_R, s_R, b_R \tag{1.4}$$

with weak hypercharges $Y(R_u) = \frac{4}{3}$, $Y(R_d) = -\frac{2}{3}$. The primes on the second components of the left-handed quark doublets is to distinguish the weak eigenstates from the mass eigenstates which will be discussed further in Section 1.1.3. The theory is described mathematically by a Lagrangian (density) of the form

$$\mathcal{L} = \mathcal{L}_\mathrm{gauge} + \mathcal{L}_l + \mathcal{L}_q \tag{1.5}$$

where

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}\sum_{l} F^{l}_{\mu\nu}F^{l\ \mu\nu} - \frac{1}{4}f_{\mu\nu}f^{\mu\nu} \tag{1.6}$$

$$\mathcal{L}_{l} = \bar{R}_{l} i\gamma^{\mu}\left(\partial_{\mu} + i\frac{g'}{2}\mathcal{A}_{\mu}Y\right)R_{l} \tag{1.7}$$

$$+\bar{L}_{l} i\gamma^{\mu}\left(\partial_{\mu} + i\frac{g'}{2}\mathcal{A}_{\mu}Y + i\frac{g}{2}\boldsymbol{\tau}\cdot\mathbf{b}_{\mu}\right)L_{l} \tag{1.8}$$

$$\mathcal{L}_{q} = \bar{R}^{(n)}_{u} i\gamma^{\mu}\left(\partial_{\mu} + i\frac{g'}{2}\mathcal{A}_{\mu}Y\right)R^{(n)}_{u} \tag{1.9}$$

$$+\bar{R}^{(n)}_{d} i\gamma^{\mu}\left(\partial_{\mu} + i\frac{g'}{2}\mathcal{A}_{\mu}Y\right)R^{(n)}_{d} \tag{1.10}$$

$$+\bar{L}^{(n)}_{q} i\gamma^{\mu}\left(\partial_{\mu} + i\frac{g'}{2}\mathcal{A}_{\mu}Y + i\frac{g}{2}\boldsymbol{\tau}\cdot\mathbf{b}_{\mu}\right)L^{(n)}_{q} \tag{1.11}$$

in which $\boldsymbol{\tau}$ denotes the Pauli isospin matrices. The coupling constants are g and g', and the field strength tensors are:

$$F^{l}_{\mu\nu} = \partial_{\nu}b^{l}_{\mu} - \partial_{\mu}b^{l}_{\nu} + g\epsilon^{jkl}b^{j}_{\mu}b^{k}_{\nu} \tag{1.12}$$

$$f_{\mu\nu} = \partial_{\nu}\mathcal{A}_{\mu} - \partial_{\mu}\mathcal{A}_{\nu} \tag{1.13}$$

for the weak isospin and weak hypercharge symmetries respectively. The gauge fields are that of a weak isovector, $\mathbf{b}_{\mu}$ and a weak isoscalar, $\mathcal{A}_{\mu}$, resulting in four massless gauge bosons $b^{1}_{\mu}, b^{2}_{\mu}, b^{3}_{\mu}$ and $\mathcal{A}_{\mu}$. These are massless because a term of the form $\frac{1}{2}m^{2}\mathcal{A}_{\mu}\mathcal{A}^{\mu}$ is not invariant under transformation of gauge. This results in several problems:

- The $W^{\pm}, Z$ bosons are observed to be massive. The only massless boson is the photon.
- The difference between the left and right-handed fields require that they transform differently, forbidding mass terms for the fermions.

In order to give the fermions and bosons mass the electro-weak symmetry must somehow be broken. By introducing a complex doublet of scalar fields of the form:

$$\phi = \begin{pmatrix}\phi^{+}\\ \phi^{0}\end{pmatrix} \tag{1.14}$$

having weak hypercharge $Y_\phi = 1$, we may supplement the EW Lagrangian with an additional term for the interaction and propagation of these scalars:

$$\mathcal{L}_{\text{scalar}} = (\mathcal{D}^\mu\phi)^\dagger(\mathcal{D}_\mu\phi) - V(\phi^\dagger\phi) \qquad \left[\mathcal{D}_\mu = \delta_\mu + i\frac{g'}{2}\mathcal{A}_\mu Y + i\frac{g}{2}\boldsymbol{\tau}\cdot\mathbf{b}_\mu\right] \tag{1.15}$$

Such that $g' = g\tan\theta_W$ with θ_W the Weinberg angle which determines the level of weak mixing. The potential may be written as:

$$V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \tag{1.16}$$

and the gauge invariant interactions between the scalar fields and the fermions takes the form:

$$\mathcal{L}_{\text{Yukawa}} = -\zeta[(\bar{L}\phi)R + \bar{R}(\phi^\dagger L)] \tag{1.17}$$

Allowing μ^2 to be negative, we can choose the vacuum state to have expectation value:

$$\langle\phi\rangle_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \qquad v = \sqrt{\frac{-\mu^2}{|\lambda|}} \tag{1.18}$$

which breaks the gauge symmetry $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y \to \mathrm{U}(1)_Q$. The ζ term in the interaction term between scalar and fermion fields is known as the Yukawa coupling, one of which exists for each fermion and determines its mass. The size of the Yukawa coupling is not predicted by the theory. The photon remains massless, but the three remaining gauge bosons acquire mass through additional longitudinal degrees of freedom provided by three of the auxiliary scalars. The physical photon, A, arises from the introduction of the Weinberg angle, through the linear combination of two of the previous gauge fields: $A = \mathcal{A}\cos\theta_W + b_3\sin\theta_W$. Orthogonal to this is the Z boson, $Z = b_3\cos\theta_W - \mathcal{A}\sin\theta_W$, and the remaining two bosons $W^\pm = (b_1 \mp ib_2)/\sqrt{2}$. The $W^\pm$ bosons acquire mass $M_W = gv/2 = ev/2\sin\theta_W$ and for the Z, $M_Z = M_W/\cos\theta_W$. This is a remarkable result, but even more so because one auxiliary scalar remains. This manifests itself in the theory as a massive spin-zero scalar particle, the Higgs boson. Again, the theory does not predict its mass.

The strong force

The remaining $SU(3)$ sector of the SM is the color[1] group of Quantum Chromodynamics (QCD). It is surprisingly simple to define, requiring as input only a coupling strength to specify it completely, yet it is host to a wealth of incredibly complex phenomena. The QCD Lagrangian is an $SU(3)$ gauge theory with color triplet quark matter fields [4]:

$$\mathcal{L}_{\text{strong}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_q \tag{1.19}$$

Where:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}\sum_i F^i_{\mu\nu} F^{i\ \mu\nu} \tag{1.20}$$

$$\mathcal{L}_q = \bar{q}^{(r)} i\gamma^\mu \left(\partial_\mu + i\frac{g_s}{2}\boldsymbol{\lambda}\cdot\mathbf{g}_\mu\right) q^{(r)} \tag{1.21}$$

Unlike in the EW sector, QCD acts upon left and right-handed quarks equally and contains no flavour-changing currents, so we specify the color triplet quark fields q of $r = 1 \ldots n_f$ different flavours without further discrimination. $\boldsymbol{\lambda}$ are the representation matrices $\lambda^i, i = 1 \ldots 8$, analagous to the isospin matrices of EW, and $\mathbf{g}_\mu$ are the gluon fields, with corresponding massless vector bosons. The field strength tensor is of the form:

$$F^i_{\mu\nu} = \partial_\mu g^i_\nu - \partial_\nu g^i_\mu - g_s f_{ijk} g^j_\mu g^k_\nu \tag{1.22}$$

Where the structure constants f_{ijk} are related to the representation matrices by $[\lambda_i, \lambda_j] = 2if_{ijk}\lambda^k$ and λ^i are normalised such that $Tr[\lambda^i, \lambda^j]\lambda^k = 4if_{ijk}$. Lastly, g_s is the strong coupling constant. The quadratic term in the QCD field strength tensor permits gluon-gluon interactions, notably not present for photons in the EW tensor of similar form [5]. The theory is non-abelian: this makes the field equations nonlinear and is responsible for the complexity of phenomena and structure of QCD. Perhaps the most interesting of these phenomena is *asymptotic freedom*. In a non-abelian theory such as QCD the coupling between the bosons and quarks is the same for all particles regardless of color or flavour, something that is not true of the EW theory. The β function, which describes the variation of the coupling as a function of the energy scale through the renormalisation

[1] The American spelling of colour is commonly used to differentiate between the quantum number associated with the $SU(3)_C$ sector of the SM and the electromagnetic phenomena in the range 380-750nm.

point μ,

$$\beta(g) = \mu \frac{dg(\mu)}{d\mu} \tag{1.23}$$

may, in the case of a non-abelian theory be negative. This has implications in the effects of screening: virtual particle pairs of very high energies created and annihilated within the vacuum interact over short timescales, and arise as loop corrections to both QED and QCD. These virtual particles have the effect of shielding bare charges in QED, so that at increasing distances (decreasing energies) the field is increasingly cancelled. In QCD the self-interactions of the gluons permit not only $q\bar{q}$ pairs to screen the color field, but for gg pairs to enhance it. Thus the coupling in QCD as a function of distance increases [6]. At energies significantly higher than the renormalisation scale:

$$\alpha_s(q^2) = \frac{g_s^2(q^2)}{4\pi} = \frac{12\pi}{33 - 2n_f \left(\ln \frac{q^2}{\mu^2}\right)} \qquad [q^2 >> \mu^2] \tag{1.24}$$

This yields another interesting phenomenon unique to QCD, *confinement*. The color fields of QCD can be considered narrow tubes of energy density $\approx$ 1 GeV/fm. At increasing distances it becomes energetically favorable to create new quark-antiquark pairs, having the effect of splitting the color field. This implies that *no isolated color charges exist*, so the quarks are confined to color-neutral bound states called hadrons, of which two types are presently observed: baryon ($q_r q_g q_b$) states and mesons ($q_r \bar{q}_{\bar{r}}$), in which quark flavour may be different. This leads to a rich landscape of particles categorised by way of the quantum numbers of their constituent quarks through the Quark Model [7].

1.1.3. Symmetries, quark mixing and the CKM mechanism

Symmetries in the Standard Model

Symmetries take a special place in physics. Noether's theorem [8] shows that for every action invariant under a continuous transformation there exists a conserved property. The conservation laws so crucial to our understanding of physics therefore stem from symmetries. While Noether's theorem applies directly to continuous symmetries, discrete symmetries are not necessarily subject to the same requirement. In fact, some

Invariant under	Process	Conserved quantity
Translation in time	$t \to t + \delta t$	Energy
Rotation	$(r, \theta) \to (r, \theta + \delta\theta)$	Angular momentum
Translation in space	$\vec{x} \to \vec{x} + \delta\vec{x}$	Linear momentum

Table 1.1.: Common conserved quantities arising from continuous symmetries in physics

have been shown to be broken within the weak interactions of the Standard Model. The three discrete symmetries of greatest interest are:

- Charge conjugation, $\mathcal{C}$, which flips the sign of all *internal* (additive) quantum numbers. Under $\mathcal{C}$ a particle becomes its antiparticle.
- Parity inversion, $\mathcal{P}$, which reverses the handedness of space, for example $\vec{r} = (x, y, z) \to -(x, y, z)$. Under $\mathcal{P}$ the helicity of a particle is reflected; a right-handed particle becomes left-handed.
- Time reversal, $\mathcal{T}$, $t \to -t$. Under time reversal, a right-handed particle is equivalent to a left-handed antiparticle, ie a complete reversal of the process.

It is immediately clear from equations (1.1) and (1.2) that $\mathcal{P}$ is maximally violated in the weak sector; if we apply $\mathcal{P}$ to a left-handed neutrino, we obtain a right handed neutrino. Assuming no neutrino mixing right-handed neutrinos do not exist within the SM, and so the weak force does not conserve $\mathcal{P}$. Similarly, $\mathcal{C}$ must be maximally violated as left-handed antineutrinos have no interaction within the SM. If we apply $\mathcal{CP}$ however, we find that the combined effect of charge conjugation and parity reversal converts a left-handed neutrino into a right-handed antineutrino. In this particular instance, $\mathcal{CP}$ is said to be conserved. An example of this is illustrated in the decay $\pi^+ \to \mu^+\nu_\mu$ in Figure 1.2. While $\mathcal{C}$, $\mathcal{P}$ symmetries are trivially broken in the EW sector of the SM, what about $\mathcal{CP}$? While this was originally assumed to be an exact symmetry, measurement of the decay of Kaons in the sixties [9] showed that $\mathcal{CP}$ is also broken, though only to a small degree. This implies that in order for $\mathcal{CPT}$ symmetry to hold (which is essential for the SM to remain a viable theory) $\mathcal{T}$ symmetry must also be broken. To date, no evidence of $\mathcal{CPT}$ symmetry violation has been found.

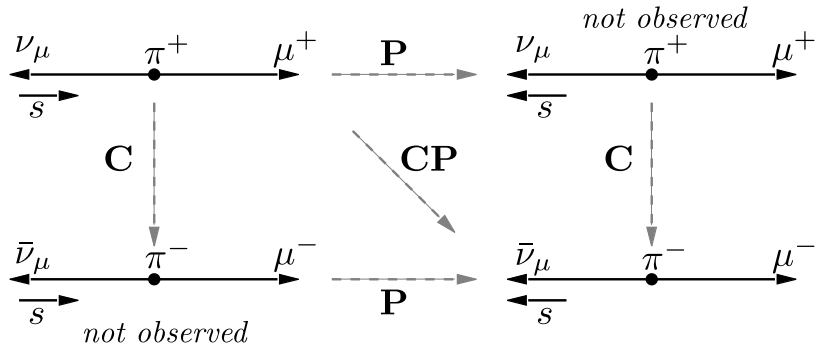


Figure 1.2.: $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{CP}$ acting upon the decay $\pi^+ \to \mu^+\nu_\mu$. The helicity of the neutrino, defined as the projection of the particle's spin s, onto its momentum, denotes the handedness of the particle.

Quark mixing, and the CKM mechanism

The only source of $\mathcal{CP}$-violation ($\not{\mathcal{CP}}$) within the Standard Model arises through the mixing of quarks[2]. In equation (1.3) the primes on the second component of each quark doublet are used to indicate the weak eigenstate. The mass eigenstates are related to the weak eigenstates through the couplings of the $W^\pm$ bosons, which arise from the Yukawa interactions with the Higgs condensate. Taking equation (1.17) and explicitly including the quark terms only:

$$\mathcal{L}_{\text{Yukawa}} = -\zeta^d_{ij}\bar{L}^i_q\phi R^j_d - \zeta^u_{ij}\bar{L}^i_q\phi^\dagger R^j_u + h.c. \tag{1.25}$$

where $h.c.$ denotes the Hermitian conjugate of the preceding terms. The physical masses of the quarks arise from diagonalisation of $\zeta^{u,d}$ by four unitary matrices, $V^{u,d}_{L,R}$:

$$M^u_{\text{diag}} = V^u_L\zeta^u V^{u\dagger}_R\frac{v}{\sqrt{2}} \qquad M^d_{\text{diag}} = V^d_L\zeta^d V^{d\dagger}_R\frac{v}{\sqrt{2}} \tag{1.26}$$

The $W^\pm$ couplings to the physical quark fields L^n_q are therefore given by:

$$V_{CKM} \equiv V^u_L V^{d\dagger}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.27}$$

[2]we assume neutrino mass to be outwith the SM

Thus the relation between the physical d, s, b and weak eigenstates d', s', b' of the quarks is given by:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.28}$$

V_{CKM} is the 3×3 complex, unitary Cabibbo-Kobayashi-Maskawa mixing matrix. The unitarity condition in combination with the ability to re-phase the quark fields implies that there are only 4 free parameters of which one is the complex mixing phase. $\mathcal{CP}$ arises from this single mixing phase, which becomes clear if we re-express the matrix in the Wolfenstein parameterisation [10], using the terms:

$$\lambda = \frac{|V_{us}|}{\sqrt{|V_{us}|^2 + |V_{ud}|^2}} \qquad A\lambda^2 = \left|\frac{V_{cb}}{V_{us}}\right| \qquad A\lambda^3(\rho + i\eta) = V_{ub}^* \tag{1.29}$$

where ρ, A, λ, η are four real parameters, $\lambda = \sin\theta_c$, the sin of the Cabibbo angle, and $i\eta$ parameterises the complex phase responsible for $\mathcal{CP}$-violation. It is also common to define the parameters $\bar{\eta} = \eta(1 - \lambda^2/2)$ and $\bar{\rho} = \rho(1 - \lambda^2/2)$. Expressing V_{CKM} in this form to $O(\lambda^3)$ with additional terms up to $O(\lambda^5)$ provides an insight into the magnitude of the elements:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \delta V_{CKM} \tag{1.30}$$

$$\delta V_{CKM} = \begin{pmatrix} 0 & 0 & 0 \\ -i\,A^2\lambda^5\eta & 0 & 0 \\ A\lambda^5\,(\rho + i\,\eta)\,/2 & A\lambda^4\,(1/2 - \rho - i\,\eta) & 0 \end{pmatrix} \tag{1.31}$$

It is important to note however, that the Wolfenstein parameterisation is *convention specific*. In this thesis, care has been taken to ensure convention-independent parameterisation of the phases, relying upon convention-specific notation only where an aid to determining the magnitudes of parameters is deemed useful.

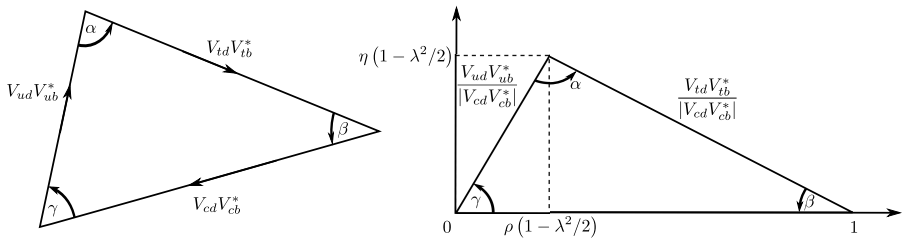


Figure 1.3.: The unitary triangle, expressed in terms of CKM elements (left) and in the $\bar{\eta}, \bar{\rho}$ plane (right)

The unitary triangles

The unitarity condition of the CKM matrix imposes the constraints $V_{ij}V_{ik}^* = \delta_{jk}$ and $V_{ij}V_{kj}^* = \delta_{ik}$, yielding six terms for which $\delta_{jk}, \delta_{ik} = 0$. These six relations can be represented as triangles in the complex plane with area $J/2$ where $\mathrm{Im}[V_{ij}V_{kl}V_{il}^*V_{kj}^*] = J\sum_m \sum_n \epsilon_{ikm}\epsilon_{jln}$. Of these six, two are of particular interest to this thesis:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.32}$$

and

$$V_{us}V_{ub}^* + V_{ts}V_{tb}^* + V_{cs}V_{cb}^* = 0 \tag{1.33}$$

The first of these is commonly referred to as 'the' unitary triangle as shown in Figure 1.3, the angles of which are defined as:

$$\alpha = \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right] \qquad \beta = \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right] \qquad \gamma = \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right] = \pi - \alpha - \beta \tag{1.34}$$

And analogous to β, the second 'bs' unitary triangle has an angle:

$$\beta_s = \arg\left[-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right] \tag{1.35}$$

1.2. Manifestation of $\mathcal{CP}$-violation in the Standard Model

The existence of $\mathcal{CP}$-violation has profound consequences: It permits unambiguous discrimination between matter and antimatter in experiments, and introduces processes by which the matter-antimatter asymmetry of our present universe may be studied. Within the SM, $\not{\mathcal{CP}}$ takes three distinct forms[3] dependent upon the process observed, but all are attributable to the single complex phase of the CKM mixing matrix.

1.2.1. $\not{\mathcal{CP}}$ in decay

We define the decay amplitude A of a meson X to final state f as being:

$$A_f = \langle f|\mathcal{H}|X\rangle \tag{1.36}$$

Where $\mathcal{H}$ denotes the Hamiltonian governing the interaction. The $\mathcal{CP}$ conjugate decay of $\bar{X}$ to $\bar{f}$ may be similarly expressed as

$$\bar{A}_{\bar{f}} = \left\langle \bar{f}|\mathcal{H}|\bar{X}\right\rangle \tag{1.37}$$

If X is charged, these are the only two allowed $\mathcal{CP}$ conjugates. The action of $\mathcal{CP}$ on the initial and final states introduces arbitrary unphysical phases dependent upon the flavour content:

$$\mathcal{CP}\,|X\rangle = e^{+i\phi_i}\left|\bar{X}\right\rangle \qquad \mathcal{CP}\,|f\rangle = e^{+i\phi_f}\left|\bar{f}\right\rangle \tag{1.38}$$

$$\mathcal{CP}\left|\bar{X}\right\rangle = e^{-i\phi_i}\,|X\rangle \qquad \mathcal{CP}\left|\bar{f}\right\rangle = e^{-i\phi_f}\,|f\rangle \tag{1.39}$$

Which means that if $\mathcal{CP}$ is conserved in the conjugate decay, the amplitudes have the same magnitude but differ by the phase of the initial and final state:

$$\bar{A}_{\bar{f}} = e^{i(\phi_f-\phi_i)}A_f \tag{1.40}$$

[3]This subsection is summarised from the excellent reviews sections in reference [7]

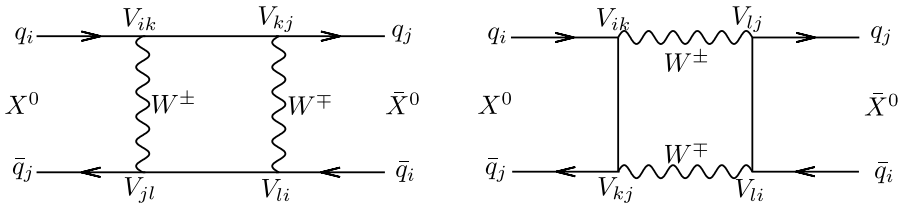


Figure 1.4.: Neutral meson mixing, as permitted by the CKM mechanism. In this instance an arbitrary neutral meson, denoted X consisting of a quark of flavour i and antiquark of flavour j mixes through the exchange of two W bosons to its $\mathcal{CP}$ conjugate, by way of (left) t-channel and (right) s-channel feynman diagrams.

The existence of $\mathcal{CP}$-violation in decay is therefore true if

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1 \tag{1.41}$$

1.2.2. $\mathcal{CP}$ in mixing

A particularly beautiful result of the CKM mechanism is that for neutral mesons, time-dependent mixing between particle and antiparticle is possible via the exchange of two charged weak currents. A meson consisting of two quarks, $X = q\bar{q}'$ may over time mix by way of the box diagram shown in Figure 1.4 to the state $\bar{X} = \bar{q}q'$. This permits another source of $\mathcal{CP}$ violation: over time the number of $X \to \bar{X}$ transitions may be lower than the $\mathcal{CP}$ conjugate $\bar{X} \to X$. If we consider a superposition of strong eigenstates at $t = 0$ of the form:

$$|\psi(t)\rangle = |X(0)\rangle + \left|\bar{X}(0)\right\rangle + |f_{i\ldots j}(0) = 0\rangle \tag{1.42}$$

The system evolves in time through a combination of mixing; $X \to \bar{X}$, $\bar{X} \to X$ and decay: $X \to f_i$, $\bar{X} \to \bar{f}_j$ where i, j represent all final states available to either meson. If we are only interested in mixing, we may neglect the final states and study only $X(t), \bar{X}(t)$. This leads to a simplification whereby we can specify a 2×2 effective Hamiltonian:

$$i\frac{d}{dt}\begin{pmatrix} |X(t)\rangle \\ \left|\bar{X}(t)\right\rangle \end{pmatrix} = \left(\mathbf{M} + \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix} |X(t)\rangle \\ \left|\bar{X}(t)\right\rangle \end{pmatrix} \tag{1.43}$$

$\mathbf{M}$ and $\mathbf{\Gamma}$ are Hermitian matrices, with $X(t)$ denoting the time evolution of the mass eigenstate of a meson produced as an X at $t = 0$, conversely for $\bar{X}(t)$. The off diagonal elements of $\mathbf{M}$ are what drives mixing. The mass eigenstates, which by convention are

referred to as $X_{\mathrm{L}}, X_{\mathrm{H}}$ may be related to the strong eigenstates $X, \bar{X}$ by way of three complex parameters:

$$|X_{\mathrm{L}}\rangle \propto p\sqrt{1-z}\,|X\rangle + q\sqrt{1+z}\,\left|\bar{X}\right\rangle \tag{1.44}$$

$$|X_{\mathrm{H}}\rangle \propto p\sqrt{1+z}\,|X\rangle - q\sqrt{1-z}\,\left|\bar{X}\right\rangle \tag{1.45}$$

where we apply the normalisation condition that for $z = 0, |p|^2 + |q|^2 = 1$. z is only nonzero should $\mathcal{CPT}$ be violated. We will assume $z = 0$ for the purposes of this thesis. It is worth defining the mass Δm and width $\Delta\Gamma$ differences of the two eigenstates:

$$\Gamma = \tau^{-1} = \frac{\Gamma_{\mathrm{H}} + \Gamma_{\mathrm{L}}}{2} \qquad \Delta m = M_{\mathrm{H}} - M_{\mathrm{L}} \qquad \Delta\Gamma = \Gamma_{\mathrm{L}} - \Gamma_{\mathrm{H}} \tag{1.46}$$

By expanding the eigenvalues in q/p and the off-diagonal elements M_{12}/Γ_{12} we obtain:

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}} \tag{1.47}$$

From which we obtain three identities:

$$\Delta m = 2|M_{12}| \qquad \Delta\Gamma = 2|\Gamma_{12}|\cos\phi \qquad \frac{q}{p} = e^{-i(\phi_M - \phi_i)} \tag{1.48}$$

where we define the phase ϕ_M as $\arg(M_{12})$, and $\phi = \arg(-M_{12}/\Gamma_{12})$. ϕ_i is the arbitrary unphysical phase introduced in equation (1.38). If $\mathcal{CP}$ is conserved, $\phi_M = 0$ and $(q/p)^2 = e^{i2\phi_i}$. The condition $|q/p| \neq 1$ indicates $\mathcal{CP}$ violation in mixing.

1.2.3. $\mathcal{CP}$ in the interference between mixing and decay

The final manifestation of $\mathcal{CP}$-violation is in the decay of neutral mesons to the same $\mathcal{CP}$ eigenstate: In this instance a combination of the previous two manifestations arises. For neutral mesons for which the same $\mathcal{CP}$ eigenstate is accessible by both X and $\bar{X}$ the decay amplitudes are:

$$A_f = \langle f|\mathcal{H}|X\rangle \tag{1.49}$$

$$\bar{A}_f = \left\langle f|\mathcal{H}|\bar{X}\right\rangle \tag{1.50}$$

If no $\mathcal{CP}$ is present in either mixing or decay of the mesons individually it is still possible for

$$\Im\mathfrak{m}(\lambda_f) = \Im\mathfrak{m}\left(\frac{q}{p}\frac{\bar{A}_f}{A_f}\right) \tag{1.51}$$

to be nonzero, thus permitting $\mathcal{CP}$.

1.3. ϕ_s in $B_s \to J/\psi\phi$

One of the key goals of HEP experiments in recent years has been to overconstrain the unitary triangle. The level of $\mathcal{CP}$-violation observed in processes susceptible to it is as yet too small to explain the matter-antimatter asymmetry present in the universe. The phase $\phi_s = -2\beta_s$ is of particular importance as it is yet to be measured to a precision that unambiguously confirms or refutes the SM prediction. ϕ_s is measurable through mixing induced $\mathcal{CP}$ in the B_s system, a derivation of which is presented in the following subsections.

1.3.1. $B_s - \bar{B}_s$ mixing

If we consider the B_s specific form of the box diagram shown in Figure 1.4, the M_{12} component is dominated by the top quark in the box, yielding

$$\phi_M = \arg(M_{12}) = \arg(V_{ts}V_{tb}^*)^2 \tag{1.52}$$

Which enters into q/p as described in Section 1.2.2. This is suppressed to the fourth power of the weak coupling constant and two additional powers of $|V_{ts}| \approx 0.04$ in the SM, so the inclusion of a new physics contribution to the box can have a sizeable effect upon the mixing phase [11].

1.3.2. β_s and $b \to c\bar{c}s$ transitions

The decay to $J/\psi\phi$ is a $b \to c\bar{c}s$ transition, which occurs at tree (T) and penguin (P_q) level where q denotes the internal mediating quark. The amplitude, being a sum of these

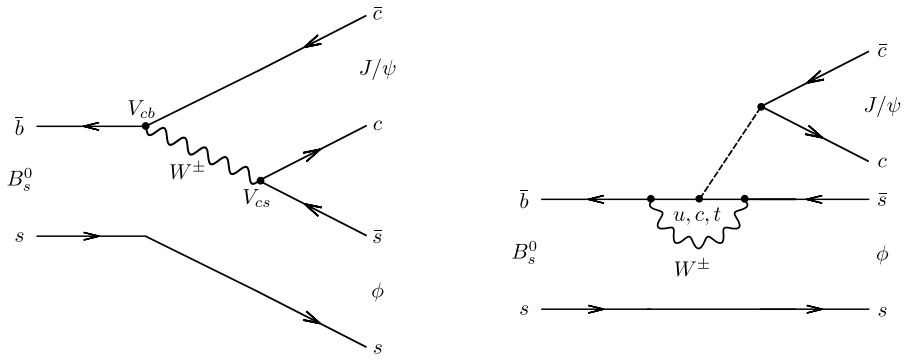


Figure 1.5.: $B_s^0 \to J/\psi\,\phi$ topologies within the SM at (left) tree level contribution and (right) penguin level.The tree-level mode is CKM favored, and dominates the decay with amplitude determined by $V_{cs}V_{cb}^*$.The penguin amplitude obtains terms from $V_{cs}V_{cb}^*$, $V_{us}V_{ub}^*$ and $V_{ts}V_{tb}^*$.

components, takes the form:

$$A_{c\bar{c}s} = V_{cs}V_{cb}^*(T + P_c) + V_{us}V_{ub}^*P_u + V_{ts}V_{tb}^*P_t \tag{1.53}$$

which may be simplified using identity equation (1.33) to:

$$A_{c\bar{c}s} = V_{cs}V_{cb}^*(T + P_c - P_t) + V_{us}V_{ub}^*(P_u - P_t) \tag{1.54}$$

In the Wolfenstein parameterisation $V_{us}V_{ub}^*$ is suppressed by a factor of λ^2 with respect to $V_{cs}V_{cb}^*$, small enough to be treated as a systematic error. The weak phase entering from the $b \to c\bar{c}s$ transition is therefore $\phi_f \approx \arg(V_{cs}V_{cb}^*)$, where the approximation arises from ignoring the additional penguin contribution. We can write

$$\frac{\bar{A}_f}{A_f} = \eta_f e^{i(2\phi_f - \phi_i)} \tag{1.55}$$

where $\eta_f = \pm 1$ is the $\mathcal{CP}$ eigenvalue of the final state. The product of q/p, $\bar{A}_f/A_f$ is therefore:

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} = \eta_f e^{-i(\phi_M - \phi_i)} e^{i(2\phi_f - \phi_i)} = \eta_f e^{-i(\phi_M - 2\phi_f)} \tag{1.56}$$

from which we obtain the overall weak phase which is referred to in the literature as $\phi_{\rm s}$[4]:

$$\phi_{\lambda_f} = \phi_{\rm s} = -\arg(\eta_f \lambda_f) = \phi_M - 2\phi_f \tag{1.57}$$

In the SM, this overall phase, neglecting penguin contributions to ϕ_f may be expressed as:

$$\phi_{\rm s} = \arg(V_{ts}V_{tb}^*)^2 - 2\arg(V_{cs}V_{cb}^*) = 2\arg\left[\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right] \equiv -2\beta_{\rm s} \tag{1.58}$$

If we return to the Wolfenstein parameterisation, we see that

$$2\arg\left[\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right] = 2\arg\left[\frac{-1+\lambda^2/2-\rho\lambda^2-i\eta\lambda^2}{1-\lambda^2/2}\right] = \frac{-2\eta\lambda^2}{1-\lambda^2/2} = -2\eta\lambda^2 - \eta\lambda^4 - O(\lambda^6) \tag{1.59}$$

The size of $\lambda = 0.2255 \pm 0.0019$ [7] permits us to neglect terms of $O(\lambda^4)$ or larger, and so in the SM the value of $-2\beta_{\rm s} \approx -2\eta\lambda^2 = (3.68 \pm 0.17) \times 10^{-2}$ is sufficiently small that deviations from $\phi_{\rm s} = -2\beta_{\rm s}$ are essentially deviations from zero- any new physics phases contributing to the mixing are likely to have a sizeable effect on this value. Similarly for $\Delta\Gamma$, where $\phi = 3.40^{+1.32}_{-0.77} \times 10^{-3}$ in equation (1.48) [12]. In the presence of new physics, Γ_{12} is unlikely to be affected as decays at tree level are robust to the existence of additional particles. The same can not be said of M_{12} however, where additional phases may enter through the box diagram. It is possible to imagine such an effect as a complex factor on M_{12}:

$$M_{12}' = M_{12} \times \Delta_s = M_{12}|\Delta_s|e^{i\phi_{NP}} \tag{1.60}$$

The additional phase then enters $\phi_{\rm s}$:

$$\phi_{\rm s}{}' = -2\beta_{\rm s} + \phi_{NP} \tag{1.61}$$

as well as quantities dependent upon ϕ, eg: $\Delta\Gamma_s$:

$$\Delta\Gamma_s{}' = 2|\Gamma_{12}|\cos(\phi + \phi_{NP}) \tag{1.62}$$

[4]The exclusive use of the symbol "ϕ" to describe phases is unfortunate as it leads to repetition. In LHCb $\phi_{\rm s}$ is used to describe the $\mathcal{CP}$ phase arising from interference between mixing and decay in $b \to c\bar{c}s$ transitions, which is *not* the same as $\phi = \arg(-M_{12}/\Gamma_{12})$ in equation (1.48).

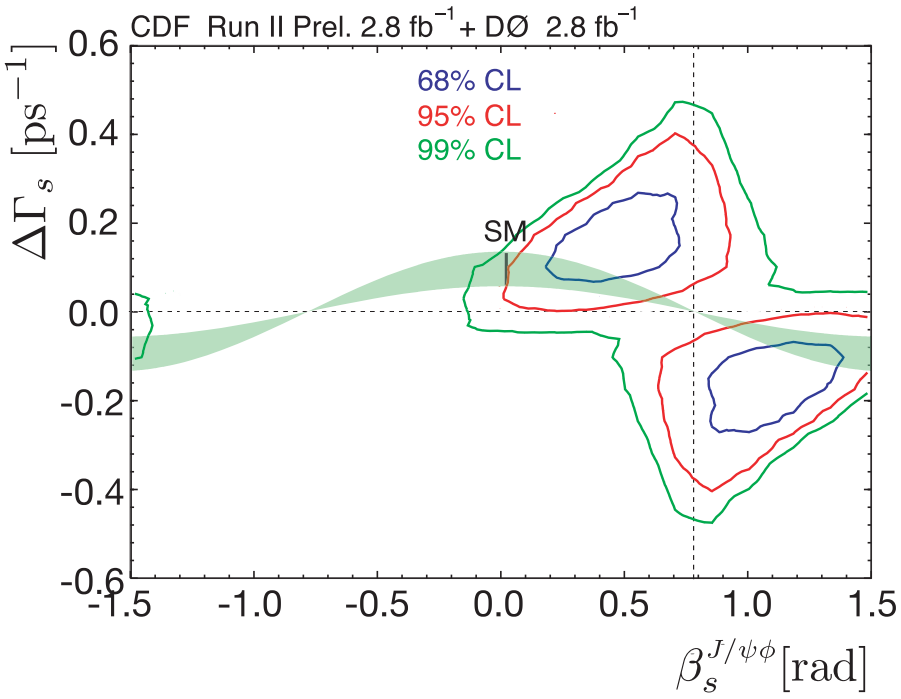


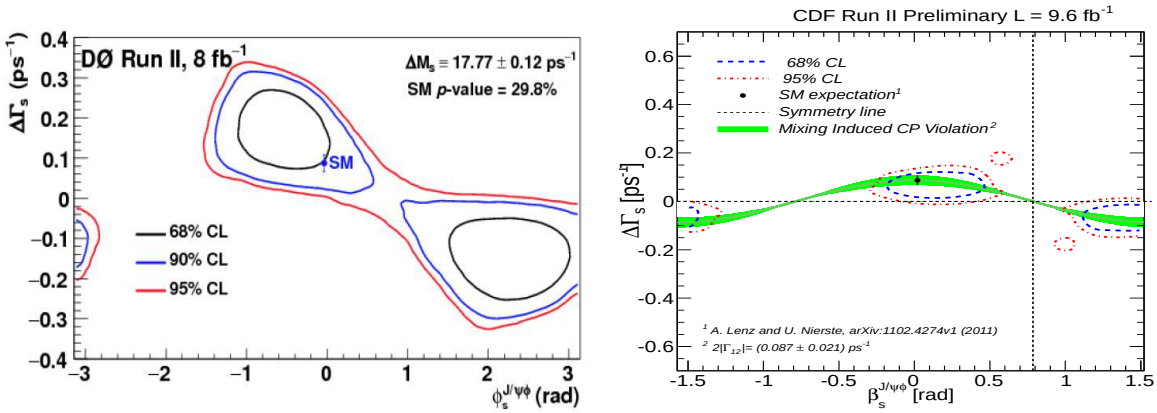
Figure 1.6.: Tevatron combined ϕ_s measurement as of Summer 2009. Shown is the confidence limits in the $\Delta\Gamma_s$, $\beta_s = -0.5\ \phi_s$ plane. The Standard Model value is also shown. The measurement deviates from the Standard Model value at 2.33 σ [13].

1.3.3. Previous ϕ_s measurements

The Tevatron experiments DØ and CDF have both studied ϕ_s throughout the course of run II. Early combined measurements with 2.8 fb^{-1} hinted at a $> 2\sigma$ significance deviation from the Standard Model expectation as shown in Figure 1.6 [13]. Since then a regime of updated measurements has decreased this significance. The most recent Tevatron results using the full Run II datasets are shown in Figure 1.7, where it can be seen that the measurements are consistent with the Standard Model at the 1σ level [14] [15]. The sensitivity of these measurements are still much poorer than the uncertainties on the Standard Model predictions however, leaving much room for improvement.

1.4. Conclusions

$\mathcal{CP}$-violation is a well understood process within the Standard Model that can be used as a test for new physics. Decay channels that permit measurement of ϕ_s are of great interest in the search for physics beyond the SM as these in particular are sensitive to sizeable deviations from the theoretically well-known value. Previous measurements of



(a) DØ, 8 fb^{-1} **(b)** CDF, 9.6 fb^{-1}

Figure 1.7.: (a) $\Delta\Gamma_s$, ϕ_s confidence limit contours from DØ with 8 fb^{-1} [14]. (b) $\Delta\Gamma_s$, $\beta_s = -0.5$ ϕ_s confidence limit contours from CDF with 9.6 fb^{-1} (Preliminary) [15]. Both results are consistent with the Standard Model expectation.

ϕ_s are inconclusive as they lack the sensitivity required to probe the Standard Model value with comparable uncertainties. Measurement of ϕ_s in $B_s^0 \to J/\psi\,\phi$ is therefore an ideal candidate for new physics discovery at LHCb.

Chapter 2.

The LHCb experiment

"Scientists have substituted mathematics for experiments, and they wander off through equation after equation, and eventually build a structure which has no relation to reality."

— Nikola Tesla

LHCb is the only experiment dedicated to heavy flavour at the LHC. This chapter will discuss heavy flavour experiments to date, offer a brief description of the LHC and describe in detail the LHCb detector. The final section of this chapter will describe the time alignment of the Ring Imaging Čerenkov detectors responsible for particle identification.

2.1. A brief history of heavy flavour collider experiments

The study of heavy flavour is not new, and with the recent announcements of the Super B [16] and Belle II [17] experiments LHCb may not be the last experiment dedicated to measuring the properties of beauty and charm mesons. While fixed-target experiments have been devised for heavy flavour measurements [18], this section will focus on collider experiments leading up to the LHC in order to set the scene for the LHCb experiment.

2.1.1. Early heavy flavour at the Υ(4S) resonance

The CLEO [19] detector on the $\sqrt{s} = 9 - 12$ GeV e^+e^- Cornell Electron-positron Storage Ring (CESR) started running at the end of the seventies, and throughout its 30-year lifetime had a successful heavy flavour programme. Early measurements together with the CUSB experiment centered on measurements around the Υ system, including the discovery of the Υ(4S) [20], which later became the workhorse of the b-factories. The Υ(4S) was observed to be a broad resonance, suggesting that it was above the B production threshold. Compelling though indirect evidence for the production of B mesons came from the abundance of leptons in the final state, a telltale sign of semileptonic decays of heavy mesons [21, 22, 23]. Direct evidence for B $\overline{\mathrm{B}}$ mesons coming from the Υ(4S) took a little longer, and was observed in 1983 [24]. The mid-to-late eighties was a particularly active time in heavy flavour, spurred by competition between CLEO and CUSB at CESR and the ARGUS [25] experiment at DESY. While CLEO laid claim to the discovery of B mesons and the $\mathrm{D_s}$ meson, ARGUS made the first measurement of $\mathrm{B^0}$-$\overline{\mathrm{B}}^0$ mixing by comparing like-sign pairs of semileptonic B decays [26].

2.1.2. The $\mathrm{Z^0}$ resonance

The late eighties and early nineties were a busy time for particle physics in general, with a number of active experiments. The $\mathrm{W^\pm}$ and $\mathrm{Z^0}$ discoveries at the SPS in the early eighties [27] set in motion the design and construction of the e^+e^- Stanford Linear Collider (SLC) [28] and Large Electron-Positron collider (LEP) [29, 30] as $\mathrm{Z^0}$-factories to further study its properties. The $\mathrm{Z^0}$ resonance is conducive to heavy flavour as it decays to two b quark jets 15% of the time [7]. These are produced back-to-back and the large $\mathrm{Z^0}$ mass leads to a significant boost imparted to the decay products, resulting in secondary vertices displaced by $\mathcal{O}$(mm). The SLD [31] experiment at the SLC soon started to measure the properties of B mesons, followed and somewhat outclassed by LEP [32]. The first evidence of Λ_b baryons came from the Aleph experiment in 1991 [33], and a year later Aleph and OPAL presented the first evidence of the existence of the $\mathrm{B_s^0}$ meson [34] [35]. In later years the LEP experiments combined measurements of $\mathcal{CP}$ asymmetries and cross-section ratios [32], and the first reasonable limits on $\mathrm{B_s^0}$ mixing were set in combination with SLD [36].

2.1.3. Heavy flavour at a hadron collider

The CDF [37] experiment on the $p\overline{p}$ Tevatron collider at Fermilab started taking data in 1988, followed in 1992 by the DØ experiment at the Tevatron's second interaction point. While e^+e^- colliders are well suited to precision measurements, discovery is more likely at proton colliders as cross-sections are much higher. The Tevatron experiments have been a remarkable success in this aspect, discovering the top quark [38], measuring B_s^0 mixing [39] and putting the strongest constraints on the mass of the Standard Model Higgs prior to the first year of LHC data [40]. The heavy flavour programmes at both CDF and DØ turned to studying ϕ_s and the rare decay $B_s^0 \to \mu^-\mu^+$ [41]. The first measurements of ϕ_s from the Tevatron showed tantalising hints of new physics in $B_s^0 \to J/\psi\phi$, reporting a 2.1σ deviation from ${\phi_s}^{SM} = -2\beta_s$ but updated measurements have decreased the significance of this deviation.

2.1.4. The B-factories

The $\Upsilon(4S)$ resonance is ideally suited to the study of B mesons, as it decays exclusively to $B\overline{B}$ and B^+B^- pairs. An e^+e^- collider running at this resonance is a copious source of B mesons as demonstrated by CLEO, but the kinematics are not ideal for time-dependent measurements as the mesons are produced almost at rest. The solution to this and the defining aspect of the B-factory design is the use of asymmetric beam energies [42], where the $\Upsilon(4S)$ is produced with a significant boost relative to the lab frame, resulting in B_d^0 mesons whose lifetime may be inferred from their flight distance. The PEP-II collider at SLAC took high-intensity beams from the SLC of 9 GeV e^- and 3.1 GeV e^+ and circulated them in physically stacked storage rings. These were brought into collision at a single interaction point within the *BABAR* detector [43]. A similar design was used at KEK with the 8 GeV e^-, 3.5 GeV e^+ KEK-B accelerator and Belle detector [44]. Due to the clean e^+e^- environment and high luminosities, large, pure samples of B meson candidates were collected by both collaborations, with the *BABAR* dataset corresponding to 468×10^6 B $\overline{\text{B}}$ pairs [45] and 657×10^6 at Belle [46], allowing precise study of $\mathcal{CP}$ in the B system in addition to probing rare decays and the first observation of a number of new particles.

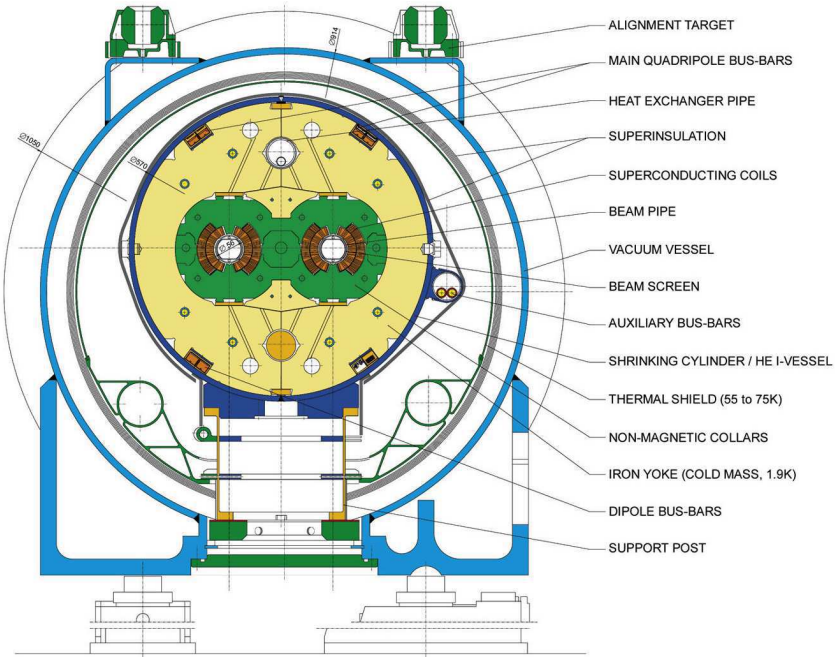


Figure 2.1.: Cross-section of an LHC dipole magnet with cold mass and vacuum chamber [47].

2.2. The LHC

The Large Hadron Collider (LHC) is a proton-proton (pp) and heavy ion collider replacing LEP as CERN's flagship facility [49]. While LEP was concerned with precision electroweak measurements, the LHC is a discovery machine and has been built to collide protons with a design luminosity of $1 \times 10^{34}\text{cm}^{-2}\,\text{s}^{-1}$ at centre-of-mass energy 14 TeV. The LHC collides bunches of protons separated by multiples of 25 ns, corresponding to a 40MHz collision rate which is a multiple of the 400MHz RF frequency used to accelerate the beams. The nominal bunch intensity is 1×10^{11} protons. The LHC requires two apertures in order to accelerate the like-charged proton/ion bunches in opposite directions. The bending dipoles of the LHC are a novel design in order to accommodate the twin apertures, sharing the same iron yoke as shown in Figure 2.1. The NbTi superconducting magnet technology requires that the cold mass of the dipoles are maintained at 1.9K in order to operate at the 8.33T required to bend the 7 TeV proton beams. Due to the large stored energy per dipole, an elaborate quench protection system has been developed [50]. The LHC relies upon a large and complex injector chain, re-using much of that exploited by LEP but modified where necessary to accept protons and lead ions as shown in Figure 2.2. The injector chain ends with bunch trains consisting of several equidistant bunches of protons at 450 GeV injected into the LHC via transfer

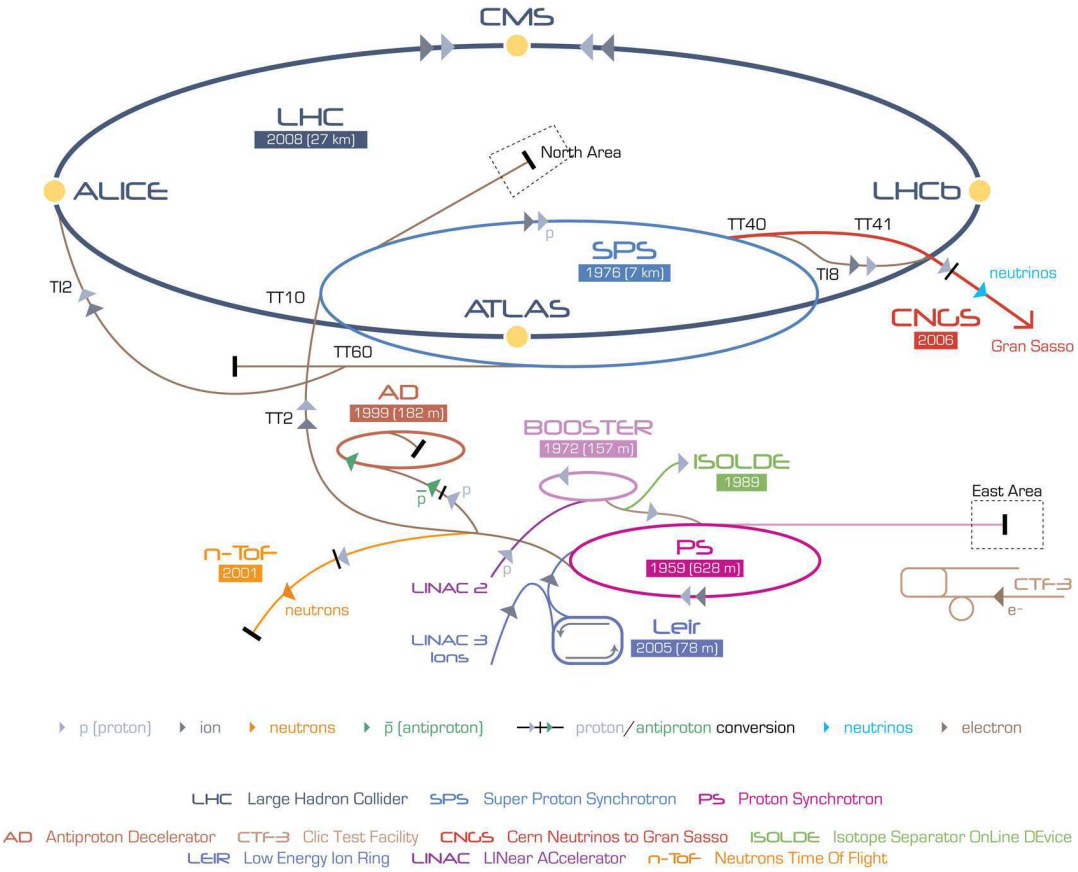
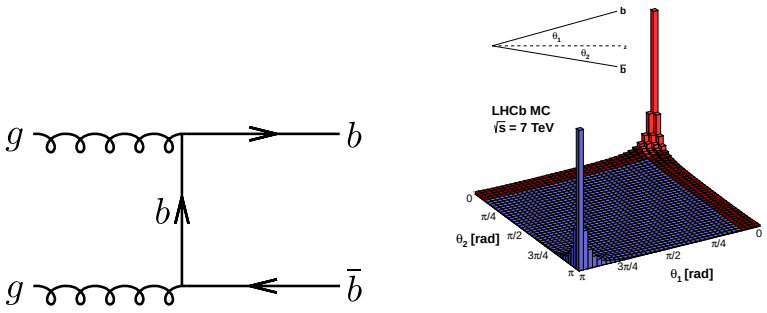


Figure 2.2.: The CERN Accelerator complex, showing the injection chain of the LHC and the interaction points at which the experiments are installed [48].

lines from the SPS. In September 2008 the LHC saw its first circulating beam at injection energy. Shortly after, during tests of the magnet ramping a catastrophic failure due to a quench in one of the current-carrying bus bar interconnects between magnets put the machine out of action until October 2009 [51]. During this time inadequacies in the quench protection system were discovered that led to the decision to limit initial operations to 3.5 TeV per beam, pending a long shutdown during 2013-2014 in which additional quench protection and liquid helium venting systems will be added. In November 2009 beams were successfully injected at 450 GeV and collided at each of the interaction points. The beam energy was also successfully ramped to 1.18 TeV giving a centre-of-mass collision energy of 2.36 TeV [52]. Operations at $\sqrt{s} = 7$ TeV began in earnest in February of 2010, with 7 TeV collisions occurring for the first time in late March of the same year. Since then machine development has focused on increasing intensity, with peak luminosity increasing first by way of higher bunch currents, then by the addition of more colliding bunches and lastly with higher bunch densities at the interaction point. The LHC collides particles at four interaction points around the main ring. At these interaction points are four main experiments, in addition to a number of smaller experiments dedicated to forward physics. The four main experiments are ATLAS, CMS, ALICE and LHCb [53, 54, 55, 56]. The first two are general purpose detectors with an emphasis on discovery. The final two are more specialised: ALICE will focus on lead-ion collisions and LHCb on precision heavy flavour and rare decays.

2.3. b-production at the LHC

The centre-of-mass energy of the LHC is large enough to produce the full range of bottom and charm mesons. The dominant production mode is through gluon-gluon fusion, shown in Figure 2.3a. Gluon-gluon fusion favours a large asymmetry between the momenta of the incident gluons, which at LHC energies will have a large momentum with respect to the mass of the $b\overline{b}$ pair. This results in the $b\overline{b}$ pair being boosted along the beam axis as observed in the lab frame. The effect of this is that $b\overline{b}$ production at the LHC is confined to a pair of cones pointing to the interaction point, extending outward in the direction of the beam axes as shown in Figure 2.3b. At the LHC the b-production cross-section is large, and is approximately linear over the range $1 \rightarrow 14$ TeV as shown in Figure 2.4. The LHC is the world's most copious source of B-mesons, with a measured $b\overline{b}$ production cross-section of $284 \pm 20 \pm 49$ µb at $\sqrt{s} = 7$ TeV [57].



(a) gluon-gluon fusion

(b) $b\bar{b}$ pair polar angle distribution

Figure 2.3.: Dominant production of $b\bar{b}$ pairs at the LHC proceeds via gluon-gluon fusion, which favours asymmetric parton momentum. This results in $b\bar{b}$ pairs with a correlated polar angle distribution [56].

2.4. The LHCb Detector

The LHCb Experiment takes advantage of the large $b\bar{b}$ production cross-section at the LHC to perform dedicated precision heavy flavour measurements. The detector geometry is optimised for the production angles shown in Figure 2.3b, adopting a single-arm spectrometer configuration to maximise detector resolution within the forward cone. The LHCb experimental layout and cavern infrastructure is shown in Figure 2.5. The co-ordinate system adopted for the LHCb experiment is right-handed with the y-axis pointing upwards and the x-axis pointing towards the center of the LHC ring. The horizontal bending plane covers the range $10 - 300$ mrad, while the vertical non-bending plane has a coverage extending to 250 mrad [59]. The choice of a single-arm configuration is purely pragmatic: The LHCb experiment inhabits the former Delphi cavern, UX85 at Intersection Point 8. To maximise precision the full cavern length is dedicated to a single spectrometer arm. This requires a modification to the LHC beam optics that moves the interaction point 11.25 m from the nominal cavern center.

2.4.1. Beampipe

The beampipe maintains a hermetic vacuum in which the LHC beams travel through the LHCb experiment. The LHCb experiment operates in the high-rapidity regime, with particles exiting the beampipe at a shallow angle: In order to minimise the number of secondary particles produced in material interactions care was taken in the design of the

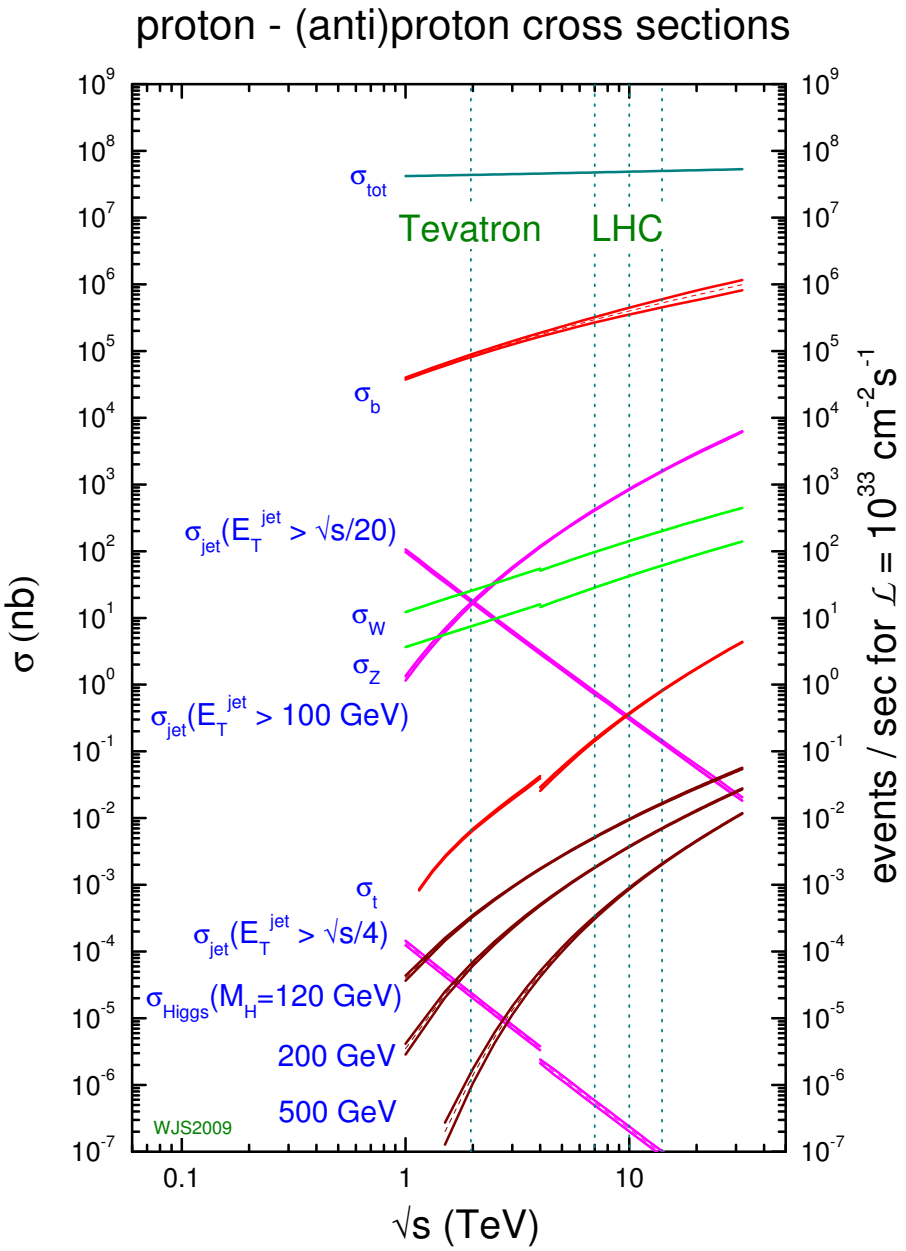


Figure 2.4.: Cross-sections for $p\bar{p}$ and pp processes as a function of center-of-mass energy [58]. Discontinuities are a result of the change from $p\bar{p}$ in the Tevatron regime to pp in the LHC regime.

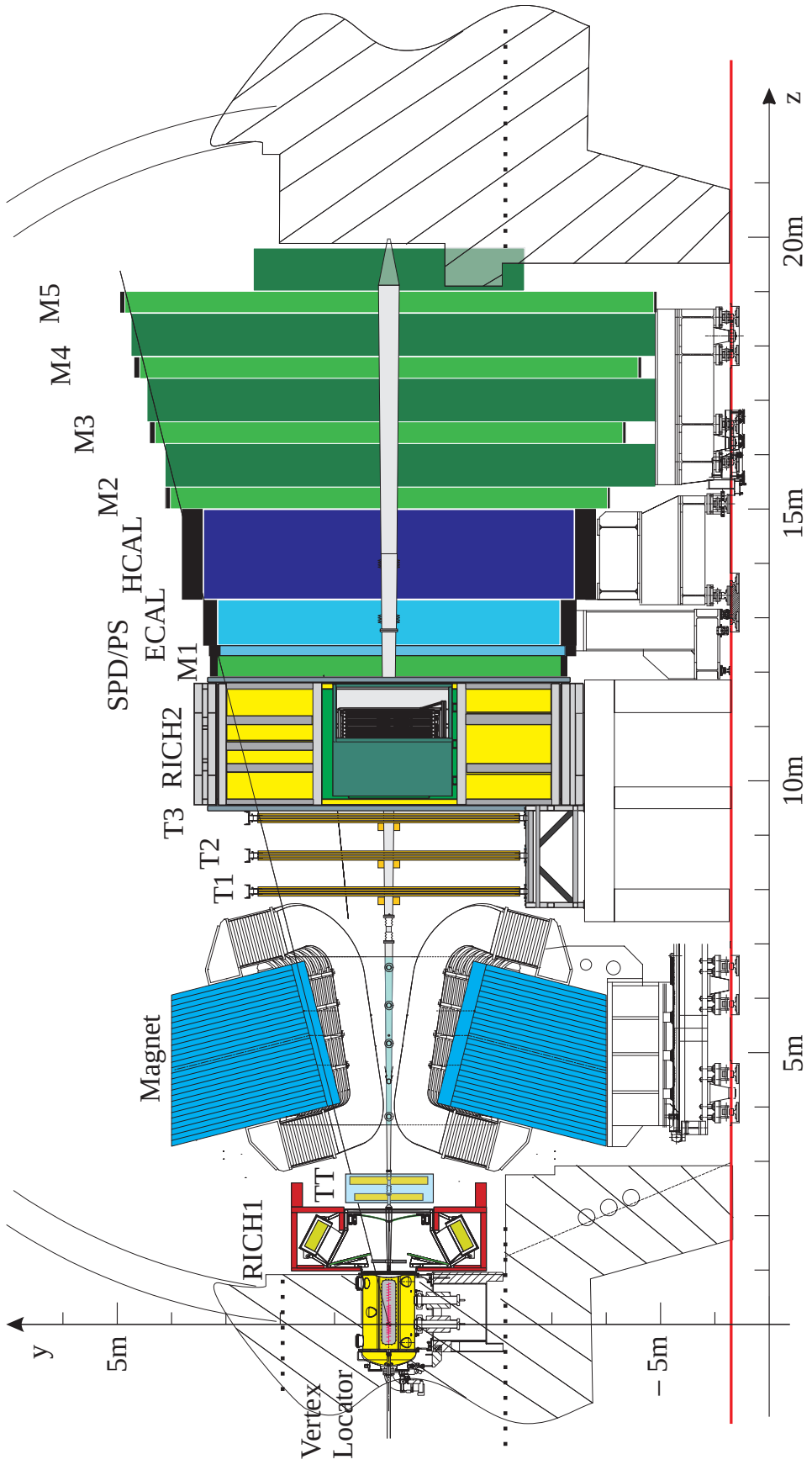


Figure 2.5.: The LHCb detector, side view.

beampipe to reduce the beampipe material traversed by particles from the interaction point. The beampipe traverses the length of the LHCb experiment starting at the vertex locator. It consists of 4 sections of which the first three nearest the interaction region are made of machined beryllium. The fourth section is made of stainless steel, starting at the Electromagnetic Calorimeter.

2.4.2. Magnet

The LHCb experiment uses a warm dipole magnet to facilitate momentum measurements of charged tracks. A warm magnet design has been chosen due to the cost and time constraints associated with the use of a superconducting magnet. The magnet defines the detector acceptance and is designed such that the field must be less than 2 mT in the region of the photon detectors of the Ring Imaging Čerenkov detectors to minimise distortion of the electrostatic focusing. The magnet as shown in Figure 2.6 consists of an iron yoke made from 100 mm sheet steel, separable into top, bottom and side sections for transportation purposes. The coils are manufactured from hollow extruded aluminium whose central channel acts as a conduit for cooling water. Each coil consists of fifteen "pancakes" wound from the conductor and formed into a conical saddle shape, held in place by aluminium clamps that permit thermal and magnetic movement. The nominal current and power dissipation are 5.85 kA and 4.2 MW respectively, resulting in a measured $\int Bdl$ of 3.615 Tm within the region $z = 2.50 - 7.95$ m. The field has been extensively mapped using an array of Hall probes across all three axes to a relative precision of about 4×10^{-4}.

2.4.3. Tracking

Tracking in the forward region in the LHC environment is both a challenge and an important component of any precision heavy flavour experiment. In LHCb, tracking is provided by four discrete subdetectors illustrated in Figure 2.7 using a variety of technologies. Tracking information is relied upon not only for momentum measurements but also to provide tracks to which calorimeter clusters and Čerenkov rings may be matched. The tracking system consists of the Vertex Locator (VELO) and Tracker Turicensis[1](TT)

[1]tu.ri.cen'sis. L. adj. *turicensis* of Turicum, the Latin name for Zurich. Previously called the Trigger Tracker, but changes to the trigger architecture meant that the L1 trigger was removed. In order to preserve the acronym the name was altered to refer to the institute having the largest involvement in the TT.

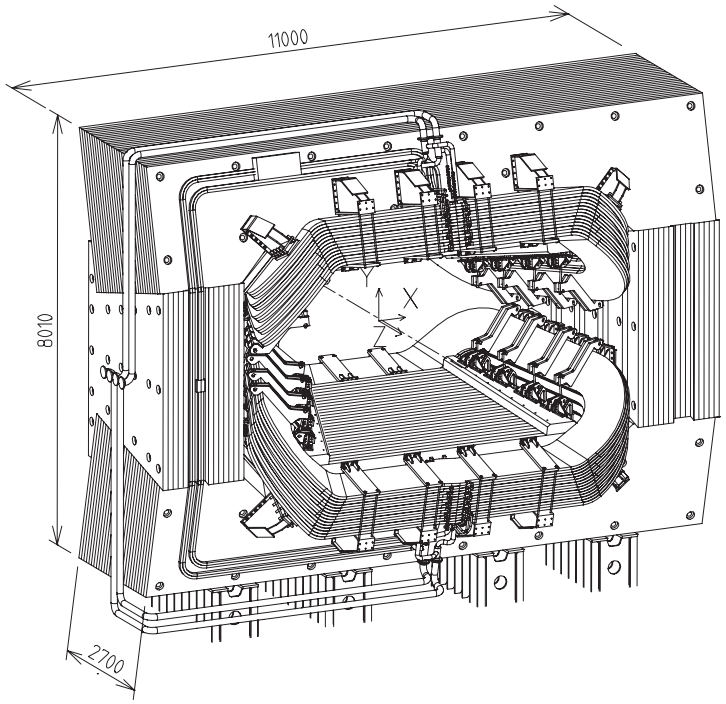


Figure 2.6.: The LHCb Magnet and supporting infrastructure (dimensions in mm) [60].

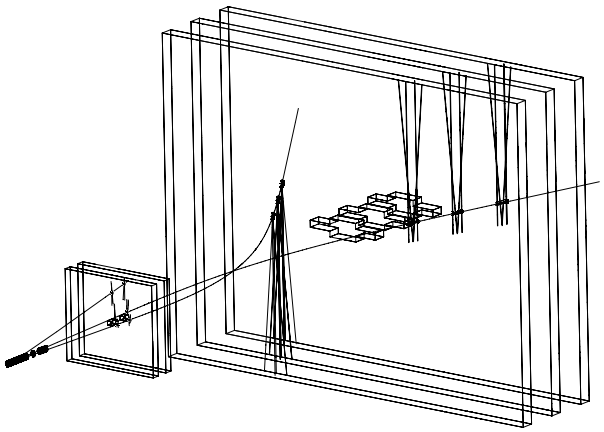


Figure 2.7.: The LHCb tracking subdetectors modelled in GEANT showing a low-multiplicity event in terms of tracks and hits. From left to right can be seen the VELO, TT and tracking stations 1-3.

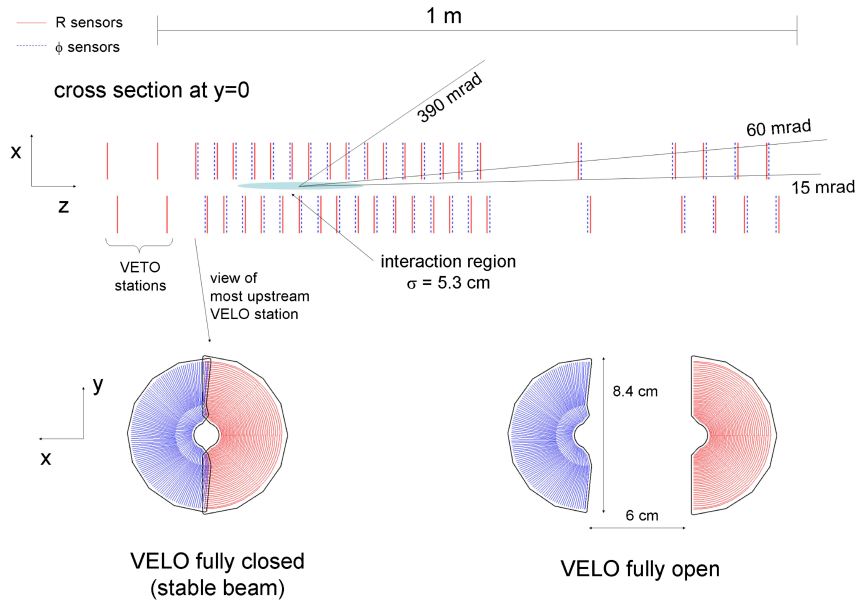


Figure 2.8.: VELO station layout, showing r, ϕ modules and acceptance [59].

upstream of the magnet, with three tracking stations T1-T3 downstream. The downstream tracking stations each consist of two regions which employ different detector technologies: The Inner tracker (IT) close to the beampipe uses silicon strip detectors to handle the higher occupancies while the Outer tracker (OT) uses straw-tubes and extends to cover the full LHCb acceptance.

Vertex Locator

A defining feature of B-hadrons at the LHC is a displaced decay vertex of $\mathcal{O}$(cm) from their production vertex. Excellent vertex resolution is required in order to precisely measure secondary vertices and is a critical component in reducing background. The B_s^0-$\overline{B}_s^0$ mixing frequency is fast, requiring propertime resolutions of $\mathcal{O}$ (50 fs) in order to measure time-dependent $\mathcal{CP}$. At LHCb the VErtex LOcator (VELO) provides tracking within the full detector acceptance close to the interaction region designed to meet these requirements. The VELO consists of 21 stations surrounding the beam axis, 6 of which are upstream of the nominal interaction point. The VELO sensor geometry is an important part of the design: A cylindrical geometry is preferable as it permits fast reconstruction in the higher-level trigger software. Each station consists of two modules, one each to the left and right-hand sides of the beam axis. Each module consists of two

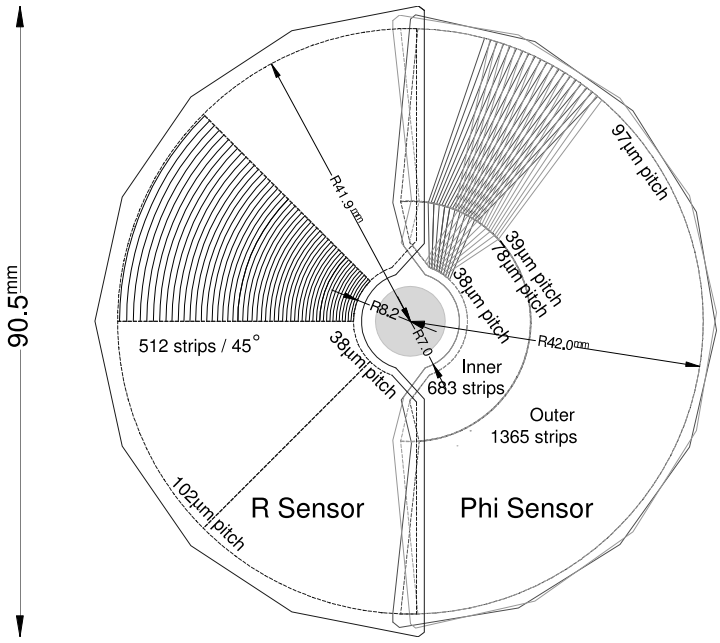


Figure 2.9.: VELO sensor geometry. Two successive ϕ sensors are shown overlaid to highlight the stereo angle design. Due to the bonding requirements of the r sensor pads the sensor is slightly larger. The overall sensitive area is however the same as that of the ϕ sensor [59].

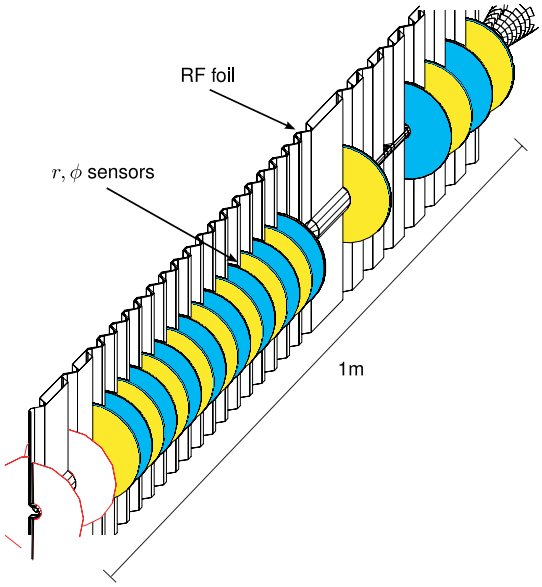
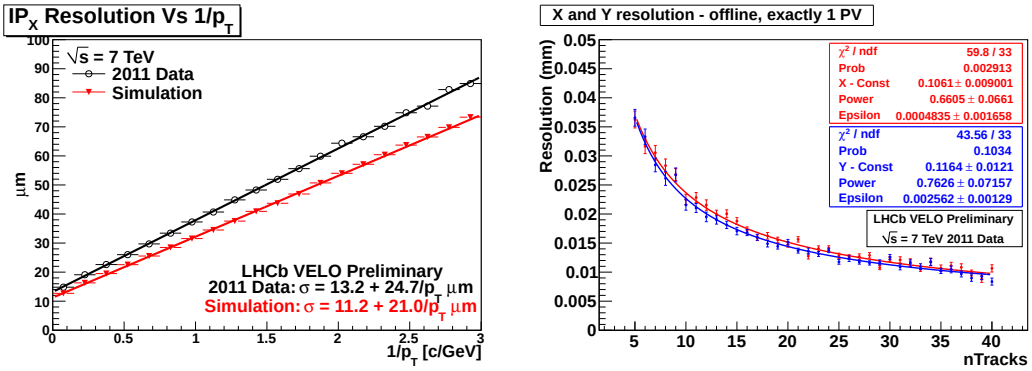


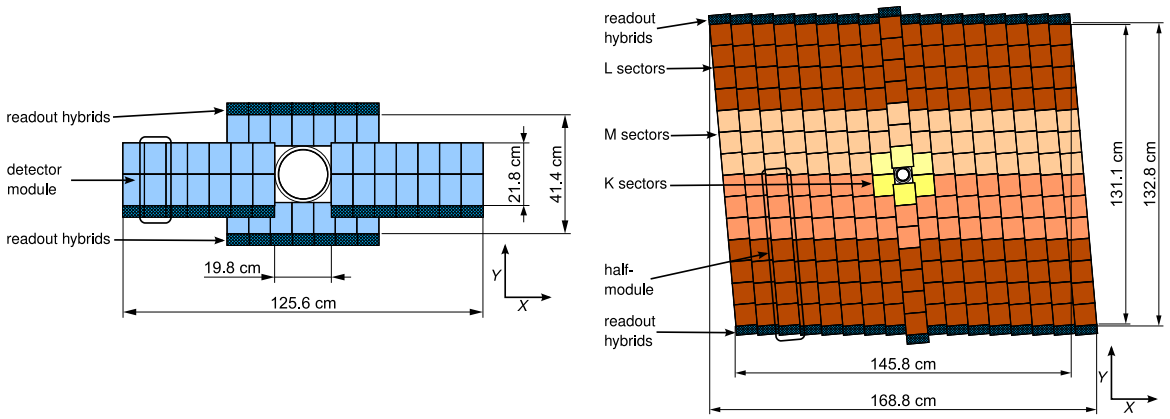
Figure 2.10.: The VELO RF foil suppresses radio frequency pickup within the VELO sensors from the beam and partitions the VELO vacuum from the LHC vacuum.



(a) σIP_x as a function of $1/p_{\mathrm{T}}$ **(b)** σPV as a function of reconstructed tracks

Figure 2.11.: (a): Impact Parameter resolution in x as a function of $1/p_{\mathrm{T}}$ for 2011 data compared to 2010 Monte-Carlo simulation. (b): Primary Vertex resolution as a function of the number of reconstructed tracks in x and y for events containing precisely one reconstructed primary vertex.

silicon strip sensors, one with radial (r) segmentation, the other with segmentation in ϕ. In addition to these 21 stations, another two stations consisting exclusively of 4 r sensors are situated further upstream for pileup veto. The module arrangement is such that a minimum of 4 modules are traversed by tracks originating at the interaction region and making an angle to the beam axis of less than 390 mrad. The r sensor consists of 512 concentric semicircular strips subdivided into four regions covering $\pi/4$ rad each in order to minimise strip capacitance and occupancy. The strip pitch decreases from the outside edge towards the beam axis so that track measurements contribute equally to impact parameter resolution. The ϕ sensor modules measure the coordinate orthogonal to the r sensor, and are divided into two sensitive regions for similar reasons to the r sensor regarding occupancy; the strip pitch of the outer region is roughly half that of the inner region, with 683 strips in the inner region and 1365 strips in the outer region. The ϕ strips are also not exactly radial in the x, y plane: A skew is introduced at 10° to the nominal axis in the inner region and an opposing 20° skew in the outer region. The skew angle is reversed in alternating ϕ sensors to provide a stereo track hit reconstruction with some sensitivity in r, permitting fast reconstruction in the trigger. The geometry of the sensors is shown in Figure 2.9. A radio-frequency (RF) box separates the LHC beam from the VELO modules to minimise RF pickup in the stations. In order to minimise the material traversed by particles coming from the interaction point the RF box is kept under vacuum which also keeps the thickness of the RF foil around the beam low as it does not need to tolerate a large pressure gradient. The RF foil is corrugated and the modules corresponding to one station are shifted by 1.5 cm to permit overlap between



(a) IT plane **(b)** TT plane

Figure 2.12.: Schematic diagrams of single planes of the Silicon Tracker. (a) One plane of the IT. (b) A single stereo layer of the TT, in this case the v-layer.

the two halves of the VELO. During data taking the VELO modules are approximately 8 mm from the beam but during injection the LHC aperture is larger than that of stable beam conditions, requiring that the VELO is retractable to a safe distance until stable conditions are reached. The halves of the VELO are retractable to a distance of 3 cm from the nominal beam axis. The ability to move the VELO halves independently also permits the sensors to surround the centre of the beam axis which is expected to deviate slightly from the nominal. The VELO impact parameter resolution in 2011 data is shown as a function of $1/p_{\mathrm{T}}$ in Figure 2.11a and compared to Monte-Carlo. Figure 2.11b illustrates the vertex resolution dependence upon the number of tracks in the vertex.

Tracker Turicensis and Inner Tracker

The Tracker Turicensis (TT) and Inner Tracker (IT) illustrated in Figure 2.12 share the same silicon sensor technology and are collectively referred to as the Silicon Tracker (ST). Each ST station consists of 4 layers of sensors. The silicon strips in the first and last "x" layers are aligned vertically while the second and third "u,v" layers are rotated by a stereo angle of +5 and -5 degrees. The strip pitch of each layer is 200 µm. The TT has a gap of $\approx$ 27 cm between the two stereo planes in order to improve spatial resolution while the separation between layers in the IT is $\sim$4 cm in each station. Common to both the TT and IT are the readout, control and monitoring systems. Readout proceeds through a readout hybrid connected by wire-bonded kapton ribbon cable to the silicon sensor.

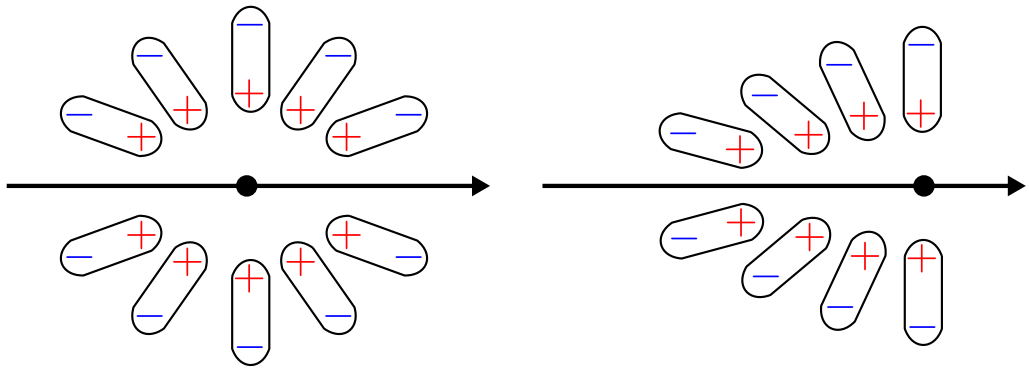
One or two sensors share a hybrid in the IT. In the TT balancing is performed such that higher occupancy regions have increased readout density: Each detector plane consists of half-modules containing seven silicon sensors sharing either two or three hybrids. Sensitive regions of the TT are labelled l, m, k, with sensors in the highest-occupancy k sector having exclusive use of a single hybrid, sensors in m share a hybrid between either 2 or 3 and sensors in l between 4. This configuration is achieved by mounting the sensors in strips of half the TT height called *half-modules* in which seven sensors and either two or three hybrids are mounted, depending upon proximity to the beampipe in the x axis as shown in Figure 2.12b.

Outer Tracker

The Outer tracker (OT) is a drift detector composed of modules containing two densely packed planes of straw tubes, staggered in order to ensure overlap. The tubes are gas-tight and filled with a 70:30 Argon-Carbon Dioxide mix providing sub-50 ns drift times in order to restrict spillover to no more than two bunch crossings. The OT geometry is similar to that of the IT, four layers of modules arranged in the same $x{-}u{-}v{-}x$ geometry where u, v planes are offset at -5 and $+5$ degrees to provide stereo hit resolution. The only tracking in y is provided by the stereo layers. The tracking acceptance of the OT extends out to the full 300 mrad horizontal and 250 mrad vertical acceptance and in to the IT boundary where the occupancy is less than 10% at nominal LHCb luminosity.

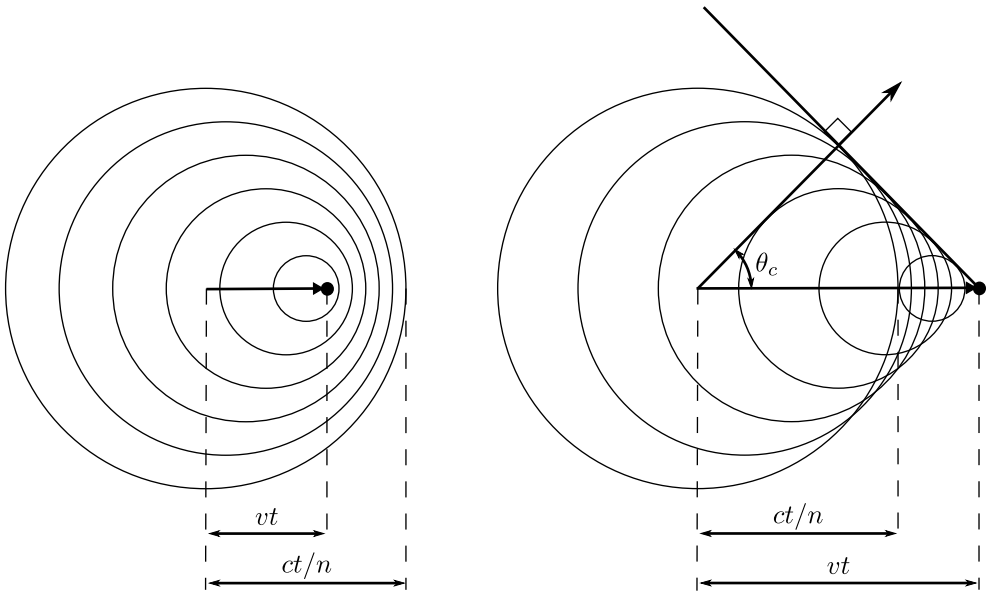
2.4.4. RICH System

The abundance of final states accessible to B-meson decays mean that a number of topologically similar decays are only distinguishable by their final state particles, for example the decay $\mathrm{B} \to \mathrm{D}^{\pm}\mathrm{K}^{\mp}$ which is used to determine the $\mathcal{CP}$ angle γ is subject to pollution from $\mathrm{B} \to \mathrm{D}^{\pm}\pi^{\mp}$ which is an order of magnitude more abundant. The need to identify these specific final states requires an emphasis on charged Particle Identification (PID) in heavy flavour experiments. At LHCb $\pi^{\pm} - \mathrm{K}^{\pm}$ separation is achieved using Ring-Imaging Čerenkov (RICH) detectors.



(a) $v < c/n$ **(b)** $v > c/n$

Figure 2.13.: Čerenkov emission from a charged particle in a dielectric, dipole arrangement. (a): For $v < c/n$ the induced dipole arrangement is symmetric and no net dipole arises. (b): For $v > c/n$ a causal connection between induced dipoles can no longer exist, and so a net dipole is created. The resulting spontaneous emission of photons is the Čerenkov effect.



(a) $v < c/n$ **(b)** $v > c/n$

Figure 2.14.: Charged particle propagation in a medium of refractive index n, in the Huygens construction. (a): The particle travels at subluminal velocity in the medium. Wavelets are incoherent and so no wavefront is produced. (b): The particle is superluminal within the medium, and constructive interference from successive wavelets results in a wavefront emitted in the direction of the Čerenkov angle θ_c [61].

The Čerenkov effect

Charged particles traversing a dielectric of refractive index n with a velocity v greater than the local speed of light c/n polarise the material in such a way as to create a net dipole (Figure 2.13) which radiates photons in a cone along the track as shown in Figure 2.14. The angle of emission is given to first order by equation (2.1) in which $\beta = v/c$ and θ_c is known as the Čerenkov angle [62]. A knowledge of n combined with a measurement of θ_c can therefore yield a measurement of β for a track traversing a radiator.

$$\cos\theta_c = \frac{1}{n\beta} \tag{2.1}$$

$v = c/n$ is a threshold for Čerenkov radiation, at which point the radiated Čerenkov photons will all be in the $\theta_c^{\text{thr}} = 0$ (forward) direction. Below this threshold Čerenkov radiation does not occur. The photon intensity and spectrum of the emitted radiation is given by the Frank-Tamm relation (equation (2.2a)) where dN_{ph} is the number of photons with energy $E \to E + dE$, α is the fine structure constant, Z and L are the particle's charge and average track length in the medium. Substituting equation (2.1) into this, we obtain the (β, n) spectral dependence given by equation (2.2b).

$$\frac{dN_{ph}}{dE} = \left(\frac{\alpha}{\hbar c}\right) Z^2 L \sin^2\theta_c \tag{2.2a}$$

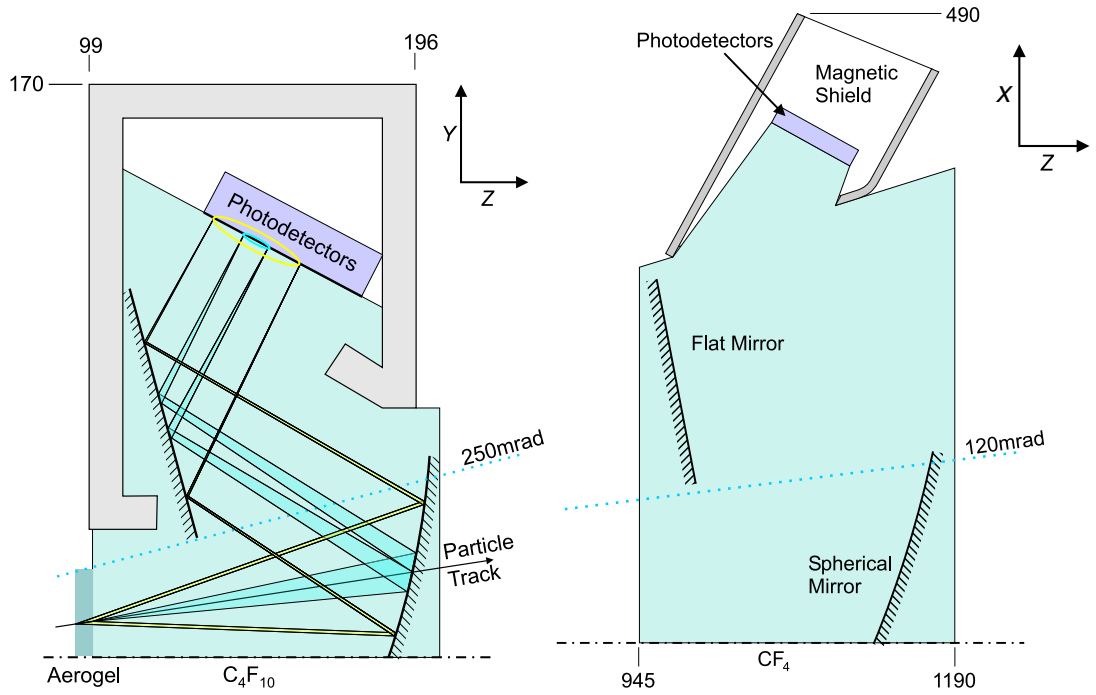
$$= \left(\frac{\alpha}{\hbar c}\right) Z^2 L \left[1 - (1/n\beta)^2\right] \tag{2.2b}$$

The LHCb RICH detectors

By reflecting the Čerenkov radiation emitted continuously along a track with a spherical mirror, all Čerenkov photons are imaged onto a single ring at the mirror's characteristic focal length. This is the defining feature of a Ring Imaging Čerenkov (RICH) detector. The ring radius determines the Čerenkov angle and thus β. Integrating equation (2.2a) over the energy bandwidth ΔE subject to the approximation that $n\beta$ is constant the number of detected Čerenkov photons as a function of Čerenkov angle is:

$$N = N_0 Z^2 L \sin^2\theta_c \tag{2.3}$$

With N_0, the detector response parameter given by equation (2.4) in which Q, T, R are the average detector efficiencies due to Quantum, Transmission and mirror Reflection



(a) RICH 1 in the y, z plane **(b)** RICH 2 in the x, z plane

Figure 2.15.: Schematic layout of the LHCb RICH detectors (one half shown, with beam axis denoted by the dashed line).

effects respectively.

$$N_0 = \left(\frac{\alpha}{\hbar c}\right) \int_E^{E+\Delta E} (QTR) dE' \tag{2.4}$$

The LHCb RICH system consists of two distinct detectors and three Čerenkov media (radiators) in order to cover the full range of acceptances and particle momenta. The geometry of the individual RICH detectors are shown in Figure 2.15 while their position with respect to the other subdetectors of the LHCb experiment is made clear in Figure 2.5. The RICH detectors of the LHCb experiment contain two sets of mirrors: The spherical focusing mirrors for ring-imaging and an additional set of flat mirrors to direct light onto the photon detector planes situated outside of the detector acceptance. RICH 1 consists of a gas enclosure containing a silica aerogel radiator of thickness $L = 5$ cm and refractive index $n = 1.03$, in addition to C_4F_{10} gas of $L = 85$ cm, $n = 1.0014$. RICH 2 is a gas enclosure of length $L = 167$ cm filled with CF_4 of refractive index $n = 1.0005$. RICH 2 provides $\pi^{\pm} - \mathrm{K}^{\pm}$ separation over the range $\sim 20 - 100$ GeV/c. The photon detectors of the RICH provide single-photon sensitivity with a spatial resolution better than 2.5 mm, with deadtime and sampling rates compatible with the LHCb bunch

crossing. RICH 1 is downstream of the VELO and provides $\pi^{\pm} - \mathrm{K}^{\pm}$ separation up to a momenta ~ 50 GeV/c. A number of prototype technologies were investigated [63, 64, 65] from which the Pixel Hybrid Photon Detector (HPD) shown in Figure 2.16 was chosen based upon a cost/risk analysis. HPDs are a novel design, combining a silicon pixel detector and integrated readout electronics with a vacuum photocathode. The HPDs of the RICH convert incident photons to electrons through a quartz window whose inner surface is coated with a thin S20 type multi-alkali photocathode, sensitive over the wavelength range $200 - 700$ nm. The liberated electrons are accelerated through a -20 kV potential and focused/demagnified onto a silicon sensor by way of additional electrodes within the HPD body. The sensor chip is 300 µm silicon divided into 8192 pixels 500×62.5 µm in size, reverse-biased to 80 V. The silicon sensor is bump-bonded to a binary readout chip which converts above-threshold charge in the pixels into a binary hit. This readout chip was designed in collaboration with the ALICE experiment and has two modes of operation: Full readout of all 8192 pixel hits, and an LHCb specific mode which performs a logical OR of 8 neighbouring pixels to form an array of 1024 pixels of size 500×500 µm^2 The demagnification factor of the electrostatic focussing is 5, resulting in an effective granularity of 2.5×2.5 mm^2 at the photocathode for each superpixel, satisfying the spatial resolution requirements.

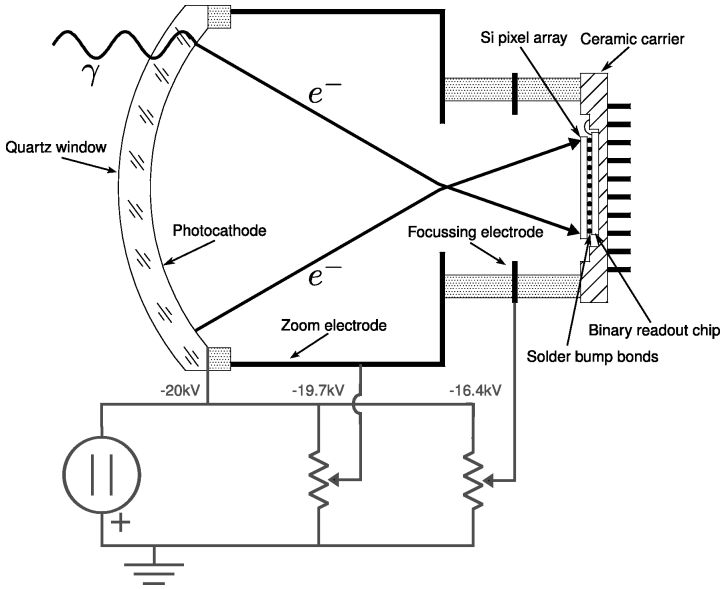


Figure 2.16.: Schematic diagram of the Pixel Hybrid Photon Detector.

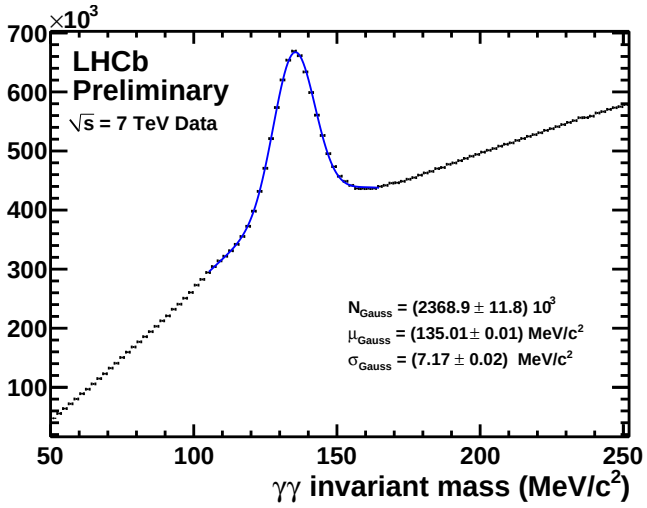


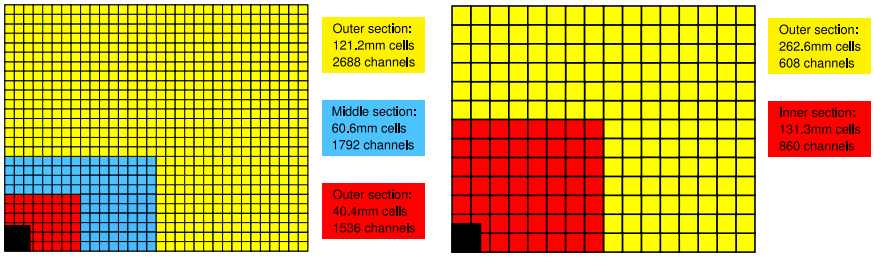
Figure 2.17.: $\pi^0 \to \gamma\gamma$ selected in the 2011 dataset using the electromagnetic calorimeter.

2.4.5. Calorimeters

The LHCb calorimeters provide particle identification for electrons, photons and hadrons as well as energy and position measurements. They are also a major part of the Level-0 trigger where they are used to select electron, photon and hadron candidates above a specified transverse energy threshold. LHCb uses a dual calorimeter structure consisting of an electromagnetic calorimeter (ECAL) followed by a hadron calorimeter (HCAL). The ECAL is supplemented by two additional detectors, the Scintillating Pad Detector (SPD), and preshower detector (PS). The calorimeters are segmented in the $x-y$ plane such that the channel density increases towards the beampipe in order to provide roughly constant angular resolution as shown in Figure 2.18. The ECAL, PS and SPD are scaled such that they have a projective channel geometry as seen from the interaction point. This makes combining measurements across the three subdetectors straightforward.

Scintillating Pad and Preshower Detectors

The Scintillating Pad Detector (SPD) consists of 15 mm thick scintillating tiles placed after the first set of muon chambers downstream from RICH 2. The SPD is sensitive to charged particles and therefore distinguishes between electrons and photons before showering occurs. Immediately after the SPD is a 12 mm thick layer of lead with radiation length 2.5 X_0, followed by another layer of scintillating tiles. This second layer of tiles in combination with the lead forms the Preshower detector (PS), which distinguishes between electromagnetic particles and charged hadrons by exploiting their



(a) ECAL **(b)** HCAL

Figure 2.18.: Channel density and module dimensions for the ECAL and HCAL. The top right-hand quarter of each calorimeter is shown, with the uninstrumented area in the bottom left corner indicating the beam pipe position.

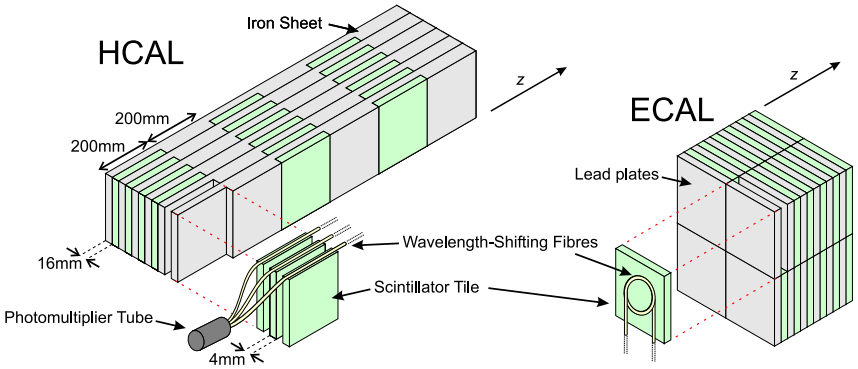


Figure 2.19.: The internal structure of the LHCb ECAL and HCAL, showing the iron/scintillator HCAL with edge read-out (left) and the lead/scintillator ECAL (right). The HCAL uses tiles whose narrow edges face the z direction while the ECAL tiles are aligned with faces orthogonal to z.

difference in interaction length. The PS and SPD employ an identical readout structure in which wavelength-shifting (WLS) fibers are embedded in the scintillator and coupled to multianode photomultiplier tubes (MAPMTs) via optical fibers. The WLS fibers are looped and both ends are connected such that they share a single PMT channel.

Electromagnetic Calorimeter

The LHCb Electromagnetic Calorimeter (ECAL) is a sampling calorimeter of alternating 4 mm thick scintillator and 2 mm lead tiles through which wavelength-shifting (WLS) fibers are threaded leading to a "shashlik" structure as shown in Figure 2.19. The wavelength-shifting fibers are read-out by photomultiplier tubes. In order to maximise energy resolution the ECAL is designed to fully contain showers from high energy pho-

tons, having a thickness of 25 radiation lengths, (X_0). The design energy resolution is $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$.

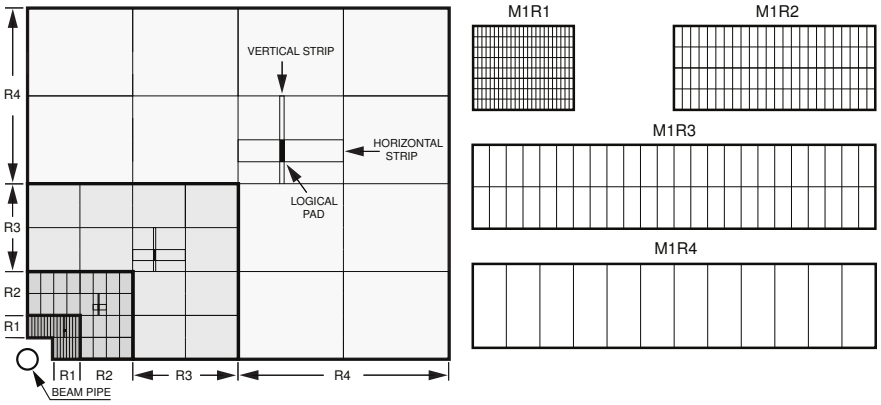
Hadron Calorimeter

The Hadron Calorimeter (HCAL) is a sampling calorimeter, using a scintillator/iron structure of which 18% is active material. The HCAL design is markedly different to the of the ECAL in that the module layers are "edge-on" in order to improve the sampling of the less collimated hadronic showers, as shown in Figure 2.19. The HCAL is also less segmented, consisting of two sensitive regions in which the inner cell density is twice that of the outer as shown in Figure 2.18b. Readout from the HCAL is the same as that of the ECAL, a WLS-PMT combination, but in the case of the HCAL the WLS fibers are mounted along the edge of the scintillator rather than through it. The trigger does not require hadronic containment and so the HCAL thickness is optimised for space constraints rather than shower containment. As a result the HCAL thickness corresponds to 5.6 hadronic interaction lengths, (λ) with energy resolution $\sigma_E/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$.

2.4.6. Muon System

Muons are an important final state for many B and B_s^0 meson decays, including the flagship LHCb channels $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \mu\mu$. They are also important for flavour tagging using semileptonic B-decays. In combination with the calorimeter, standalone high p_T muon tracks make up the first stage of the trigger. The muon system consists of five stations labelled M1-M5 in Figure 2.5, the first of which is upstream of the calorimeters to provide an enhanced p_T measurement. The remaining stations are situated behind the calorimeters and interspersed with 80 cm iron absorbers, corresponding to 20 λ total from M1 to M5. This leads to a minimum muon momentum of 6 GeV/c required to traverse all five stations.

As with the calorimeter, the muon system is segmented in order to keep channel occupancy constant across the full LHCb acceptance. Each muon station consists of four regions of increasing pad density towards the beampipe as shown in Figure 2.20a. Within a muon station each region is composed of muon chambers divided into pads, the logical pad structure for station M1 is shown in Figure 2.20b. The average pad density per station also varies; stations M2 and M3 double the number of pads in x



(a) Station M2 quadrant **(b)** Station M1 chambers

Figure 2.20.: Muon system channel density. (a) One quadrant of station M2 indicating the sensitive regions R1-R4 in which the channel density successively halves outwards from the beampipe. Each rectangle represents one chamber. (b) The pad structure of four chambers from station M1. Stations M2 and M3 have double the x-axis pad density per chamber while M4 and M5 have halve it.

of station M1 while stations M4 and M5 halve it. Stations M1-M3 require this higher spatial resolution to provide a p_{T} resolution of 20% for muon candidates, while stations M4 and M5 are only required for the identification of more penetrating particles. The muon system uses Multi-Wire Proportional Chambers (MWPCs) for all five stations with four gas gaps in stations M2-M5 and two gaps in M1 in order to reduce the material in front of the calorimeters. The M1 station also uses Gas-electron Multiplier (GEM) detectors consisting of two layers of triple-GEM detectors per chamber for the innermost (R1) region where a high flux requires radiation-damage resistant technologies. Both the MWPC and GEM chambers use an $Ar/CO_2/CF_4$ gas mixture resulting in 95% efficiency in a 20 ns window for the MWPCs and 96% for the GEMs. The LHCb muon system provides excellent resolution as illustrated by the dimuon mass spectrum in Figure 2.21.

2.4.7. Level 0 Trigger and readout

The LHCb trigger is composed of two stages: A hardware Level 0 (L0) trigger synchronous with the LHC bunch crossing frequency of 40 MHz designed to reduce event rates to 1 MHz and a flexible software Higher Level Trigger (HLT) which uses the full detector information to further reduce the event rate to 3 kHz suitable for offline stor-

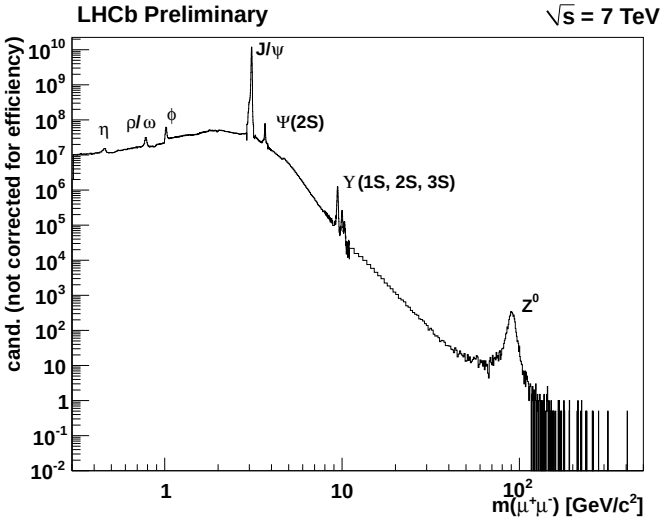


Figure 2.21.: The LHCb dimuon invariant mass spectrum with 2010 data, highlighting the excellent mass resolution over the full range. The Υ system is fully resolved.

age. The HLT will be discussed later in an analysis-specific context. The Level 0 trigger employs three custom processors, one each for the L0 Calorimeter triggers and the L0 Muon trigger. These provide information to the L0 decision unit (L0DU) where the decision to keep or discard each event is made.

L0 Calorimeter trigger

The L0 Calorimeter trigger is designed to accept high transverse-energy (E_{T}) particles. Front-end electronics boards compute the E_{T} sum of 2×2 cell clusters over an 8×4 cell region each in the ECAL and HCAL, selecting the highest E_{T} sum. The front-end boards pass their highest E_{T} clusters to a validation card which merges similar cluster information from the PS and SPD to ECAL clusters in order to determine electromagnetic particle ID. The validation card also merges ECAL energy with identified HCAL clusters to compute hadron E_{T}. The SPD multiplicity is computed in order to veto events with too many interactions which take a disproportionate time to reconstruct in the HLT. Finally a selection crate determines the highest E_{T} for each particle type: hadron, π^0, γ and e^-. Due to the timing requirements of the L0 trigger this procedure is carried out in dedicated hardware both on and off-detector. The on-detector front-end electronics are radiation-hard and require a three-way decision in order to avoid single-event upsets. Off-detector validation is performed on FPGAs communicating with the

front-end hardware by way of 20 LVDS links running at 280 MHz, the result of which is passed to the selection crates via 1.6 Gbps optical links.

L0 Muon trigger

The L0 Muon trigger finds the highest and second-highest p_{T} muon tracks. Station M3 seeds hits for the track-finding algorithm. Each hit in M3 is extrapolated to corresponding positions in stations M2, M4 and M5 based upon a straight-line from the nominal interaction point. Hits are then looked for in a field of interest (FOI) centered around this position in each station. The size of the field of interest is adjustable and depends upon preset minimum bias retention and background rates as well as the station under consideration. For M1 and M2 the field of interest is specified in x only, while stations M4 and M5 are open in both planes. The track p_{T} is determined from the hit position in M1 where the FOI is identified by a straight-line extrapolation from hits in M3 and M2.

L0 Pileup

The pileup detector consists of the first two sensor planes of the VELO situated upstream of the interaction point as indicated in Figure 2.8. These planes are sensitive in r with only coarse (45°) sensitivity in ϕ. Tracks from a Primary Vertex along the beam axis leave hits in the two planes at radii r_a and r_b as shown in Figure 2.22. For each combination of hits r_{a_i}, r_{b_j} in the same octant of both sensors the z-axis crossing $z_{i,j}$ can be determined by equation (2.5). Binning this distribution and performing a simple peak search allows the number of interactions per bunch crossing to be determined.

$$z_{v_{i,j}} = \frac{\left(\frac{r_{b_j}}{r_{a_i}}\right) z_a - z_b}{\left(\frac{r_{b_j}}{r_{a_i}}\right) - 1} \tag{2.5}$$

The resolution of z_v is limited by the hit resolution of the radial measurements and multiple scattering in in the silicon to around 3mm. After each iteration of the peak search the previous peak is masked allowing the next peak to be found.

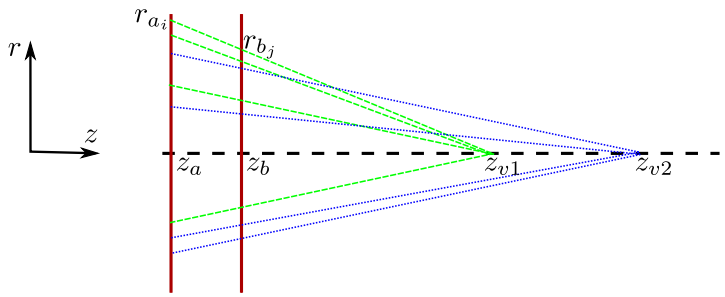


Figure 2.22.: Determination of pileup using the Pileup Detector. The first two r sensors of the VELO measure the z-axis crossing of backwards-pointing tracks according to equation (2.5) for all combinations of r_{a_i}, r_{b_j} in a common octant. Binning these z-axis crossings allows a simple peak search to determine the number of multiple interactions per bunch crossing.

L0 Decision Unit

The L0 Decision Unit (L0DU) receives information from the Muon and Calorimeter triggers at 40 MHz and applies a configurable algorithm to determine the trigger decision. This decision is then passed to the readout supervisor which includes additional self-triggers used for calibration and luminosity determination. The readout supervisor is ultimately responsible for forwarding the L0 decision, and can veto depending upon the buffer level and subdetector availability in addition to prescale settings. The L0DU is configured prior to each run using a Trigger Configuration Key (TCK). The TCK identifies uniquely each configuration for bookkeeping and verification purposes. The trigger configuration for L0 is dependent upon the expected beam conditions to be delivered by LHC, and specifies the individual Calorimeter E_{T} thresholds for electrons, pions, photons, hadrons as well as the SPD multiplicity cut and muon track p_{T} for a given FOI.

2.5. Time Alignment of the LHCb RICH detectors

As with any large-scale distributed electronic system requiring synchronous operation, timing is crucial for the LHCb experiment. When data is read out from the subdetectors and their individual sensors it should be done in such a way as to maximise collection efficiency and synchronised to the 40 MHz bunch crossing frequency. Time alignment must be performed to ensure that this is the case, both for individual subdetectors and for the experiment as a whole. The LHCb RICH subdetectors are no exception: HPDs

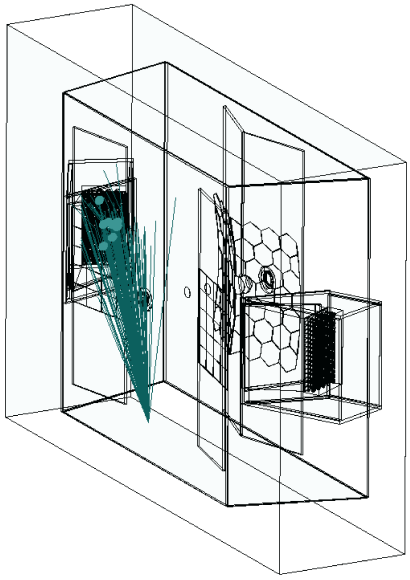
are subject to variation in photon collection efficiency as a function of time between illumination and readout. This time-dependent efficiency is strongly coupled to leakage current: The bias voltage across the silicon sensor of an HPD decreases as leakage current increases, which in turn increases the drift time of electrons through the silicon [66]. This section describes the procedure used and results of time alignment of the HPDs of RICH 1 and RICH 2.

2.5.1. RICH readout architecture

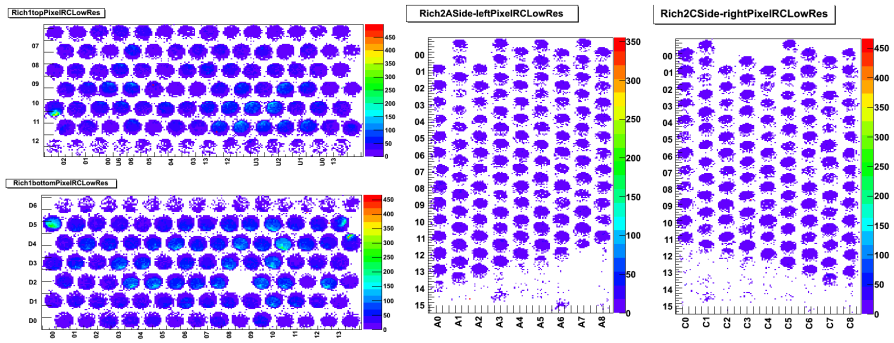
The RICH electronics system is designed to readout, control and supply power to the HPDs. The Level-0 (L0) electronics are on-detector and provide high voltage, silicon bias voltages readout of the silicon sensor. The Level-1 (L1) electronics are off-detector, housed in the cavern electronics barracks behind the radiation wall. The L1 implements multiplexing, data compression and outputs in a format compatible with the LHCb DAQ network architecture.

Timing control

The L0 electronics provides the timing control capability. A single L0 unit consists of two HPDs mounted on a single Level-0 board which provides Low Voltage (LV) and High Voltage (HV) distribution. In addition the L0 board acts as an interface between HPDs and the experiment control system (ECS), Timing and Fast Control (TFC) and is responsible for data transmission to the Level-1 electronics. The control, timing and readout interfacing is performed by the Pixel Interface (PINT), a dedicated, radiation-hard FPGA. The PINT chip receives data from two HPDs, adds event information and data integrity checking and transfers this data off-board via gigabit optical links. Input to the PINT consists of an additional optical link which supplies TFC data. This comes by way of an additional chip on the L0 board, the TTCrx, which generates the 40MHz clock, timing and calibration signals required to read-out the HPDs [67]. Upon receipt of a trigger signal, the TTCrx chip instructs the PINT to read-out both HPDs. The TTCrx has two on-board registers responsible for applying a delay prior to forwarding the trigger. These registers implement a coarse delay in 25 ns intervals equivalent to one bunch crossing over up to 15 steps, and a fine delay of stepsize 104 ps over a 25 ns range. Because each L0 board is shared by two HPDs and readout is governed by a single TTCrx/PINT per board it is not possible to alter timing characteristics of an individual



(a) RICH 2 fiber placement



(b) RICH hitmap

Figure 2.23.: (a) Visualisation of the RICH 2 detector. In green are simulated photons emitted from the nominal post-commissioning fiber position. The steep angle causes the lower HPDs to be in shadow. (b)RICH cumulative hitmap for all four detector planes taken during pulsed laser illumination. The downstream columns of the RICH 1 top and bottom enclosures (left) and the lower portions of the RICH 2 A and C sides (right) suffer from poor illumination due to fiber placement, positioned this way to ensure that they are outside the LHCb acceptance. The missing HPD in the RICH 1 bottom enclosure is due to a faulty VCSEL, and was soon replaced.

HPD. To reduce the relative timing offset between HPDs sharing a L0 board, HPDs have been paired according to leakage current as measured at the Photon Detector Test Facility in Edinburgh prior to installation in the RICH detectors.

Pulsed laser control and fiber distribution

Prior to completion of the LHC and LHCb, alignment of the RICH was performed using a pulsed laser system. This consists of a Hamamatsu picosecond light pulser (PLP10-065C) controller and laser diode installed in the barracks of the experimental cavern. This system is capable of producing pulses of 100ps FWHM at a wavelength of 655nm. The laser output is split and fed to the RICH HPD enclosures via fiber-optic cables. In RICH 1 four illuminating fibers are used, two for each HPD plane as space is limited and proximity of the beampipe prevents even illumination with only a single fiber. For RICH 2 simulation determined an optimal positioning that only requires one fiber per plane, but this is subject to the constraint that towards the end of the commissioning phase it must be moved out of the LHCb acceptance in preparation for beam to the position indicated in Figure 2.23a [68]. As a consequence of these constraints, RICH 1 receives slightly less illumination on columns furthest from the interaction region where a shadow is cast by the edge of the enclosure and RICH 2 HPDs towards the bottom of the enclosure are effectively in the shadow of the RICH 2 radiator volume wall as shown in Figure 2.23b. The laser is interfaced to the LHCb standard controls network by way of a custom electronics board nicknamed the Johan Laser Interface (JOLI) board after its designer. This board contains a TTCrx chip whose purpose is to receive trigger signals and timing delays for the laser. The TTCrx is in turn interfaced to a Serial Protocol for the Experimental Control System (SPECS) slave controller [69]. The SPECS interface is a general-purpose system for the slow-control of experimental hardware, designed to permit configuration of the experiment through a uniform software and hardware architecture.

2.5.2. Timing Scan Control Software and Strategy

The LHCb experiment is controlled via a large-scale software Finite-State Machine (FSM) implemented within the Supervisory Control And Data Acquisition (SCADA) framework common to the LHC experiments known as the Joint Controls Project (JCOP) which builds upon a commercially available product, PVSS [70]. PVSS is responsible for storing hardware configuration data, transmitting it to the experiment subsystems via SPECS, monitoring the on- and off-detector hardware, configuring the farms and transitioning through the FSM states necessary to run LHCb during data taking. The FSM is modular, permitting individual subdetectors to run in standalone configuration for testing and calibration purposes, with separate “Recipes” stored for different specific

tasks. Recipes have been written for timing scans of the RICH detector, common to which are the concept of automated stepping. The RICH subdetectors are configured in standalone mode, using either their own clock to provide triggers when the pulsed laser is deployed or a parasitic L0 trigger taken from the experiment L0 Decision Unit when taking collision data. The user configures the PVSS recipe to take a specified number of triggers per step over a specified number of steps. In the case of scans using the pulsed laser, the TTCrx on the JOLI board is configured with its own delay register values so that upon receipt of a trigger signal it is pulsed at a delay that puts it roughly in time with the global LHC trigger. When PVSS has been configured the run is started, and the L0 TTCrx registers are configured according to the first step in the recipe. The run then proceeds until the prerequisite number of triggers have been sent to the L0 boards for this step. At this point the run is paused and the L0 TTCrx registers are incremented by the recipe's step width. Depending upon the specific recipe used a step can correspond to a single 25 ns coarse delay increment, a single fine delay increment of 104 ps, or multiples/combinations thereof. The run continues for the same number of triggers, and the registers are incremented again. This process continues until the total number of required steps have been recorded. At this point the run stops and the data is stored either locally or offline in preparation for analysis. Different timing scan resolutions and strategies have their own PVSS recipes, described below:

- `CALIB|COARSETIMEALIGN` and `CALIB|FINETIMEALIGN`: In the early stages of commissioning when the timing was uncalibrated it was common to use 15 coarse time steps, the recipe for which was called `CALIB|COARSETIMEALIGN`. This is equivalent to the maximum time range over which a TTCrx, and therefore L0 board, can be stepped. Once all boards were aligned to within 75 ns (3 coarse steps) per RICH, `CALIB|FINETIMEALIGN` could then be used, which increments the fine delay TTCrx register by multiples of 10, equivalent to $1.04(\sim 1)$ ns, modulo the coarse delay in order to cover the full 75 ns range. Both of these recipes overwrite the presently stored TTCrx delay registers, so all L0 boards are set to the same timing at each step.

- `CALIB|TIMESCANATNOMINAL`: Once an initial low resolution timing scan has been performed and the updated TTCrx delay register values have been stored it is a more interesting exercise to investigate the timing in an envelope around the stored values. For this purpose additional recipes have been developed that step each L0 board around the stored "nominal" value. These have been written for 1 ns, 5 ns and 25 ns steppings with a range of ± 25 ns for the 1 ns scan and ± 50 ns for

the 5 and 25 ns scans. This permits a more rapid evaluation of the RICH timing and allows the operator to spot systematic drift from the previously applied timing settings.

2.5.3. Timing Scan Analysis Software

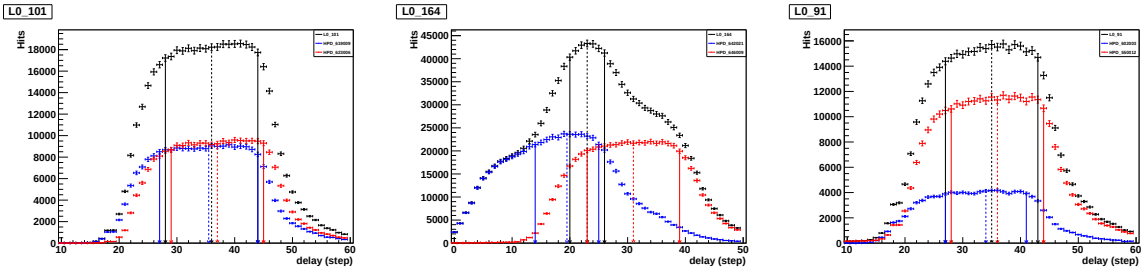
The timing scan analysis software consists of two separate tools: An algorithm that processes raw LHCb data into a more portable format and an offline analysis executable that produces statistics, diagnostic plots and the new TTCrx register values required to align the RICH detector, one per L0 board.

Data taken during a timing scan consists of a series of sequential raw events of size $n \times m$ where n is the number of triggers requested per step and m is the number of steps. The timing scan analysis algorithm `Rich/RichTimingScanAnalysis` processes these raw events within GAUDI, the LHCb analysis framework. The output of this algorithm is a histogram, one each per L0 board and HPD of m bins. Each bin contains the total number of recorded pixel hits over the n collected triggers, or optionally a clustering algorithm is employed in which case only clustered hits of a user-specifiable size range are saved. The histograms are written out to a standard ROOT Ntuple with a common naming convention to undergo further processing.

The offline analysis executable, packaged within `Rich/RichTimingScanAnalysis` and called `TimeScan.cc` locates two bins per HPD and L0 board from which to determine the timing: The *turn-on* and *turn-off* bins. These are defined as the first and last bin for which a specified fraction of the background-subtracted maximum number of hits is reached. Typically this fraction is taken to be 90%. The *midpoint* is the center of these two bins, and considered to be the optimal timing position of the L0 board. The contribution of each HPD to an L0 board's timing histogram is intentionally not normalised: The HPD with higher occupancy contributes proportionally more to the midpoint calculation than its lower-occupancy partner. In the case where two HPDs have slightly different timing characteristics this leads to a slight increase in efficiency for the HPD in the higher occupancy region. The analysis executable also calculates the individual midpoints for each HPD and checks that they deviate by no more than a conservative 5 steps from each other. If they do a warning is printed on-screen and the timing profile of this board and HPDs is written to disk for closer inspection. The midpoint, turn-on and turn-off profile for all boards in a given RICH are saved, and a Timing Alignment Plot is produced showing the midpoint, turn-on and turn-off for

each board together. This is a useful diagnostic as the boards are processed and plotted sequentially according to the mounting pattern of the RICH detectors. Neighbouring boards on the plot are physical neighbours in the RICH, ordered by position on each column, column position on each enclosure and enclosure position per RICH from left to right. The final output from `TimeScan.cc` is a machine-readable text file containing the TTCrx delay differences required to align the L0 boards of each RICH based on their pre-alignment value. These files can be read directly in as PVSS datapoints, allowing a rapid deployment of new timings.

A typical timing histogram for an L0 board as produced by `TimeScan.cc` is shown in Figure 2.24a. Here the HPDs are correctly paired with only a few ns deviation in timing characteristics. The majority of L0 boards in RICH 1 and RICH 2 exhibit this behaviour. The overall shape of the timing profile is that of a slightly asymmetric top-hat distribution with a longer trailing edge caused by *charge sharing* and *backscatter*. Charge sharing between adjacent pixels on the silicon sensor dilutes the number of electron-hole pairs per pixel such that it takes longer to overcome the predefined threshold that determines a pixel "hit". In scans where clusters of size > 2 pixels were rejected the trailing edge was observed to diminish consistent with this effect, but resulted in no change to the position of the turn-off bin with respect to scans without clustering. Backscattered photoelectrons have an increased path length leading to a similar effect. The plateau width (the width of the $> 90\%$ efficiency regime) per-HPD is approximately 16 ns which gives a large safety margin to account for timing drift and minor HPD misalignments. Figure 2.24b shows an extreme case: In this particular instance one of the HPDs had been replaced due to vacuum degradation at a time when HPDs with similar leakage current characteristics were not available. For this particular pair the occupancies of both HPDs are similar, but their individual midpoints are separated by 12 ns, at the edge of acceptable timing compatibility. Both HPDs contribute with approximately similar weight to the L0 midpoint determination and at their present position are both $\geq 90\%$ efficient but any deviation from the midpoint due to drift or variation in leakage current over the lifetime of the HPDs will rapidly reduce efficiency. Figure 2.24c shows the case of two HPDs with different occupancies. This can come from the intrinsic quantum efficiency differences inherent in the manufacture of the HPD or from transition between areas of different physical occupancy in the detector (although a case as extreme as this is more likely to be due to the former). In such cases the higher occupancy HPD contributes more to the L0 timing profile and the alignment will be biased to favour this HPD.



(a) Typical board response **(b)** Different timings **(c)** Different occupancies

Figure 2.24.: Time alignment profiles produced by `TimeScan.cc`. In each profile black datapoints indicate the total hits for this L0 board. Datapoints in blue indicate the first HPD on the board and red indicates the second. Solid arrows indicate the turn-on and turn-off bins while dashed arrows indicate the midpoint between these.

2.5.4. Time alignment results prior to LHC startup

Towards the end of the commissioning stage time alignment was performed using the pulsed laser. The procedure consisted of scans with increasingly finer resolution, starting with a coarse time scan of 25 ns step width and finishing with full alignment to 1 ns resolution. The laser alignment scans cannot align RICH 1 and RICH 2 together, or even the detector planes of a given RICH: Fiber length differences in the distribution of the pulsed laser to RICH 1 top and bottom planes is $\mathcal{O}(1\text{m})$, leading to an appreciably different average midpoint profile for the two planes. There is an even greater difference between RICH 1 and RICH 2. The fiber lengths between the RICH 2 A and C sides are similar however. When time aligning with laser care was taken to maintain the average midpoint for each of RICH 1 top, RICH 1 bottom and RICH 2 to prevent introducing false delays.

The 25 ns scan determined that the average timing delay due to fiber between RICH 1 and RICH 2 was 3-4 coarse delay steps, equivalent to $\sim$75 ns. All HPDs had turn-on and turn-off in the same 25 ns bin per RICH, allowing finer scans to proceed without realignment. In order to perform scans at a finer resolution with the laser it was necessary to take two separate scans, one for each RICH, in which the laser TTCrx was delayed by these additional steps in the case of RICH 2 so that the full range of RICH 2 timing profiles could be covered. The final laser alignment procedure as performed in November 2009 just prior to the first collisions is shown in Figure 2.25 and Figure 2.26. Figure 2.25a and Figure 2.26a show the pre-alignment turn-on, turn-off and midpoints of all active L0 boards in RICH 1 and RICH 2 respectively. RICH 1 can be seen to have a higher

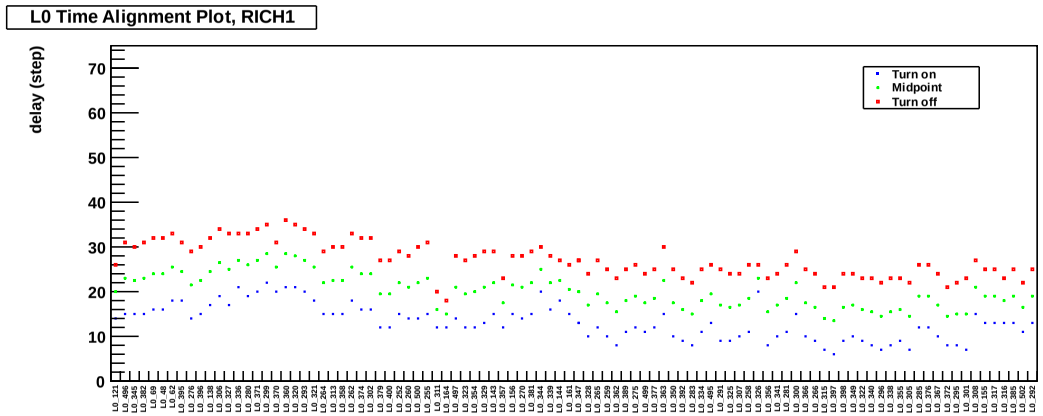
average midpoint value for the first half of the plot (the down box) than for the second (the upper box). This is due to the increased fiber length to this plane's fiber positions which are mounted in the opposite plane. For RICH 2 there are L0 boards that appear to turn-on earlier than their neighbours. This is due to the poor illumination of their HPDs as shown in Figure 2.23b. The timing profiles of these boards are effectively a measure of their dark count which causes the midpoint determination to choose the first upwards fluctuation.

Because a global timing cannot be derived from the laser scans the timing procedure in this instance was to preserve the average delay of each plane while minimising the spread. L0 boards were therefore delayed by their difference from the mean timing value. For RICH 1 this mean was 18 ns for the upper plane and 23 ns for the lower plane. The results of the laser alignment procedure for RICH 1 are visible in comparison between Figure 2.25b and Figure 2.25d where the RMS per plane drops from ~ 3 ns to < 1 ns while the mean midpoint is preserved for each plane as shown in Figure 2.25c.

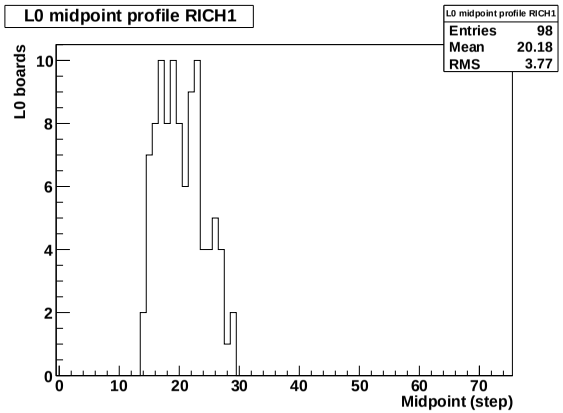
For RICH 2 the mean midpoint prior to alignment was determined to be 35 ns (+50 ns relative to RICH 1) after excluding shadowed L0 boards which manifest as the leading tail in Figure 2.26b. The mean midpoint for the A and C sides of RICH 2 prior to alignment differed by less than 1 ns, making a per-plane alignment unnecessary. Post-alignment with laser the timing resolution for RICH 2 ~ 2.1 ns, centred on the same mean midpoint as shown in Figure 2.26d.

2.5.5. Time alignment with collision data

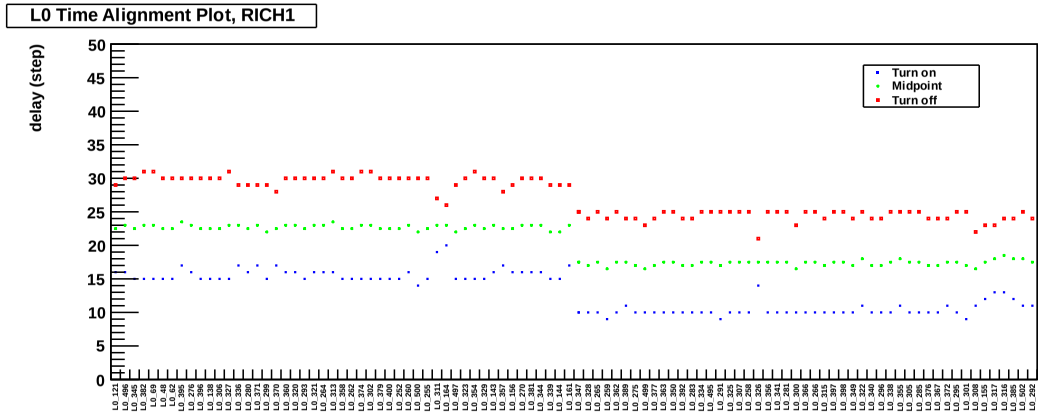
During the start-up period time alignment with beam and beam-gas events was performed in parallel to the alignment of the other LHCb subdetectors. Alignment to beam is simpler than that of laser as the timing that maximises collection efficiency for each HPD with beam is optimal without the need to account for false delays. There is no need to account for unphysical delays as was the case with the Laser timing. Initially, alignment scans were performed using the `CALIB|TIMESCANATNOMINAL` recipe on 450 GeV beam-gas events as triggered by the CALO just prior to "first collisions" in late November of 2009, and again confirmed with 450 GeV beam-beam collision data. Due to the limited rate and rapidly varying beam conditions associated with the start-up of the LHC 5 ns resolution was used, and the results showed that the laser alignment translated to an average midpoint per-rich of 0 ± 5 ns when tested with beam. The width of



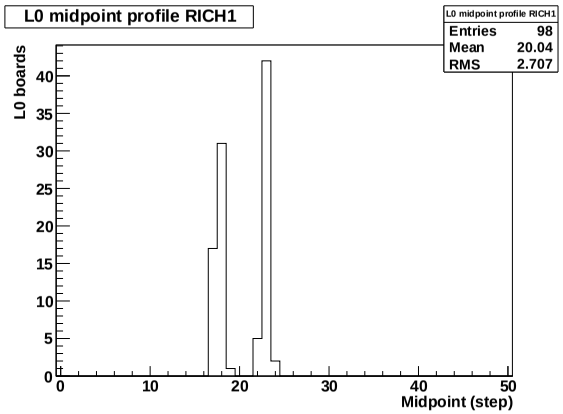
(a) Timing plot before alignment



(b) Midpoint profile before alignment

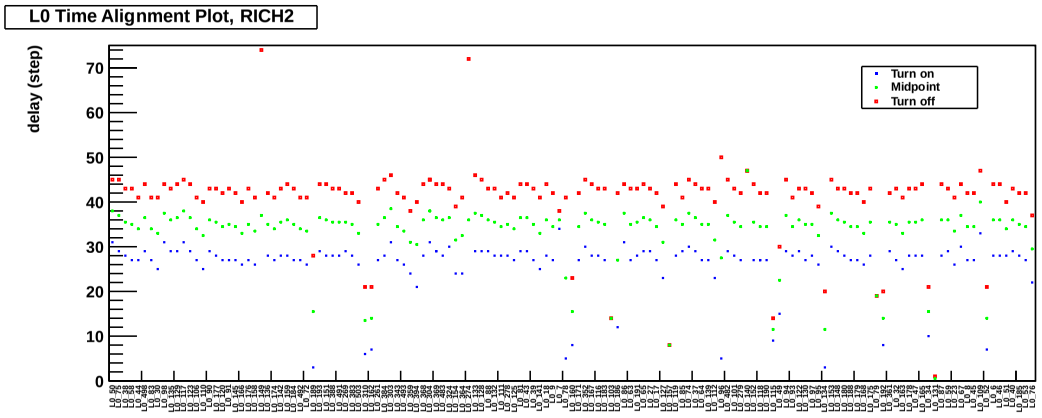


(c) Timing plot post-alignment

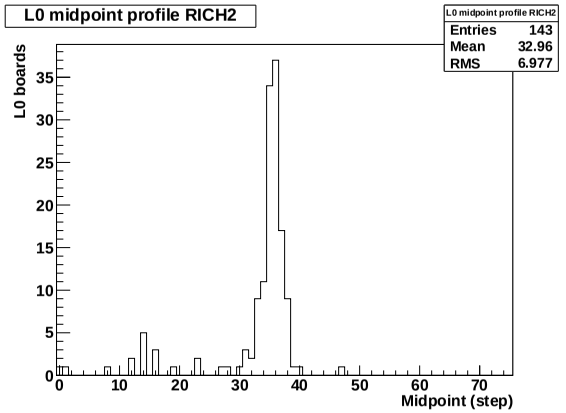


(d) Midpoint profile post-alignment

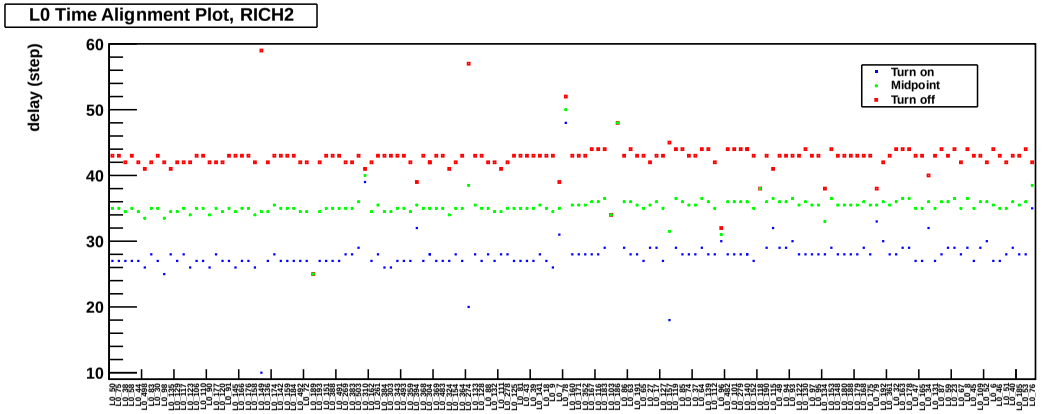
Figure 2.25.: 1 ns time alignment scan for RICH 1 taken using the pulsed laser. Each datapoint is determined from 2000 recorded triggers per step. There is a clear gradient from left-to-right in (a) due to a combination of path-length difference for photons emitted from the fiber and a difference in fiber lengths between the upper and lower HPD planes. This delay is deliberately preserved during alignment as the fiber delay will not be present in collisions as is shown in (c) and (d) where two distinct peaks are observed, separated by the mean timing difference between the top and bottom planes.



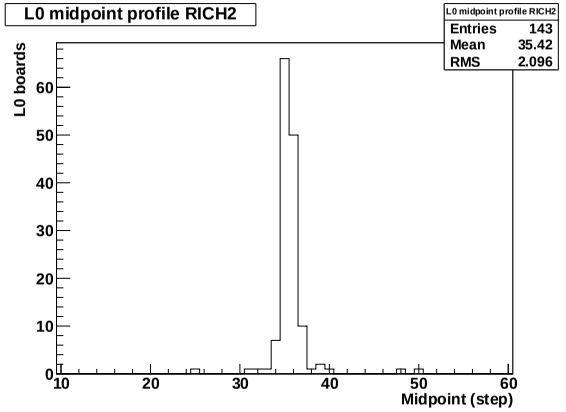
(a) Timing plot before alignment



(b) Midpoint profile before alignment



(c) Timing plot post-alignment



(d) Midpoint profile post-alignment

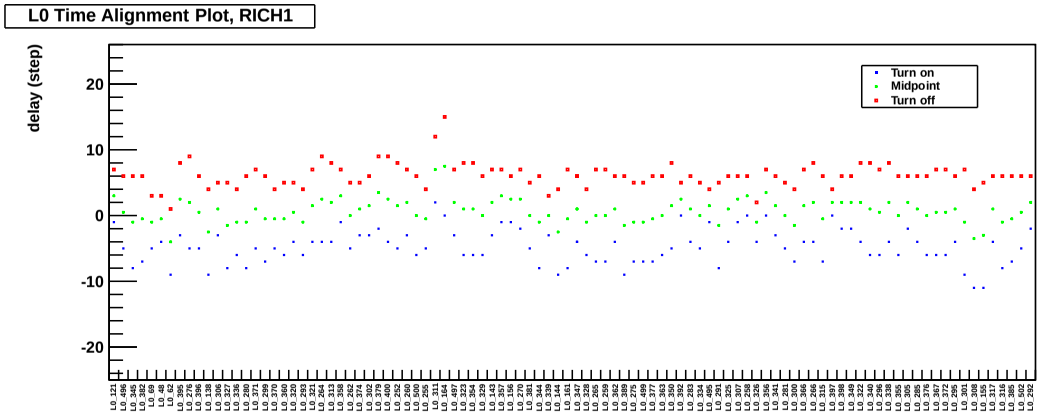
Figure 2.26.: 1 ns time alignment scan for RICH 2 taken using the pulsed laser. Each datapoint is determined from 2000 recorded triggers per step. A number of HPDs at the bottom of the A and C side enclosures are poorly illuminated, leading to badly defined midpoint estimates visible in (a) and (b). These are accounted for by moving their timings to correspond to the average timing of all other L0 boards in RICH 2.

the timing plateau for L0 boards being ~ 16 ns meant that 5 ns was considered suitable for physics.

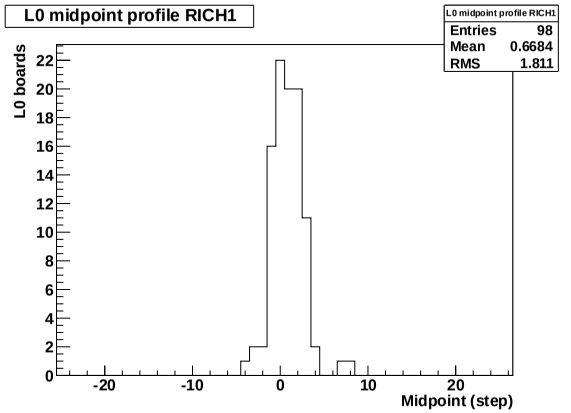
After the winter shutdown In April of 2010 the LHC went to 3.5 TeV collisions and increased the collision rate, allowing for a finer resolution alignment. `CALIB|TIMESCANATNOMINAL` was again used, collecting 5000 triggers per timing step over 51 1 ns steps. The results of the initial scan are available in Figure 2.27b and Figure 2.28b for RICH 1 and RICH 2 respectively. There is evidence of a periodic structure in the initial alignment plot for RICH 2 as shown in Figure 2.28a. This is due to the introduction of an additional false delay during the laser alignment procedure: There is a path-length difference for photons emitted by the laser incident upon HPDs in the top of the A and C side enclosures with respect to those in the bottom. This difference is of $\mathcal{O}(1\text{ m})$, leading to a ~ 3 ns delay due to the fiber placement and RICH 2 geometry. In collision data this delay is not present, so the laser effectively aligns the top and bottom of each RICH 2 enclosure to no better than ~ 3 ns resolution when this effect is not accounted for. After applying the updated TTCrx delays this periodic effect is no longer present in RICH 2, and the RICH subdetectors are aligned to 1.1 ns and 0.8 ns about the optimal timing respectively as evident from Figure 2.27d and Figure 2.28d.

2.5.6. Conclusions

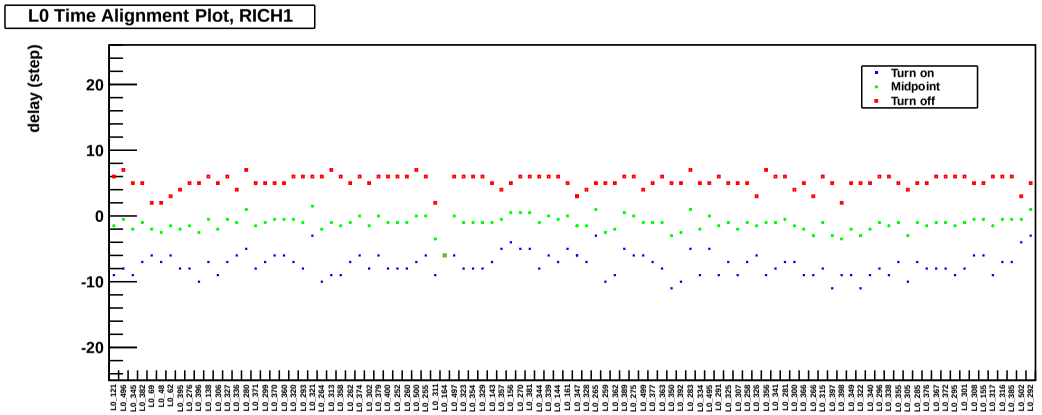
The LHCb RICH subdetectors require time alignment in order to maximise readout efficiency. This has been performed using a combination of a pulsed laser and pp collision data prior to physics data taking using self-contained analysis software and tools to manipulate the TTCrx delay registers of individual L0 boards. The final alignment of the RICH detectors at the start of the 2010 data taking period resulted in RICH 1 being aligned to -1.1 ± 1.1 ns and RICH 2 aligned to -0.2 ± 0.8 ns of the optimal readout delay with respect to the LHC bunch crossing.



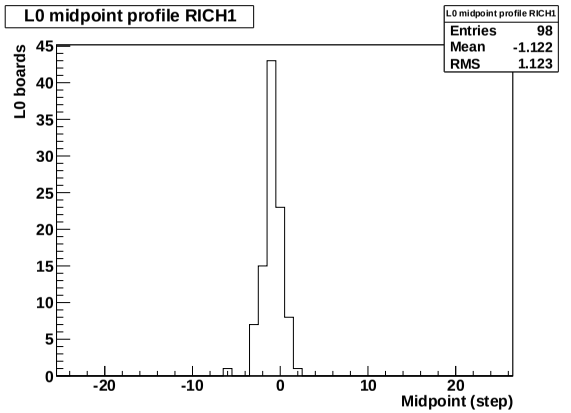
(a) Timing plot before alignment



(b) Midpoint profile before alignment

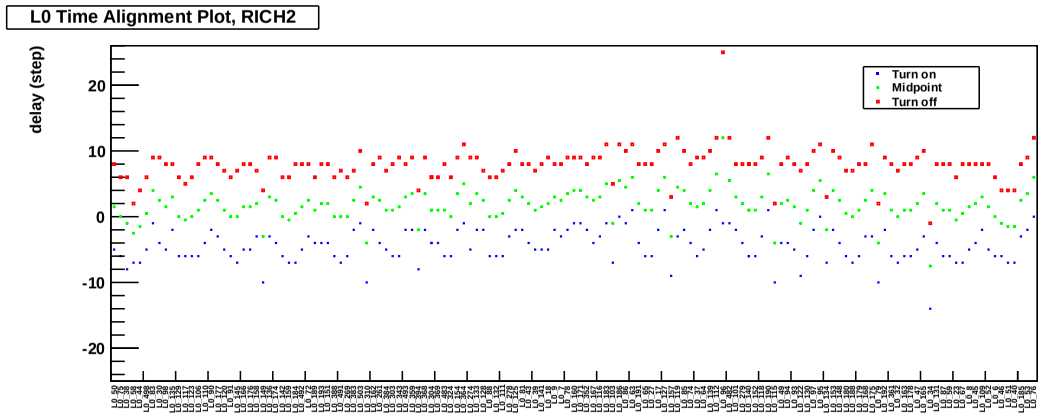


(c) Timing plot post-alignment

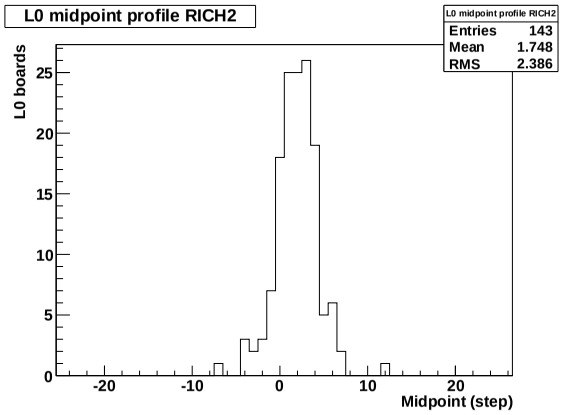


(d) Midpoint profile post-alignment

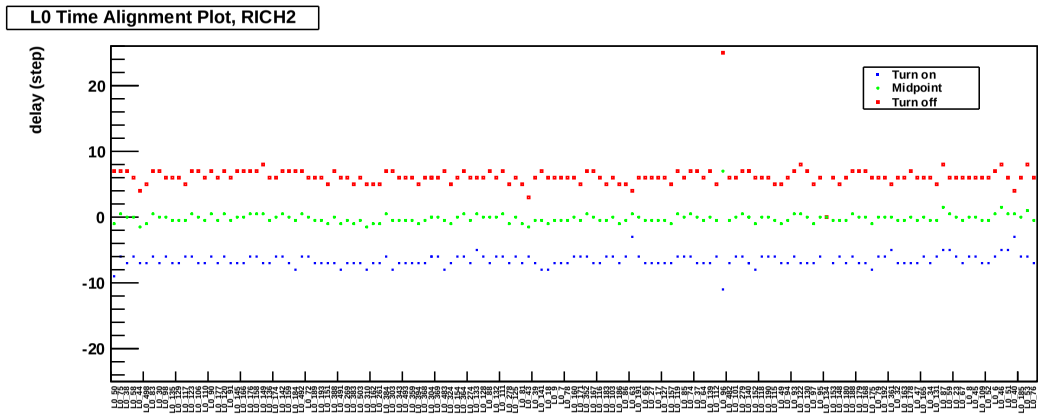
Figure 2.27.: 1 ns time alignment scan for RICH 1 taken using 3.5 TeV collision data. Each datapoint is determined from 5000 recorded triggers per step. The alignment procedure reduces the RMS spread of the timing to within -1.1 ± 1.1 ns of the optimial timing.



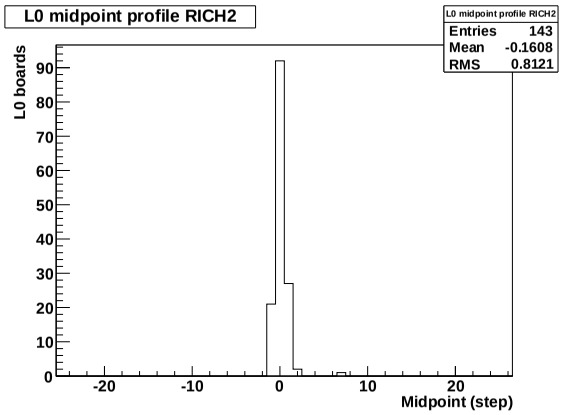
(a) Timing plot before alignment



(b) Midpoint profile before alignment



(c) Timing plot post-alignment



(d) Midpoint profile post-alignment

Figure 2.28.: 1 ns time alignment scan for RICH 2 taken using 3.5 TeV collision data. The initial timing scan in (a) contains evidence of fine structure. This is due to neglecting the path length difference of photons emitted from the fiber during alignment with laser, where a $\sim$3 ns difference between the top and bottom of the RICH 2 enclosures is present. Alignment with beam removes this artificial delay, showing no further structure in (c). The alignment procedure reduces the spread of the timing to within -0.2 ± 0.8 ns of the optimal timing.

Chapter 3.

Inclusive $\mathrm{D}^{\pm}$, $\mathrm{D}_{\mathrm{s}}^{\pm}$ production cross-sections at $\sqrt{s} = 7$ TeV

"Facito aliquid operis, ut te semper diabolus inveniat occupatum."
— St. Jerome

The start-up phase of any experiment is an exciting time: With the initial data samples collected it is desirable to measure well-understood processes to calibrate the detector and ensure that it is performing to expectation. It is also an opportunity to make measurements on the outskirts of the defined physics scope. The process $\phi \to \mathrm{K}^+\mathrm{K}^-$ is useful for calibration of the kaon particle identification (PID) performance, but studies prior to data taking suggested that selecting and reconstructing this mode inclusively would not result in sufficient purity for this purpose. A higher purity could be obtained by selecting and reconstructing $\mathrm{D}_{\mathrm{s}}^{\pm} \to \phi\pi^{\pm}$ in which the ϕ decays to two kaons, and at the time of data-taking I had such a selection prepared. It was with this selection that the first $\mathrm{D}_{\mathrm{s}}^{\pm}$ candidates were reconstructed at LHCb. This chapter presents a measurement of the inclusive $\mathrm{D}_{\mathrm{s}}^{\pm}$ production cross-section and the $\mathrm{D}_{\mathrm{s}}^{\pm}/\mathrm{D}^{\pm}$ cross-section ratio as determined from the decays $\mathrm{D}^{\pm}, \mathrm{D}_{\mathrm{s}}^{\pm} \to \phi\{\mathrm{K}^+\mathrm{K}^-\}\pi^{\pm}$. A general-purpose tool that recursively optimises cuts by maximising a user-specifiable figure-of-merit using either Monte-Carlo or a fully data-driven approach is introduced. This tool, `CROP`, allowed a rapid selection to be developed prior to tuning of the LHCb simulation.

Magnet Polarity	LHCb Run numbers	LHC Fill numbers	Total triggers
Up	71474, 71476-79, 71481,	1089	
	71485-94, 71525-27, 71530	1090	82,241,048
Down	71807–13, 71815, 71816,	1101	38,524,560

Table 3.1.: The first LHCb datasets collected at 7 TeV, as used by the charm production cross-section measurements.

Mode	Total Events	
	Magnet Down	Magnet Up
Minimum Bias	10,699,961	10,650,460
$D_s^{\pm} \to K^+K^-\pi^{\pm}$	759,996	778,989
$D^{\pm} \to K^+K^-\pi^{\pm}$	782,992	742,495

Table 3.2.: Monte-Carlo datasets from the Beam3500GeV-VeloClosed-Mag(Down/Up)-Nu1, 2010-Sim03Reco03-withTruth configuration and processing, as used by the charm production cross-section measurements.

3.1. Datasets and selection

Due to the low intensities of the early LHC operating regime a "microbias" trigger could be used, which accepts all inelastic pp collisions without prescale. These triggers are 100% efficient for the modes $D_s^{\pm}, D^{\pm} \to \phi\pi^{\pm}$. A total integrated luminosity of 1.81 nb^{-1} has been collected in this regime, of which approximately two thirds were collected with an upwards-pointing magnetic field orientation as shown in Table 3.1.

3.1.1. Preselection

Prior to finalisation of the full dataset, an initial preselection for $D_s^{\pm} \to \phi\{K^+K^-\}\pi^{\pm}$ was determined by-eye to be used as a starting point for further studies. This was developed initially using approximately 50% of the magnet up sample, a total of 43M triggers, in combination with minimum-bias Monte-Carlo. At this stage only very basic criteria are imposed upon the reconstructed decay. This initial preselection is performed

in the DaVinci online analysis framework as listed in the first column of Table 3.3 and is described as follows:

- All tracks entering the combination should have a χ^2 in the track fitting algorithm of less than 10 per degree of freedom.
- Candidates combined to form the ϕ should have a difference between the K and π RICH hypothesis of $\Delta_{LL}(\mathrm{K} - \pi) > 10$.
- The $\mathrm{K}^+\mathrm{K}^-$ candidate pair should have an invariant mass within 75 MeV of the nominal ϕ mass.
- The invariant mass of the $\mathrm{K}^+\mathrm{K}^-\pi^{\pm}$ combination should be within 200 MeV of the nominal $\mathrm{D_s}$ mass.
- The cosine of the angle between a reconstructed candidate's momentum vector and the direction defined by the candidate vertex with respect to the Primary Vertex (PV), known as the DIRA, should be > 0.995.

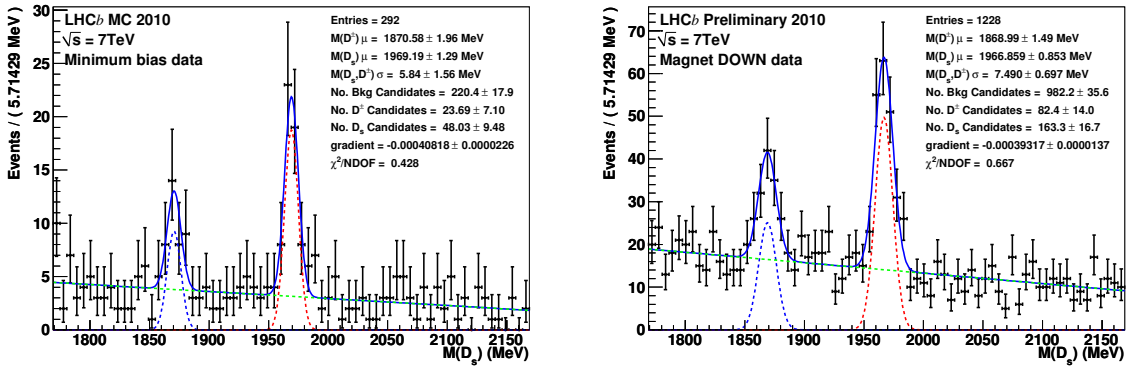
These candidates are written out from the analysis framework in `ROOT::TNtuple`[1] format where a selection is performed by-eye in order to produce peaks in the $D_s^{\pm}$, $D^{\pm}$ mass regions. The additional criteria applied for this initial selection as listed in the second column of Table 3.3 are as follows:

- All daughter tracks are subject to the loose requirement that they do not come from the Primary Vertex (PV), by having an Impact Parameter (IP) greater than twice its own uncertainty: PV IP $\chi^2 > 2.0$.
- The DIRA is further tightened to > 0.9999.
- The candidate $D^{\pm}$ or $D_s^{\pm}$ should have a flight distance between the primary vertex and the candidate vertex greater than 35 times its uncertainty: FD $\chi^2 > 35.0$.
- The χ^2 of the vertex fit to the candidate $D^{\pm}$ or $D_s^{\pm}$ should be < 5.0.
- The largest distance of closest approach (DOCA) between the $\mathrm{K}^+\mathrm{K}^-\pi^{\pm}$ candidates should be less than 0.2 mm.
- The p_{T} of the candidate $D^{\pm}$ or $D_s^{\pm}$ should be greater than 1 GeV.

[1] An Ntuple is a data storage format in which each entry or candidate has the same number of stored elements, for example momentum 4-vector, impact parameter, vertex position, daughter particle ID hypotheses, etc.

cut	initial preselection	initial selection
K $\Delta_{LL}(K-\pi)$	> 10.0	-
K,π Track χ^2/DOF	< 10.0	-
K,π PV IP χ^2	-	> 2.0
$D_s^{\pm}$,$D^{\pm}$ DIRA	> 0.995	> 0.9999
$D_s^{\pm}$,$D^{\pm}$ Flight. Dist. χ^2	-	> 35.0
$D_s^{\pm}$,$D^{\pm}$ Vertex χ^2	-	< 5.0
$D_s^{\pm}$,$D^{\pm}$ DOCA	-	< 0.2 mm
$D_s^{\pm}$,$D^{\pm}$ p_T	-	> 1.0 GeV
ϕ Mass window	± 75 MeV	± 20 MeV
$D_s^{\pm}$,$D^{\pm}$ Mass window	± 200 MeV	-

Table 3.3.: Initial $D_s^{\pm}$,$D^{\pm}$ selection criteria determined by-eye from MC2010 minimum-bias Monte-Carlo



(a) Monte-Carlo **(b)** Data

Figure 3.1.: Invariant $K^+K^-\pi^{\pm}$ mass distributions of candidates passing the initial selection in Table 3.3 applied to 10.65M minimum bias magnet up Monte-Carlo events and the first 43M events collected by the LHCb experiment in 2010. The signal is modelled by single Gaussian functions with a shared mass resolution for the $D_s^{\pm}$ and $D^{\pm}$ peaks. Background is modelled by a linear function.

- The K^+K^- combination invariant mass is further reduced to ± 20 MeV.

With this selection the first $D_s^{\pm}$,$D^{\pm}$ candidates were observed in LHCb. The $K^+K^-\pi^{\pm}$ invariant mass distribution is shown in Figure 3.1b, where a fit to two Gaussians over a

Candidate	Fit	Truth-matched
Background	220 ± 18	222
$D^{\pm}$	24 ± 7	18
$D_s^{\pm}$	48 ± 9	52

Table 3.4.: Yields returned by the fit shown in Figure 3.1a directly compared to the truth-matched yields for all MC candidates with full MC association to the $\phi\pi^{\pm}$ final state.

cut	Preselection
D_s, D^+ mass	$1770 < M < 2070$ MeV
ϕ mass	$1000 < M < 1040$ MeV
D_s, D^+ DIRA	> 0.9999
D_s, D^+ Flight. Dist. χ^2	> 60
π IP PV χ^2	> 10
K IP PV χ^2	> 2
D_s, D^+ Vert χ^2/DoF	< 15
$K\Delta_{LL}(K - \pi)$	> 5
K, π Track χ^2/DoF	< 10.0

Table 3.5.: $D_s^{\pm}$,$D^{\pm}$ Preselection criteria designed to use no p_T cuts on either the parent or daughter candidates, with a retention of $\sim 10 \times 10^{-4}$ on microbias-triggered data for the first 1.81 nb^{-1} of data recorded by LHCb at $\sqrt{s} = 7$ TeV.

linear background is overlaid, in which the $D_s^{\pm}$ and $D^{\pm}$ share a common mass resolution but all other parameters are permitted to float. The $D_s^{\pm}$ and $D^{\pm}$ peaks are clearly visible. As a cross-check to ensure that the selection was correctly selecting $D_s^+, D^+ \to \phi\pi^+$ and their charge conjugates the same selection was applied to 10.65 million magnet up minimum-bias Monte-Carlo candidates from which it is possible to determine the generated decay information. This dataset is listed in Table 3.2. The Monte-Carlo mass distribution is shown in Figure 3.1a using the same fit as to data. The fit to Monte-Carlo is consistent with the true $D_s^{\pm}$,$D^{\pm}$ yields, as indicated in Table 3.4, and the Monte-Carlo distributions agree with those found in the data up to differences in mean and resolution due to the preliminary status of the alignment and mass calibration at the start of 7 TeV running. This initial selection formed the basis for a more robust preselection which could be applied online to the entire 1.81 nb^{-1} and leave room for offline cut optimisation. The p_{T} cut was removed in order to investigate the possibility of a cross-section measurement at low transverse momenta. The DOCA cut was also removed, and the Vertex χ^2 requirement was loosened. Due to concerns over the differences in particle ID between data and Monte-Carlo the kaon PID cut was also loosened to $\Delta_{LL}(\text{K} - \pi) > 5.0$. In order to reduce the retention rate in the absence of these cuts, the parent flight distance cut was increased to $FD\chi^2 > 60$ and the pion impact parameter requirement was increased to $PV\ IP\chi^2 > 10$. The K^+ K^- invariant mass was tightened to ± 20 MeV in order to reduce the nonresonant background. The preselection criteria are listed in Table 3.5. Applying this preselection to the full 1.81 nb^{-1} of microbias-triggered data listed in Table 3.1, consisting of 121M triggers total, 10316 candidates remain. The $\text{K}^+\text{K}^-\pi^{\pm}$ invariant mass spectrum is shown in Figure 3.2, where the same fit as before is used but there are now sufficient statistics to allow the mass resolutions to be fit to separately.

3.2. MC-Free selection using CROP, the Cut Recursive OPtimiser

An offline selection optimisation strategy was applied in order to improve signal purity prior to extraction of yields for the cross-section determination. Optimisation is typically performed by maximising a figure-of-merit (FoM) using Monte-Carlo samples where signal and background distributions are known. At the time of this analysis the alignment and calibration of the experiment was not yet mature, resulting in discrepancies between

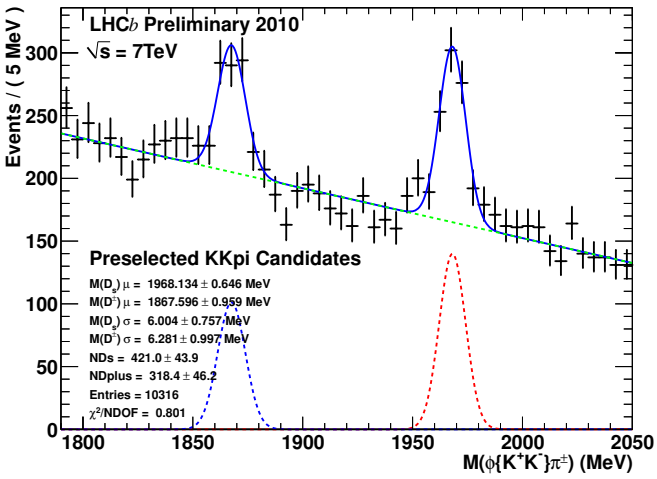


Figure 3.2.: $K^+K^-\pi^{\pm}$ invariant mass spectrum for candidates passing preselection from the first 1.81 nb^{-1} of data recorded by LHCb at $\sqrt{s} = 7$ TeV. The fit described before is overlaid, with separate resolutions for the $D_s^{\pm}$ and $D^{\pm}$. Of the 10316 candidates total, the fit extracts 421 ± 44 $D_s^{\pm}$ candidates and 318 ± 46 $D^{\pm}$ candidates.

the processed data and Monte-Carlo produced within the LHCb simulation framework. Due to these differences an optimisation strategy on Monte-Carlo was considered sub-optimal. The use of multivariate selection techniques was also deemed unsuitable as insufficient quantities of both data and Monte-Carlo were unavailable to perform the rigorous cross-checks necessary to ensure that such a procedure would be unbiased and adequately trained. These issues called for a data-driven, simple optimisation strategy.

3.2.1. SimpleTools

Prior to the LHC startup, a number of tools were developed to aid in the selection and analysis of simulated LHCb data. These tools are designed as a collection of small, independent C++ binaries sharing a compatible input and output structure in accordance with the UNIX philosophy [71], intended to simplify and allow the scripting of common analysis tasks. Together these make up the `SimpleTools` package [72]. Input files are *NTuples*, a dataset storage format in which each candidate decay is specified by a single row, and columns for each variable are filled once each per candidate. Of particular relevance to this analysis are:

- `tuplesampler`: Accepts as input an ntuple, a random number generator seed, a sampling fraction and two output filenames. The input ntuple is sampled at random using the `ROOT::TRandom3` implementation of the Mersenne twister pseudo-random number generator [73] such that on average the first output file contains the specified sampling fraction of events and the second file is populated with the remainder.

- `sepper`: Determines and ranks according to the separation power a user-specifiable list of variables given any number of signal and background distributions. The separation power, $\langle S^2 \rangle$ is defined as [74]:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{\left(f_S(y) - f_B(y)\right)^2}{f_S(y) + f_B(y)} dy \tag{3.1}$$

  Where $f_S(y), f_B(y)$ denote the distributions of signal and background in the variable y respectively. The separation power classifies variables based on overlap such that distributions which are identical in signal and background and therefore are unsuitable to use as a cut have separation powers close to 0. Variables that have no overlap between signal and background receive a rank of 1.

- `CROP`, The Cut Recursive OPtimiser: `CROP` takes as input any number of signal or background samples, a list of discriminating variables and recursively maximises a user-specifiable figure of merit by applying rectangular cuts to the specified variables. `CROP` has a number of features that makes it preferable to other rectangular cut based optimisers, namely:

  - Speed: `CROP` uses an initialisation stage to start the recursion in a region close to the true maxima. It also uses several ordering methods to choose the next discriminating variable to reoptimise, which does not affect the final result, but which does improve the time taken to converge.

  - Transparency: `CROP` uses rectangular cuts on user-specified discriminating variables without internal transforms. During optimisation the per-variable signal efficiency, background rejection, figure-of-merit and signal efficiency as a function of background rejection are plotted and presented to the user.

  - Modularity: The figure of merit to be maximised can be provided by the user, or a pre-written FoM may be chosen. `CROP` is written in an object-oriented fashion permitting further customisation.

– Simplicity: `CROP` uses flat text files as inputs to define discriminating variables, ranges and signal/background sample locations.

Common to a number of `SimpleTools` applications is the ability to include a per-event weighting. This is user-specifiable as either a constant numeric weight, as a column in the input ntuple or as a function of several ntuple columns.

3.2.2. $\mathcal{S}$-weights and the $\mathcal{S}$-plot technique

The $\mathcal{S}$-plot technique permits statistical extraction of variable distributions on a per-species basis given a single discriminating variable from which a likelihood fit can be made [75]. The simplest use case is the extraction of the signal-only variable distributions from a sample of mixed signal and background, but the technique can be extended to more than two species. The $\mathcal{S}$-plot technique relies upon the production of $\mathcal{S}$-weights of the form:

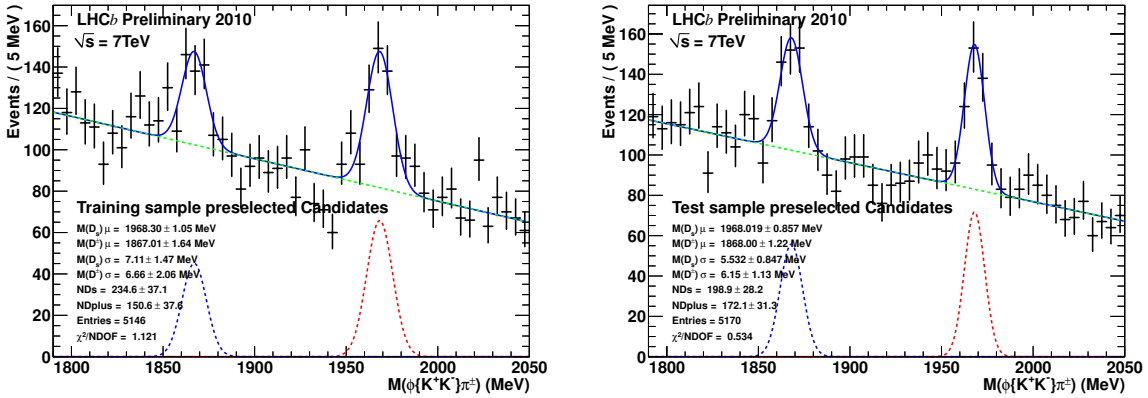
$$\mathcal{S}_i(\vec{x}_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{ij} P_j(\vec{x}_e)}{\sum_{k=1}^{N_s} n_k P_k(\vec{x}_e)} \tag{3.2}$$

Where:

- $\mathcal{S}_i(\vec{x}_e)$ is the $\mathcal{S}$-weight of the e^{th} event.
- The subscript $i = 1 \ldots N_s$ denotes the individual species one wishes to determine the weight for out of N_s total species in the sample.
- The PDFs $P_i(\vec{x})$ describe the shape of the distribution of each species in the discriminating observable(s) $\vec{x}$, where the yield of each species as determined by the likelihood fit to the combined PDF is $\sum_{i=1}^{N_s} n_i = n$. $P_i(\vec{x}_e)$ is the value of the PDF for the i^{th} species evaluated at $\vec{x}_e$.
- The $N_s \times N_s$ matrix $\mathbf{V}_{ij}$ is the covariance matrix of the species yields, which can be determined by inverting the second derivative of the likelihood function:

$$\mathbf{V}_{ij}^{-1} = \sum_{e=1}^{n} \frac{P_i(\vec{x}_e) P_j(\vec{x}_e)}{(\sum_{k=1}^{N_s} n_k P_k(\vec{x}_e))^2} = \frac{\delta^2(-\mathcal{L})}{\delta n_i \delta n_j} \tag{3.3}$$

$\mathcal{S}$-plots have a number of properties that makes their use in data-driven analyses desirable. The procedure to obtain $\mathcal{S}$-plots is as follows:



(a) Training sample **(b)** Test sample

Figure 3.3.: Mass spectra of the training and test samples of $K^+K^-\pi^{\pm}$ candidates passing the preselection in Table 3.5. From the overlaid fits $\mathcal{S}$-weights are extracted for the $D^{\pm}$,$D_s^{\pm}$ and background models.

- A discriminating variable is chosen that can be modelled for each species of event, and which is uncorrelated with the variables whose distributions one wishes to extract.
- A maximum likelihood fit is performed on the discriminating variable distribution that permits extraction of the species yields, one for each species.
- $\mathcal{S}$-weights are calculated using the covariance matrix of the fit as described above.
- The variable that one wishes to unfold is histogrammed, with each entry weighted by the $\mathcal{S}$-weight of the species that one wishes to plot.
- The error bars assigned to a given bin are the square-root of the sum of the square of the $\mathcal{S}$-weights in that bin.

In fact, the $\mathcal{S}$-weights are useful for much more than simply unfolding the variable distributions. If `SimpleTools` programs are instructed to use a per-event weighting consisting of the $\mathcal{S}$-weights of a given species, they can operate without the need for Monte-Carlo truth information, permitting a completely data-driven analysis.

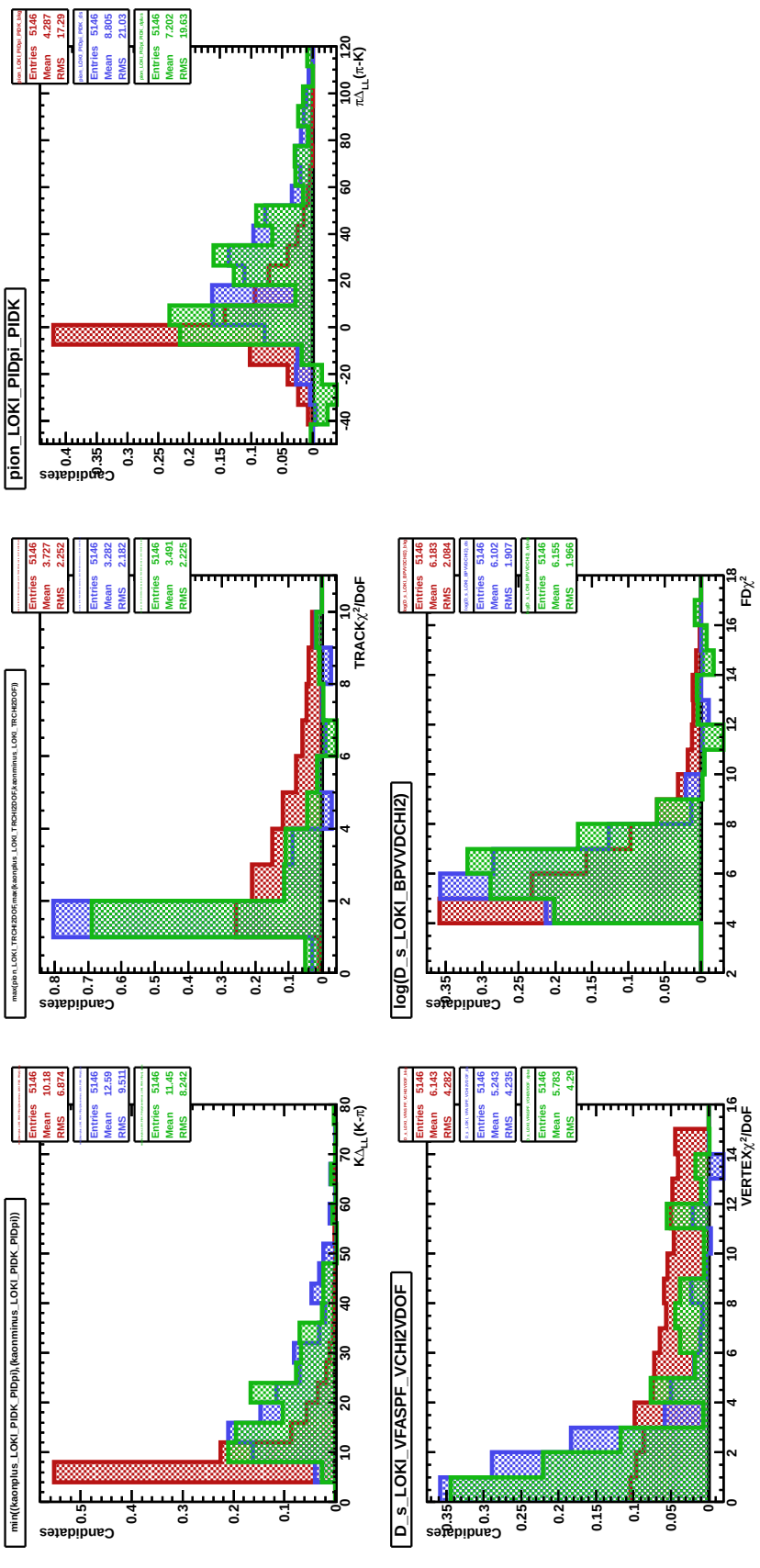


Figure 3.4.: $\mathcal{S}$-plots of discriminating variables determined to be suitable for data-driven selection optimisation of $D_s^{\pm}$, $D^{\pm}$ training sample candidates, normalised such that each species has unit area. $D_s^{\pm}$ and $D^{\pm}$ candidate $\mathcal{S}$-weights are shown in blue and green respectively, with background in red.

Discriminating Variable	Separation power
$K\Delta_{LL}(K-\pi)$	0.41
K, π Track χ^2/DoF	0.40
$\pi\Delta_{LL}(\pi-K)$	0.39
D_s, D^+ ln(Flight. Dist. χ^2)	0.27
D_s, D^+ Vert χ^2/DoF	0.1

Table 3.6.: Separation powers determined from equation (3.1) for $\mathcal{S}$-plots of discriminating variables determined to be suitable for data-driven selection optimisation of $D_s^{\pm}$,$D^{\pm}$ training sample candidates. The natural logarithm of the parent flight distance χ^2 is taken as this distribution is easier to optimise when using a uniform binning.

3.2.3. Selection optimisation

The preselected candidates written to ntuple from the full 1.81 nb^{-1} are split into two samples at random using `tuplesampler`, using a sampling ratio of 0.5. These are known as the *training* and *test* samples respectively, shown in Figure 3.3. The training sample $\mathcal{S}$-weights for $D_s^{\pm}$, $D^{\pm}$ and Background candidates are obtained by fitting to the mass distribution in Figure 3.3a. These serve as the input to `sepper`, which ranks the separation power of variables contained in the preselected ntuple. The highest-ranked variables in terms of separation power excluding p_T, momentum, rapidity and lifetime are checked for correlations with the mass, momentum and rapidity variables, leaving five variables suitable for data-driven optimisation as listed in Table 3.6. The training sample $\mathcal{S}$-plots for these variables are shown in Figure 3.4. The training sample and discriminating variables serve as input to `CROP`. For this analysis the Figure of Merit to be maximised is the signal significance, defined as the number of signal candidates over the Poissonian error on the total number of candidates, $S/\sqrt{S+B}$. The signal $\mathcal{S}$-weight distributions for both the $D_s^{\pm}$ and $D^{\pm}$ candidates determined from the fit are used simultaneously during the optimisation, and all species are reweighted by a factor of two so that the maximised $S/\sqrt{S+B}$, while determined from the training sample, is correct for the combination of both training and test samples. After each cut is re-optimised in the ensemble the $S/\sqrt{S+B}$ rises until a stable maxima is reached. The stability of the maxima is checked by randomising the optimisation order and repeating the procedure. The maximum signal significance obtained as determined from the pre-selection training sample $\mathcal{S}$-weights is 16.3 ± 1.2, and the optimal ensemble of selection

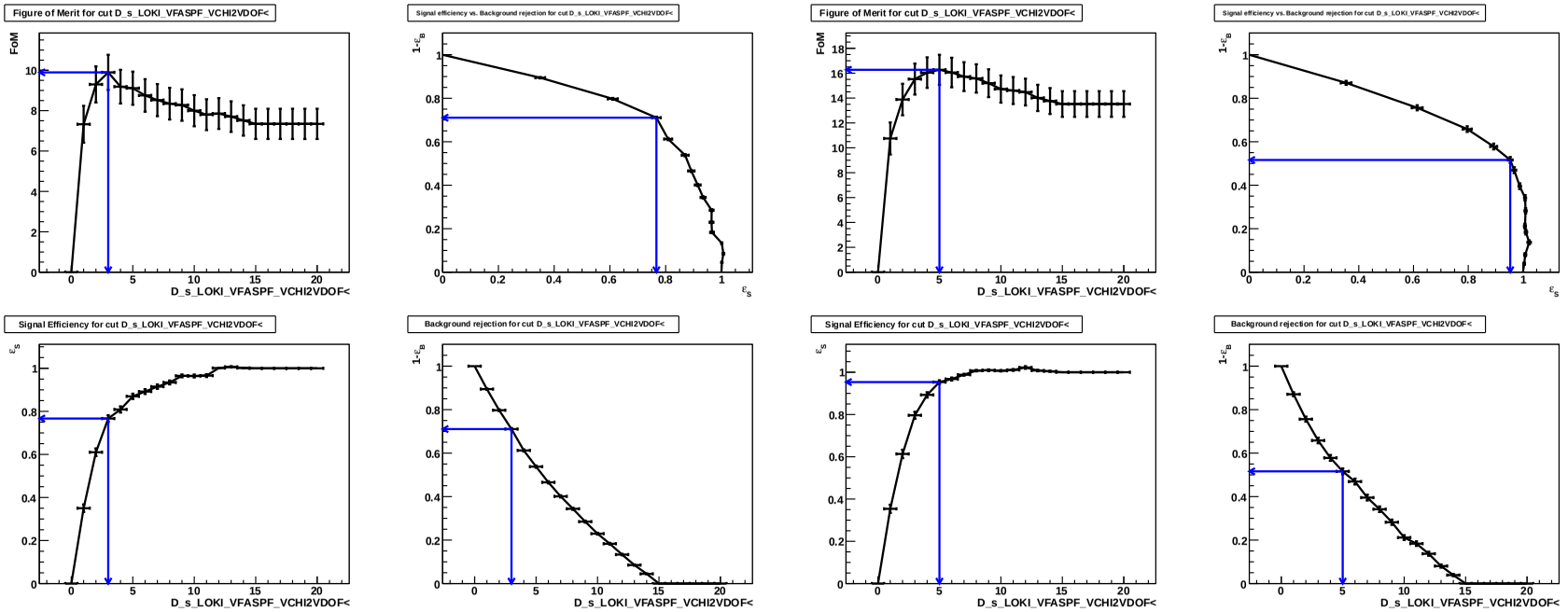
selection cut	value
$K\Delta_{LL}(K-\pi)$	> 9
D_s, D^+ Vert χ^2/DoF	< 5
K, π Track χ^2/DoF	< 4
D_s, D^+ Flight. Dist. χ^2	> 67
$\pi\Delta_{LL}(\pi-K)$	> -2

Table 3.7.: $D_s^{\pm}$, $D^{\pm}$ MC-free optimised selection.

cuts is listed in Table 3.7. After applying this selection and refitting the selected training sample, this changes slightly to 15.8 ± 1.15 based upon the new fit. The extracted yields, efficiencies and $S/\sqrt{S+B}$ for each of the training, test and combined samples are shown in Table 3.8. The good agreement between the training and test sample post optimisation indicates the absence of overtraining. The invariant mass distributions of the training and test samples after applying this selection may be found in Figure 3.6, and the combined fit to the full 1.81 nb^{-1} after selection is presented in Figure 3.7. In this sample we find 330.6 ± 20.8 $D_s^{\pm}$ and 218.2 ± 18.1 $D^{\pm}$ candidates which will be used for the determination of the cross-section and cross-section ratios.

3.3. Extraction of prompt charm yields

The selected $D^{\pm}$, $D_s^{\pm}$ candidates are a combination of those coming directly from the primary vertex and those which come from the decays of B hadrons. In order to separate the prompt candidates from those coming from secondary vertices, an additional discriminating variable is needed, from which prompt and secondary $\mathcal{S}$-weights can be determined. The Impact Parameter (IP) provides such discriminating power [57]. Selected $D_s^{\pm}$, $D^{\pm}$ candidates from 2010 Monte-Carlo samples are used to determine the lineshape of the secondary contribution as there are insufficient events in the 1.81 nb^{-1} of data to do so. Figure 3.8 illustrates that the use of an asymmetric Gaussian function (in which the width of each side of the Gaussian is a separate parameter) is sufficient to model this for both $D^{\pm}$ and $D_s^{\pm}$ candidates. The prompt charm distributions are also parameterised by an asymmetric Gaussian function whose mean is at larger values of $\ln(IP)$. The prompt and secondary asymmetric Gaussian function parameterisations are identical for both $D_s^{\pm}$ and $D^{\pm}$ candidates. Due to instabilities in the fit related to



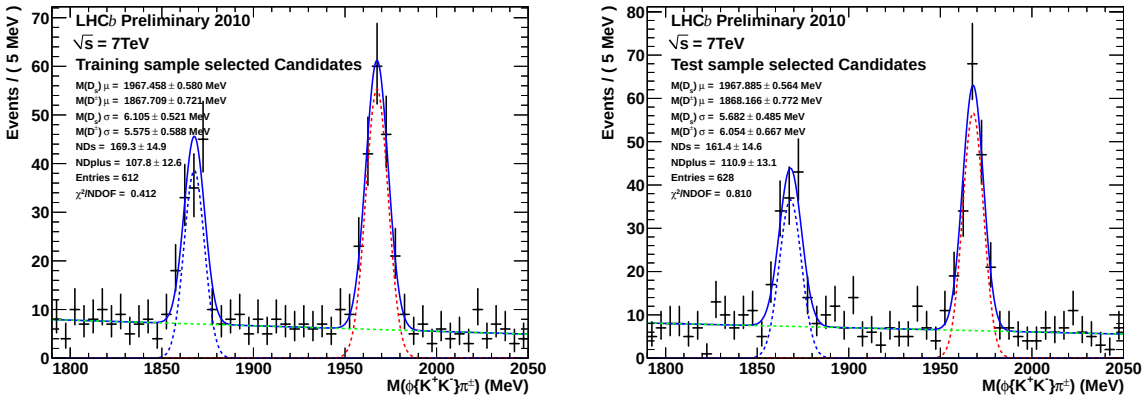
(a) Prior to ensemble optimisation

(b) After ensemble optimisation

Figure 3.5.: The effect of the optimisation procedure on a single cut in the ensemble, in this case the parent vertex χ^2 cut. Figure 3.5a shows the status of this cut after all cuts in the ensemble have been individually optimised and applied at their optimal value if used alone. Clockwise, from top left: $S/\sqrt{S+B}$ as a function of the cut position, the background rejection as a function of signal efficiency, the background rejection as a function of cut position and the signal efficiency as a function of cut position. Figure 3.5b shows the same cut after recursive reoptimisation has finished, leading to a higher $S/\sqrt{S+B}$.

Sample	$D_s^{\pm}$ Cands.	$D^{\pm}$ Cands.	Bkg. Cands	$S/\sqrt{S+B}$
Training Sample (preselection)	234.6 ± 37.1	150.6 ± 37.6	4760.7 ± 86.5	$7.35 \pm 0.75^*$
Test Sample (preselection)	198.9 ± 28.2	172.1 ± 31.3	4799.1 ± 79.4	$7.75 \pm 0.75^*$
All Preselected Candidates	421.0 ± 43.9	318.4 ± 46.2	9577 ± 115	7.55 ± 0.53
Training Sample (selection)	169.3 ± 14.9	107.8 ± 12.6	334.9 ± 21.1	$15.8 \pm 1.15^*$
Test Sample (selection)	161.4 ± 14.6	110.9 ± 13.1	355.7 ± 21.8	$15.2 \pm 1.13^*$
All Selected Candidates	330.6 ± 20.8	218.2 ± 18.1	691.1 ± 30.4	15.5 ± 0.81
	$D_s^{\pm}$ Eff.	$D^{\pm}$ Eff.	Background Eff.	
Training Sample	$72.2 \pm 2.9\%$	$71.6 \pm 3.7\%$	$7.03 \pm 0.37\%$	
Test Sample	$81.1 \pm 2.8\%$	$64.4 \pm 3.6\%$	$7.41 \pm 0.38\%$	
Total Selection Efficiency	$78.5 \pm 2.0\%$	$68.5 \pm 2.6\%$	$7.22 \pm 0.26\%$	

Table 3.8.: Yields extracted from fits shown in Figures 3.2, 3.3, 3.6 and 3.7. The * denotes $S/\sqrt{S+B}$ values computed using $\mathcal{S}$-weights that have been scaled by a factor of two in order to be consistent with the total training + test sample used for analysis. The species yields are statistically compatible between samples, and the $S/\sqrt{S+B}$ is consistent at selection level indicating no overtraining. The selection efficiencies are defined as the ratio of selection yield to preselection yield where uncertainties are assumed to be binomial.



(a) Training sample **(b)** Test sample

Figure 3.6.: Invariant $K^+K^-\pi^{\pm}$ mass distributions of the training and test samples of candidates passing the data-driven selection. The selection was optimised exclusively on the training sample and shows no signs of overtraining when applied to the test sample.

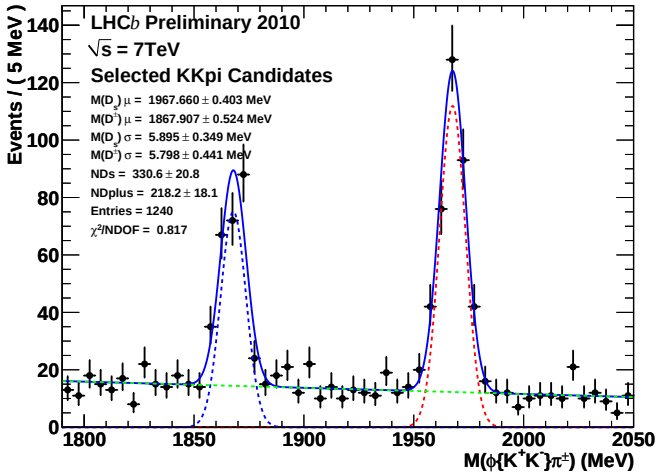
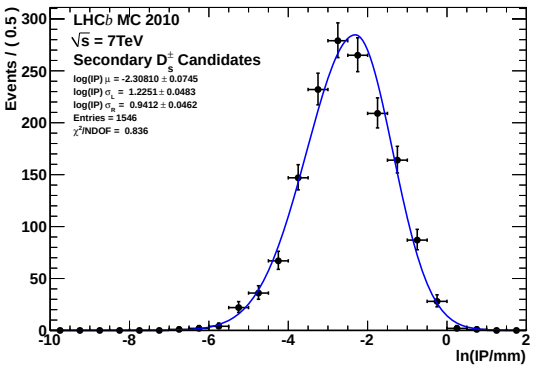
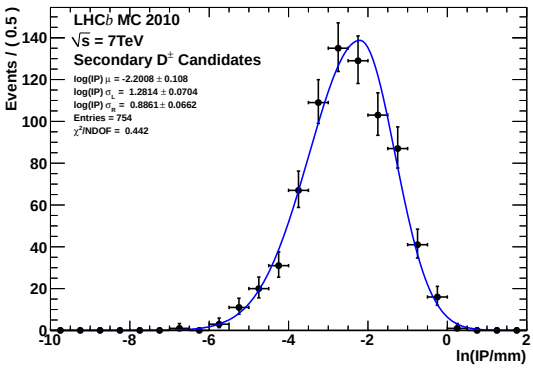


Figure 3.7.: $K^+K^-\pi^{\pm}$ invariant mass spectrum for candidates passing the data-driven optimised selection from the first 1.81 nb^{-1} of data recorded by LHCb at $\sqrt{s} = 7$ TeV. The $D_s^{\pm}$ signal distribution is modelled by a single Gaussian (red, dashed). The $D^{\pm}$ signal distribution also uses a single Gaussian shown in blue, dashed. The background component is modelled as a linear function.

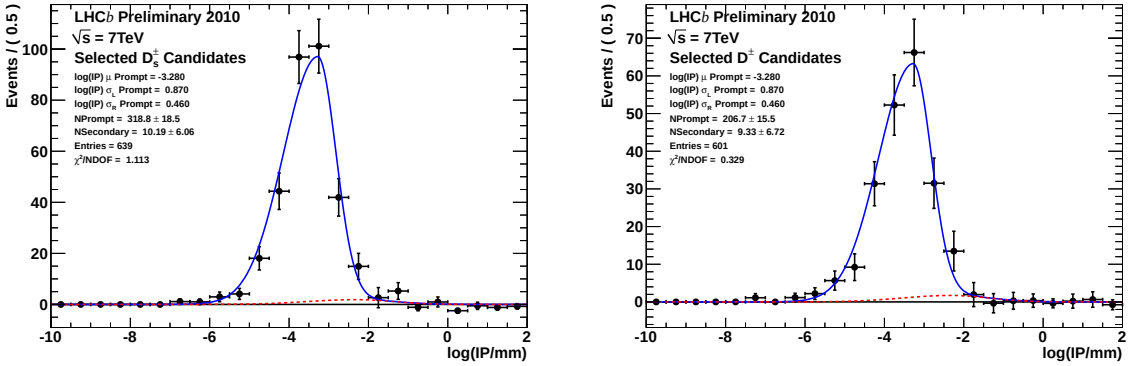


(a) MC $D_s^{\pm}$ secondaries



(b) MC $D^{\pm}$ secondaries

Figure 3.8.: The natural logarithm of the Impact Parameter distributions of selected Monte-Carlo $D_s^{\pm}$, $D^{\pm}$ candidates that have come from decays of long-lived particles. The distributions are modelled by asymmetric Gaussian functions sharing a common mean but with different widths for the left and right hand sides.



(a) $D_s^{\pm}$ $\mathcal{S}$-plot **(b)** $D^{\pm}$ $\mathcal{S}$-plot

Figure 3.9.: The natural logarithm of the Impact Parameter $\mathcal{S}$-plots of selected $D_s^{\pm}$, $D^{\pm}$ candidates used to separate prompt candidates from those produced in the decays of long-lived particles. The secondary PDF lineshape is determined from Monte-Carlo simulated data. Both prompt and secondary distributions are modelled by asymmetric Gaussian functions.

the use of bifurcated Gaussians, all parameters except for the prompt and secondary yields are fixed when extracting the prompt $\mathcal{S}$-weights. The means of the prompt and secondary components are floated prior to this in order to determine the relative mean offset between data and Monte-Carlo, which was found to be approximately $0.5\ln(IP)$. The fits to data are consistent with a small ($\approx 5\%$) secondary contribution as shown in Figure 3.9. The final prompt yields are extracted in bins of pseudorapidity and p_{T} using the product of the $\mathcal{S}$-weights from the mass fit which determines signal candidates, and the $\ln(IP)$ fit which determines prompt candidates. Making the $\mathcal{S}$-plot in bins of transverse momentum, p_{T}, and rapidity, y results in the prompt signal yields listed in Tables 3.9 and 3.10.

3.4. Efficiency estimation

The raw yields as measured by the detector are subject to trigger efficiencies, cut efficiencies and detector effects. The majority of these effects are well modelled by Monte-Carlo except for particle ID which is dependent upon varying conditions such as cavern pressure, temperature and Čerenkov gas purity, requiring calibration. The corrected yields

p_T (MeV/c)	$D_s^{\pm}$ raw yield $N_{raw}(D_s^{\pm})$	$D^{\pm}$ raw yield $N_{raw}(D^{\pm})$
$(0, 1000)$	2.4 ± 5.3	13.5 ± 5.2
$(1000, 2000)$	66.7 ± 9.8	34.6 ± 7.3
$(2000, 3000)$	66.9 ± 9.7	55.5 ± 8.5
$(3000, 4000)$	75.6 ± 9.7	43.2 ± 7.2
$(4000, 5000)$	38.9 ± 7.0	24.3 ± 5.6
$(5000, 6000)$	24.7 ± 5.8	8.4 ± 3.4
$(6000, 7000)$	13.5 ± 4.1	8.1 ± 3.1
$(7000, 8000)$	12.9 ± 4.0	3.6 ± 2.2

Table 3.9.: Prompt $D_s^{\pm}$,$D^{\pm}$ raw yields in bins of p_T, integrated over y on the range (2.0,4.5) extracted from the product of $\mathcal{S}$-weights determined from the K^+ K^- $\pi^{\pm}$ mass spectrum and $ln(IP)$ distributions.

y	$D_s^{\pm}$ yield $N_{raw}(D_s^{\pm})$	$D^{\pm}$ yield $N_{raw}(D^{\pm})$
$(2, 2.5)$	39.5 ± 7.1	34.5 ± 6.6
$(2.5, 3)$	92.1 ± 10.9	68.2 ± 9.2
$(3, 3.5)$	122.4 ± 12.4	50.1 ± 8.7
$(3.5, 4)$	44.1 ± 9.4	35.4 ± 6.9
$(4, 4.5)$	3.6 ± 4.0	3.1 ± 3.0

Table 3.10.: Prompt $D_s^{\pm}$,$D^{\pm}$ raw yields in bins of y, integrated over p_T on the range $(0, 8000)$MeV/c extracted from the product of $\mathcal{S}$-weights determined from the K^+ K^- $\pi^{\pm}$ mass spectrum and $ln(IP)$ distributions.

in transverse momentum, p_{T}, and rapidity, y can be determined using equation (3.4):

$$N_{corr} = \frac{N_{raw}}{\epsilon_{det} \times \epsilon_{L0} \times \epsilon_{HLT} \times \epsilon_{presel} \times \epsilon_{sel}} \tag{3.4}$$

Where the Level-0 and Higher level trigger efficiencies $\epsilon_{L0} \times \epsilon_{HLT} = 1$ for this analysis. $\epsilon_{det} \times \epsilon_{presel} \times \epsilon_{sel}$, the product of detector, preselection and selection efficiencies can be determined from Monte-Carlo candidates at selection level using a modified selection that excludes any PID cuts, resulting in a single efficiency term $\epsilon_{sel'}$. Due to the way in which the LHCb simulation is configured an additional geometric acceptance cut must be accounted for, ϵ_{acc}. The correction applied to obtain the efficiency corrected yields is therefore:

$$N_{corr} = \frac{N_{raw}}{\epsilon_{acc} \times \epsilon_{sel'} \times \epsilon_{PID}} \tag{3.5}$$

Where ϵ_{acc} is determined from a generator-level study, $\epsilon_{sel'}$ from signal Monte-Carlo and ϵ_{PID} is determined by reweighting signal Monte-Carlo using a data-driven PID calibration.

3.4.1. LHCb detector acceptance

The LHCb simulation framework has been designed to maximise efficiency in the production of events for Monte-Carlo analyses. Generated signal candidates are only reconstructed if they fall within the detector acceptance in order to reduce the computing load spent on events that cannot be fully reconstructed. This acceptance cut ensures that all charged decay products of the signal mode are within the LHCb acceptance of $10 < \theta < 400$ mrad. In order to determine the efficiency of the selection on Monte-Carlo particles this cut needs to be included. For both signal Monte-Carlo types listed in Table 3.2, 0.4 M events of each magnet polarity have been generated without the acceptance cut in order to determine the acceptance efficiency. Tables in Appendix A.1 present the generated and accepted yields and acceptance efficiency ϵ_{acc} for $\mathrm{D}_{\mathrm{s}}^{\pm}$ and $\mathrm{D}^{\pm}$ candidates in bins of p_{T} and y. Typical values are 80-100%.

3.4.2. Selection efficiency

The combined detector, selection and preselection efficiency is determined by applying the selection cuts listed in Tables 3.5 and 3.7 excluding PID cuts to truth-matched Monte-Carlo signal candidates resulting in the yield $N_{sel'}$ and comparing them to all candidates generated within the detector acceptance in the absence of any cuts, N_{acc}. The resulting efficiency $\epsilon_{sel'}$ in bins of y and p_{T} is tabulated for $\mathrm{D}_{\mathrm{s}}^{\pm}$ and $\mathrm{D}^{\pm}$ candidates in Appendix A.2. Typical values are $\sim 5\%$ for $\mathrm{D}_{\mathrm{s}}^{\pm}$ and 10% for $\mathrm{D}^{\pm}$, increasing with p_{T} and flat in y.

3.4.3. Particle ID efficiency

The RICH Particle ID (PID) efficiency is not well modelled in Monte-Carlo as it is dependent upon conditions such as cavern temperature and pressure and the Čerenkov gas purity, all of which vary as a function of time. In order to determine the full efficiency corrected yields the particle ID efficiency must be determined from data collected during the same period as that used by the analysis. The PID calibration procedure involves the following steps:

- Large samples of relatively pure $\mathrm{K}^{\pm}$, $\pi^{\pm}$ are selected without using PID criteria in the decay modes $\phi \to \mathrm{K}^{+}\mathrm{K}^{-}$, $\mathrm{K}_{\mathrm{S}}^{0} \to \pi^{+}\pi^{-}$ in the 1.81 nb^{-1} sample of data.
- These samples are binned in the kinematic variables p_{T}, η.
- Signal $\mathcal{S}$-plots of the $\Delta_{LL}(\mathrm{K}-\pi)$ for $\mathrm{K}^{\pm}$ and $\Delta_{LL}(\pi-\mathrm{K})$ for $\pi^{\pm}$ distributions are made in each bin using fits to the ϕ, $\mathrm{K}_{\mathrm{S}}^{0}$ mass distributions.
- The Particle ID efficiency and statistical uncertainty as a function of p_{T}, η is determined for a given PID cut using these $\mathcal{S}$-plots.
- The signal Monte-Carlo candidates selected without PID are assigned an efficiency and uncertainty for each daughter based on their kinematics: $\epsilon_{PID}(p_{\mathrm{T}}, \eta) \pm \delta\epsilon_{PID}(p_{\mathrm{T}}, \eta)$.
- The per-candidate $\mathrm{D}_{\mathrm{s}}^{\pm}$, $\mathrm{D}^{\pm}$ PID efficiency weighting is determined as the product of the daughter efficiencies, as well as the propagated uncertainty: $\omega_{PID} = \epsilon_{PID}^{\mathrm{K}^{+}} \times \epsilon_{PID}^{\mathrm{K}^{-}} \times \epsilon_{PID}^{\pi^{\pm}}$.

- The calibrated PID efficiency in p_{T}, y is determined from the ratio of the sum-of-weights in a given bin to the number of unweighted candidates in that bin. The uncertainty on the sum-of-weights is propagated as a systematic.

The choice of binning in η and p_{T} is dependent upon the calibration sample size, with a criteria that the statistical uncertainty should be not more than 5% for the central bins. Tables A.9 and A.10 show the efficiencies and choice of binning for $\mathrm{K}^{\pm}$ and $\pi^{\pm}$ candidates as obtained from the 1.81 nb^{-1} sample. These are propagated to combined PID efficiencies for the selected $\mathrm{D}_{\mathrm{s}}^{\pm}$,$\mathrm{D}^{\pm}$ Monte-Carlo samples in bins of p_{T} and y listed in Appendix A.3 where the PID statistical uncertainty is propagated as a correlated systematic uncertainty.

3.5. Cross-section determination

Dividing the raw yields by the product of the acceptance, selection and Particle ID efficiencies results in the corrected yields listed in Appendix A.4. The cross sections are then determined from the corrected yields in bins of p_{T} and y as:

$$\sigma(p_{\mathrm{T}}, y) = \frac{N_{corr}(p_{\mathrm{T}}, y)}{\int \mathcal{L} \times \mathcal{B}(\phi\pi^{\pm})} \tag{3.6}$$

Where $\int \mathcal{L} = 1.81\ \mathrm{nb}^{-1} \pm 10\%$, and $\mathcal{B}(\phi\pi^{\pm})$ is the branching ratio for $\mathrm{D}_{\mathrm{s}}^{\pm}$,$\mathrm{D}^{\pm}$ candidates to the $\phi\ \pi^{\pm}$ final state. These are taken to be:

$$\mathcal{B}(\mathrm{D}_{\mathrm{s}}^{\pm} \to \phi\pi^{\pm}) = 2.24 \pm 0.11 \pm 0.06\%\ [76] \tag{3.7}$$

$$\mathcal{B}(\mathrm{D}^{\pm} \to \phi\pi^{\pm}) = 0.271 \pm 0.011\%\ [7] \tag{3.8}$$

Applying these to the efficiency-corrected yields results in the cross-sections and cross-section ratios in bins of y and p_{T} in Tables 3.11 and 3.12 where the uncertainties are statistical only.

3.6. Systematic uncertainties

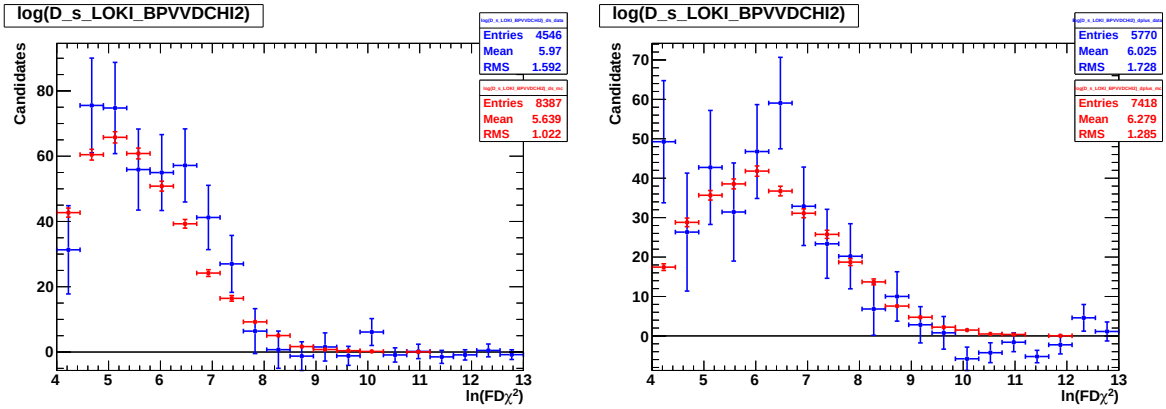
Several sources of systematic uncertainties are considered:

p_T (MeV/c)	$D_s^{\pm}$ cross-section (µb) $\sigma(D_s^{\pm})$	$D^{\pm}$ cross-section (µb) $\sigma(D^{\pm})$	Cross-section ratio $\frac{\sigma(D^{\pm})}{\sigma(D_s^{\pm})}$
$(0, 1000)$	7 ± 16	133 ± 52	18 ± 40
$(1000, 2000)$	94 ± 14	167 ± 36	1.79 ± 0.47
$(2000, 3000)$	42.1 ± 6.2	140 ± 22	3.32 ± 0.71
$(3000, 4000)$	28.4 ± 3.7	71 ± 12	2.52 ± 0.54
$(4000, 5000)$	10.9 ± 2.0	34.5 ± 8.0	3.18 ± 0.94
$(5000, 6000)$	6.0 ± 1.4	11.0 ± 4.5	1.85 ± 0.87
$(6000, 7000)$	2.82 ± 0.87	9.4 ± 3.6	3.3 ± 1.6
$(7000, 8000)$	2.46 ± 0.77	3.5 ± 2.1	1.44 ± 0.98

Table 3.11.: Prompt $D_s^{\pm}$, $D^{\pm}$ cross-section in bins of p_T, integrated over y on the range (2.0,4.5). Uncertainties are statistical only.

y	$D_s^{\pm}$ cross-section (µb) $\sigma(D_s^{\pm})$	$D^{\pm}$ cross-section (µb) $\sigma(D^{\pm})$	Cross-section ratio $\frac{\sigma(D^{\pm})}{\sigma(D_s^{\pm})}$
$(2, 2.5)$	61 ± 11	242 ± 47	3.9 ± 1.1
$(2.5, 3)$	54 ± 6.5	175 ± 24	3.26 ± 0.60
$(3, 3.5)$	60 ± 6.2	106 ± 19	1.77 ± 0.36
$(3.5, 4)$	32 ± 6.8	106 ± 21	3.35 ± 0.98
$(4, 4.5)$	9 ± 10	33 ± 32	3.6 ± 5.4

Table 3.12.: Prompt $D_s^{\pm}$, $D^{\pm}$ cross-section in bins of y, integrated over p_T on the range $(0, 8000)$MeV/c. Uncertainties are statistical only.



(a) $D_s^{\pm}$ candidates **(b)** $D^{\pm}$ candidates

Figure 3.10.: $\ln(FD\chi^2)$ distributions in data and Monte-Carlo for $D_s^{\pm}$,$D^{\pm}$ candidates at preselection level, where there is already a cut at $\ln(60) = 4.1$. The data distributions in blue are $\mathcal{S}$-plots, while the Monte-Carlo distributions in red are truth-matched candidates normalised to the data yields.

- Luminosity: The LHCb luminosity determination imparts a 10% systematic uncertainty based upon knowledge of the number of pp collisions per bunch crossing and the Van-der-Meer scan technique [77].
- Tracking Efficiency: At the time of this analysis the systematic uncertainty on the tracking efficiency as recommended by the tracking group was 3% per track, totalling 9% for the $K^+K^-\pi^{\pm}$ final state.
- PID Calibration: The average statistical uncertainty per-bin of p_T, y on the calibration sample is applied as a correlated systematic uncertainty.
- Branching ratios: The relative error on the BR measurements are taken as systematic uncertainties, corresponding to 5.8% for $D_s^{\pm}$ and 4.1% for $D^{\pm}$ respectively.
- Monte-Carlo statistical uncertainties are taken as a systematic on the Monte-Carlo efficiency determination.
- Differences between data and Monte-Carlo are assigned a correlated systematic uncertainty, described in the following subsection.

3.6.1. Data-MC differences

The majority of cuts applied at selection level are to distributions that are consistent between Monte-Carlo and data, requiring no additional calibration. The flight distance χ^2 cut is the sole exception, showing some deviation as indicated in Figure 3.10 between the $\mathcal{S}$-plot in data and the Monte-Carlo truth matched signal distribution. In order to determine a systematic uncertainty the efficiency-corrected yields listed in Tables A.15 and A.17, were again determined with this cut tightened to $\ln(90) = 4.6$ from the previous value of $\ln(67) = 4.2$. In the absence of any differences we would expect the ratio as a function of p_{T} to be flat and consistent with 1.0. We take the deviation from this as a systematic uncertainty. A fit to the ratio in bins of p_{T} gives a mean acceptance of 1.06 ± 0.03, $\chi^2/DoF = 1.6/7$ for $D_s^{\pm}$ candidates and 1.12 ± 0.05, $\chi^2/DoF = 2.8/7$ for $D^{\pm}$ candidates, resulting in 6% and 12% uncorrelated systematics respectively.

3.6.2. Total Systematic uncertainties

For the cross-section measurements the combined uncorrelated systematic uncertainties are listed in Table 3.16, with the combined correlated and uncorrelated systematics listed in Tables 3.17 and 3.18 for $D_s^{\pm}$ candidates, Tables 3.19 and 3.20 for $D^{\pm}$ candidates. For the cross-section ratio the luminosity and tracking efficiency uncertainties cancel. Partial cancellation of the flight distance acceptance systematic is determined by fitting the ratio of the $D_s^{\pm}$,$D^{\pm}$ flight distance acceptances in p_{T}, resulting in a distribution whose ratio is consistent with a linear fit whose mean is 1.07 ± 0.04, $\chi^2/DoF = 1.2/8$. This partial cancellation is taken as a 7% systematic uncertainty on the data-MC difference in the cross-section ratio. Using the same technique the systematic uncertainty due to PID efficiencies cancels completely, having a fit result of 0.995 ± 0.011, $\chi^2/DoF = 1.6/8$. The uncorrelated systematic uncertainties for the cross-section ratio are listed in Table 3.13, with total systematic uncertainties listed in Table 3.14 and Table 3.15 for p_{T} and y respectively.

Source	Ratio Systematic uncert. (%)
$\mathcal{B}(D_s^{\pm})$	5.8
$\mathcal{B}(D^{\pm})$	4.1
Flight. Dist	7.0
Total	10.0

Table 3.13.: Uncorrelated systematic uncertainties assigned to the $D^{\pm}/D_s^{\pm}$ cross-section ratio.

Ratio	Systematic uncert. (%)	
p_T (MeV/c)	MC	Total
$(0, 1000)$	5.2	11.3
$(1000, 2000)$	3.5	10.6
$(2000, 3000)$	3.3	10.5
$(3000, 4000)$	3.7	10.6
$(4000, 5000)$	4.6	11.0
$(5000, 6000)$	6.0	11.6
$(6000, 7000)$	7.3	12.4
$(7000, 8000)$	9.3	13.6

Table 3.14.: Total systematic uncertainties assigned to the $D^{\pm}/D_s^{\pm}$ cross-section ratio in bins of p_T.

$D_s^{\pm}$	Systematic uncert. (%)	
y	MC	Total
$(2, 2.5)$	4.5	11.0
$(2.5, 3)$	2.9	10.4
$(3, 3.5)$	2.9	10.4
$(3.5, 4)$	3.7	10.6
$(4, 4.5)$	6.6	12.0

Table 3.15.: Total systematic uncertainties assigned to the $D^{\pm}/D_s^{\pm}$ cross-section ratio in bins of y.

Source	$D_s^{\pm}$ Systematic uncert. (%)	$D^{\pm}$ Systematic uncert. (%)
$\mathcal{B}(\phi\pi^{\pm})$	5.8	4.1
Luminosity	10	10
Tracking	9	9
Flight. Dist	6	12
Total	15.8	18.5

Table 3.16.: Uncorrelated systematic uncertainties assigned to the $D_s^{\pm}$, $D^{\pm}$ cross-section measurements.

$D_s^{\pm}$	Systematic uncert. (%)			
p_T (MeV/c)	MC	PID	Total Corr.	Total
$(0, 1000)$	3.9	1.1	4	16.3
$(1000, 2000)$	2.6	1.2	2.9	16.1
$(2000, 3000)$	2.3	1.4	2.7	16.0
$(3000, 4000)$	2.5	2.1	3.3	16.1
$(4000, 5000)$	3.0	2.8	4.1	16.3
$(5000, 6000)$	3.8	3.6	5.2	16.6
$(6000, 7000)$	4.8	3.9	6.1	16.9
$(7000, 8000)$	5.9	4.1	7.2	17.4

Table 3.17.: Total systematic uncertainties assigned to the $D_s^{\pm}$ cross-section in bins of p_T.

$D_s^{\pm}$	Systematic uncert. (%)			
y	MC	PID	Total Corr.	Total
$(2, 2.5)$	3.0	2.8	4.2	16.3
$(2.5, 3)$	1.9	2.0	2.8	16.0
$(3, 3.5)$	2.0	1.8	2.7	16.0
$(3.5, 4)$	2.6	2.0	3.3	16.1
$(4, 4.5)$	4.5	3.7	5.8	16.8

Table 3.18.: Total systematic uncertainties assigned to the $D_s^{\pm}$ cross-section in bins of y.

$D^{\pm}$	Systematic uncert. (%)			
p_T (MeV/c)	MC	PID	Total Corr.	Total
$(0, 1000)$	3.3	1.1	3.5	18.8
$(1000, 2000)$	2.3	1.2	2.6	18.7
$(2000, 3000)$	2.4	1.4	2.7	18.7
$(3000, 4000)$	2.6	2.1	3.3	18.8
$(4000, 5000)$	3.5	2.8	4.5	19.0
$(5000, 6000)$	4.6	3.5	5.7	19.4
$(6000, 7000)$	5.5	4.0	6.9	19.7
$(7000, 8000)$	7.1	4.0	8.2	20.2

Table 3.19.: Total systematic uncertainties assigned to the $D^{\pm}$ cross-section in bins of p_T.

$D^{\pm}$	Systematic uncert. (%)			
y	MC	PID	Total Corr.	Total
$(2, 2.5)$	3.2	2.5	4.1	19.0
$(2.5, 3)$	2.1	1.7	2.7	18.7
$(3, 3.5)$	2.0	1.6	2.6	18.7
$(3.5, 4)$	2.6	1.9	3.2	18.8
$(4, 4.5)$	4.9	3.5	6.0	19.4

Table 3.20.: Total systematic uncertainties assigned to the $D^{\pm}$ cross-section in bins of y.

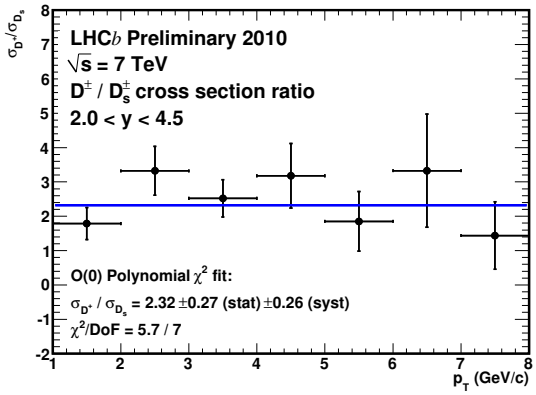


Figure 3.11.: $D^{\pm}/D_s^{\pm}$ Cross-section ratio in bins of p_T integrated over the range $2.0 < y < 4.5$. The ratio is flat on the range, as indicated by a fit to an $\mathcal{O}(0)$ polynomial.

3.7. Results

3.7.1. Cross-section ratio

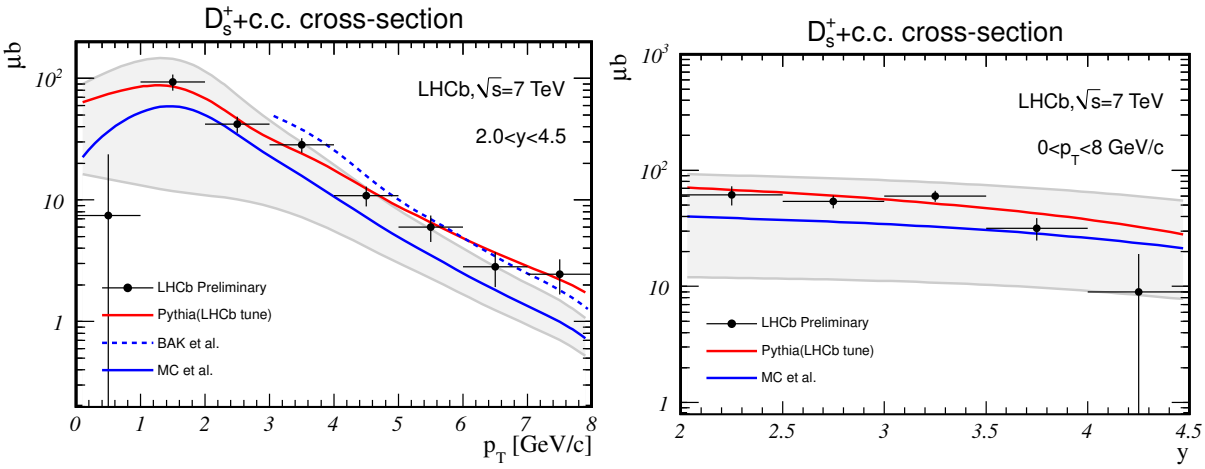
The cross-section ratio is observed to be flat in p_T as shown in Figure 3.11. The result of a linear fit to this ratio determines the average $D^{\pm}/D_s^{\pm}$ cross-section ratio over the range $1000 < p_T(\text{MeV}/c) < 8000$, $2.0 < y < 4.5$ to be:

$$\frac{\sigma(D^{\pm})}{\sigma(D_s^{\pm})} = 2.32 \pm 0.27(\text{stat}) \pm 0.26(\text{syst}) \tag{3.9}$$

This is consistent with the ratio of transition probabilities for charm quarks to $D^{\pm}$, $D_s^{\pm}$ mesons: $f(c \to D^+)/f(c \to D_s^+) = 3.08 \pm 0.70$ [7].

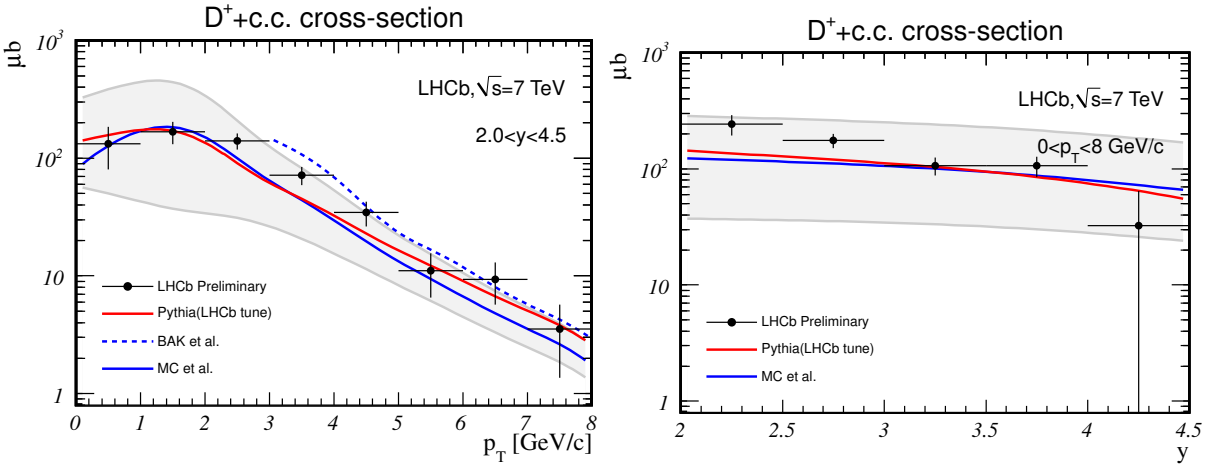
3.7.2. Cross sections in p_T and y

The cross-section measurements are presented in Figures 3.12 and 3.13. These are compared to theoretical expectations calculated at NLO by [78] and [79] integrated over bins in y and p_T. The calculations of reference [79] contain estimates of theoretical uncertainties from mass and scale variations. Uncertainties due to the choice of the parton distribution functions are not included and are expected to be small. Uncertainties due to higher order QCD effects, which can be estimated by comparing FONLL-calculations with shower Monte Carlos are not included either. While expected to be small in most



(a) $D_s^{\pm}$ cross-section in p_T **(b)** $D_s^{\pm}$ cross-section in y

Figure 3.12.: $D_s^{\pm}$ Cross sections in bins of p_T and y. Black points are the experimental result. The blue dashed line indicates the calculation performed by [78], available only in p_T. The Blue solid line with error bands in grey is from [79]. The cross-section determined from the LHCb tune of PYTHIA is shown in red.



(a) $D^{\pm}$ cross-section in p_T **(b)** $D^{\pm}$ cross-section in y

Figure 3.13.: $D^{\pm}$ Cross sections in bins of p_T and y. Black points are the experimental result. The blue dashed line indicates the calculation performed by [78], available only in p_T. The Blue solid line with error bands in grey is from [79]. The cross-section determined from the LHCb tune of PYTHIA is shown in red.

regions of phase space, they may become important at large rapidities. Additionally, we compare the results to the LHCb tuning of PYTHIA [80].

Chapter 4.

Experimental determination of ϕ_s in the decay $B_s^0 \to J/\psi\,\phi$

"In [chess], where the pieces have different and "bizarre" motions, with various and variable values, what is only complex, is mistaken (a not unusual error) for what is profound."

— Edgar Allan Poe

Any departure from the SM expectation of $\phi_s = -2\beta_s$ would be an unambiguous signal of physics beyond the SM as discussed in Chapter 1, but the act of measuring ϕ_s is far from straightforward. This chapter will describe the time-dependent angular analysis required to perform such a measurement, starting with the selection procedure to obtain a sample of $B_s^0 \to J/\psi\,\phi$ candidates and following with a description of the fitting techniques used to extract ϕ_s.

4.1. Datasets, triggers and selection

The dataset used in this analysis consists of 0.37 fb^{-1} of proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV collected during the first half of 2011. The data were processed under Reco10-Stripping13b with DaVinci v28r2p2. At the time of this analysis Monte-Carlo simulated data consistent with the 2011 data taking conditions was unavailable. The 2010 Monte-Carlo production originally produced in the MC10-Sim01 round was reprocessed using the Trigger Configuration Key (TCK) 0x00730035,

Decay Type	$B_s^0 \to J/\psi(\to \mu^-\mu^+)\phi(\to K^+K^-)$
Monte-Carlo Configuration	MC10-Sim01
Gauss version	29r0
Boole version	21r0
Brunel version	39r4
Moore version	12r6
Moore TCK	0x00730035
DaVinci version	28r3

Table 4.1.: The full software configuration used to generate Monte-Carlo $B_s^0 \to J/\psi\,\phi$ signal events for the ϕ_s analysis.

Physics parameter	value
$\phi_s = -2\beta_s$	-0.04 rad
$\Delta\Gamma_s$	0.060 ps^{-1}
Γ_s	0.6793 ps^{-1}
Δm_s	17.8 ps^{-1}
$\lvert A_0\rvert^2$	0.60
$\lvert A_\parallel\rvert^2$	0.24
$\lvert A_\perp\rvert^2$	0.16
δ_0	0 rad
$\delta_\parallel$	2.50 rad
$\delta_\perp$	-0.17 rad

Table 4.2.: Values of physics parameters used in the generation of Monte-Carlo $B_s^0 \to J/\psi\,\phi$ samples. These parameters are defined in Section 4.3.1

which is consistent with the trigger conditions during data taking. Table 4.1 lists the software configuration used to produce the Monte-Carlo, and Table 4.2 lists the values of the physics parameters with which it was generated.

Decay mode	Cut parameter	Stripping	Selection
all tracks	$\chi^2_{\text{track}}/\text{nDoF}$	< 5	< 4
	clone distance	–	> 5000
$J/\psi \to \mu^-\mu^+$	$\Delta_{LL}(\mu - \pi)$	> 0	> 0
	$min(p_T(\mu^+), p_T(\mu^-))$ [GeV/c]	–	> 0.5
	$\chi^2_{\text{vtx}}/\text{nDoF}(J/\psi)$	< 16	< 16
	$\lvert M(\mu^+\mu^-) - M(J/\psi)[\text{MeV}/c^2]\rvert$	< 80	$\in [3030, 3150]$
$\phi \to K^+K^-$	$\Delta_{LL}(K - \pi)$	> -2	> 0
	p_T (ϕ) [GeV/c]	> 1	> 1
	$M(\phi)$ [MeV/c^2]	$\in [980, 1050]$	$\in [1007.46, 1031.46]$
	$\chi^2_{\text{vtx}}/\text{nDoF}(\phi)$	< 16	< 16
$B_s^0 \to J/\psi\,\phi$	$M(B_s^0)[\text{MeV}/c^2]$	$\in [5100, 5550]$	$\in [5200, 5550]$
	$\chi^2_{\text{vtx}}/\text{nDoF}(B_s^0)$	< 10	< 10
	$\chi^2_{\text{DTF(B+PV)}}/\text{nDoF}(B_s^0)$	–	< 5
	$\chi^2_{\text{IP}}(B_s^0)$	–	< 25
	$\chi^2_{\text{IP,next}}(B_s^0)$	–	> 50
	t [ps]	-	> 0.3
Candidates passing		3405560	14042
	proper time unbiased		11029
	proper time biased		1754

Table 4.3.: Stripping and selection criteria for $B_s^0 \to J/\psi\,\phi$ candidates. The last rows indicate the number of candidates passing specific trigger lines described in Section 4.1.2.

4.1.1. Stripping and selection

The selection procedure for $B_s^0 \to J/\psi\,\phi$ has evolved iteratively from MC studies made prior to the LHC startup [81] The selection is designed to maximise yield in the signal channel while retaining compatibility with control channels used for flavour tagging calibration. The selection is listed in Table 4.3 and summarised here. For tracks used to make daughter candidates:

- The track fit χ^2 per degree of freedom must be less than 4.
- The Kullback-Liebler distance [82] between tracks must be greater than 5000 to reject clones.

The $J/\psi \to \mu^+\mu^-$ candidates are made by requiring that:

- The particle ID of candidates must favour muons over pions.
- The transverse momentum of muon candidates must be greater than 500 MeV/c.
- The vertex fit to the J/ψ using the two $\mu^\pm$ tracks must have a χ^2 per degree of freedom less than 16.
- The J/ψ candidate invariant mass must be on the range $3030 < M(\mu^+\mu^-) < 3150$ MeV/c^2.

$\phi \to K^+K^-$ candidates are made requiring that:

- The particle ID of candidates must favour Kaons over pions.
- The vertex fit to the ϕ using the two $K^\pm$ tracks must have a χ^2 per degree of freedom less than 16.
- The transverse momentum of the ϕ candidate must be greater than 1 GeV/c.
- The invariant mass of the K^+K^- combination must be within 12 MeV/c^2 of the PDG value of the ϕ mass.

Finally the $B_s^0 \to J/\psi\phi$ combination is subject to the following cuts:

- The B_s^0 vertex fit χ^2 per degree of freedom must be less than 10.

- The global kinematic fit to the B_s^0 candidate in which the primary vertex is constrained prior to fitting the decay tree must have an impact parameter χ^2 per degree of freedom less than 5.
- The impact parameter χ^2 with respect to the best primary vertex must be less than 25.
- The impact parameter χ^2 with respect to any other primary vertex must be greater than 50.
- The invariant mass of the $J/\psi\phi$ combination must be on the range $5200 < M(J/\psi\phi) < 5550$ MeV/c^2.

The candidate multiplicity per event containing at least one candidate is on average 1.06 regardless of the event complexity. In order to remove duplicate candidates a single $B_s^0 \to J/\psi\,\phi$ candidate is chosen per event. We chose the "best" candidate to keep based on the smallest χ^2 returned by the simultaneous fit to the fully reconstructed decay chain's vertices, momenta and track properties [83]. This procedure picks the correct candidate $> 99.5\%$ of the time when applied to Monte-Carlo data where the candidate multiplicity is 1.04. Lastly, a proper time cut of 0.3 ps^{-1} is applied offline in order to reject J/ψ candidates coming directly from the primary vertex which pick up two additional tracks to form a fake B_s^0 with zero lifetime.

4.1.2. Triggers

The presence of two muons in the $B_s^0 \to J/\psi\,\phi$ final state is fortuitous as the muon system discussed in Chapter 2 permits fast and efficient triggering of this type of decay. Two trigger combinations are used, colloquially referred to as the *proper time biased* and *proper time unbiased* triggers respectively, derived from the combination of trigger criteria listed in Tables 4.4, 4.5 and 4.6. The proper time biased trigger consists of the combination `(Hlt2DiMuonJPsi && (Hlt1TrackMuon || Hlt1TrackAllL0))`, while the proper time unbiased trigger consists of the combination `(Hlt2DiMuonJPsi && Hlt1DiMuonHighMass)`. In both cases it is a requirement that the passing trigger was explicitly of the type "Trigger on Signal" (TOS), meaning that the trigger decision was explicitly and only due to the properties of the $B_s^0 \to J/\psi\,\phi$ candidate. The proper time unbiased trigger line is approximately 83% efficient at HLT1. The loss of 17% of candidates at HLT1 incurred by using only proper time unbiased lines motivates the inclusion of the additional proper time proper time biased trigger. This recovers an additional

cut	`Hlt1DiMuonHighMass`
L0 decision	muon || dimuon
$\mu^\pm$ p_T	> 0.5 GeV/c
$\mu^\pm$ p	> 6.0 GeV/c
$\mu^\pm$ IsMuon	True
$M(\mu^-\mu^+)$	> 2.7 GeV/c^2

Table 4.4.: HLT1 requirements for the trigger line used in the proper time unbiased trigger.

cut	`Hlt1TrackAllL0`	`Hlt1TrackMuon`
L0 decision	physics	muon || dimuon
IP	> 0.1mm	> 0.1mm
# Tracker hits	> 16	
# Velo hits	> 9	-
# Missed Velo hits	< 3	-
p_T	> 1.7 GeV/c	> 1.0 GeV/c
p	> 10.0 GeV/c	> 8.0 GeV/c
$IP\chi^2$	> 16	> 16
IsMuon	-	True

Table 4.5.: HLT1 requirements for trigger lines used in the proper time biased trigger.

15% of selected candidates where the use of an impact parameter cut biases the lifetime distribution. Both the proper time biased and proper time unbiased lines use the same `HLT2DiMuonJPsi` line which is 92% efficient on TOS candidates using the cuts listed in Table 4.6.

cut	`Hlt2DiMuonJPsi`
$\lvert M(J/\psi) - M(\mu^-\mu^+)\rvert$	< 150MeV/c^2
J/ψ $\chi^2_{\rm vtx}$/nDoF	< 25
$\mu^\pm$ Track χ^2/nDoF	< 5

Table 4.6.: HLT2 requirements for both proper time biased and proper time unbiased triggers used to select $B_s^0 \to J/\psi\,\phi$ candidates.

4.2. Fitting

4.2.1. The Likelihood

The statistical extraction of parameters from a set of observables can be achieved in a number of ways depending upon the specifics of a given analysis. In High Energy Physics the most common approach is to use the technique of Maximum Likelihood [84] [85]. The likelihood function is defined as:

$$\mathcal{L}(\vec{\alpha}) = \prod_{i}^{N} \left[P(\vec{x}_i; \vec{\alpha})\right] \tag{4.1}$$

Where $\vec{x}$ is a set of observables such that every event has a value for each observable $\vec{x}_i$ and N is the number of events. $P(\vec{x}; \vec{\alpha})$ is the probability density function (PDF) where $\vec{\alpha}$ is the set of unknown parameters that we wish to determine. The notation $\vec{x}; \vec{\alpha}$ is used to imply that the PDF is conditional on the values of $\vec{\alpha}$. The subscript i is included when explicitly referring to the value of the PDF at a point specified by $\vec{x}_i$, and excluded when referring to the general form of the PDF over all phase-space. The PDF has two main properties: It is equal to or greater than 0 across the entire phase space of dimension $\vec{x}$, and it is normalised to unity. Typically when constructing such a PDF the normalisation must be explicitly computed:

$$P(\vec{x}; \vec{\alpha}) = \frac{p(\vec{x}; \vec{\alpha})}{\int p(\vec{x}; \vec{\alpha}) d\vec{x}} \tag{4.2}$$

Where $p(\vec{x}; \vec{\alpha})$ is the unnormalised PDF describing the observables $\vec{x}$ in terms of parameters $\vec{\alpha}$. By maximising $\mathcal{L}(\vec{\alpha})$ with respect to the parameters $\vec{\alpha}$, we obtain a fit to $P(\vec{x}; \vec{\alpha})$. The probability for an individual event $P(\vec{x}_i; \vec{\alpha})$ can be small, which leads to issues with computational precision when the product is taken over the entire dataset. For this reason it is more common to use the log-likelihood in fits as the sum is less likely to lead to problems with precision:

$$\ln \mathcal{L}(\vec{\alpha}) = \sum_{i}^{N} \ln \left[P(\vec{x}_i; \vec{\alpha})\right] \tag{4.3}$$

It is more common to minimise the negative log-likelihood (NLL) as a number of function minimisation algorithms are readily available [86]. The procedure then consists of finding

the lowest value of $-\ln \mathcal{L}(\vec{\alpha})$ for which

$$\frac{d}{d\vec{\alpha}} \sum_{i}^{N} -\ln\left[P(\vec{x}_i;\vec{\alpha})\right] = 0 \tag{4.4}$$

$P(\vec{x};\vec{\alpha})$ may consist of the product of independent, uncorrelated normalised probability distributions based upon the dimensionality of $\vec{x}$ such that:

$$P(\vec{x};\vec{\alpha}) = P(\vec{x}^1;\vec{\alpha}^1) \times P(\vec{x}^2;\vec{\alpha}^2) \times P(\vec{x}^3;\vec{\alpha}^3)\ldots \tag{4.5}$$

where $\vec{x}^j$ and $\vec{\alpha}^j$ are the ensemble of observables and parameters specifying the j_{th} PDF. It is therefore possible to build multidimensional PDFs from products of PDFs with one or more dimension assuming that the combined PDF can be factorised into components.

4.2.2. Fitting to several species

For a dataset consisting of two subsets (for example signal and background) having the same dimensionality $\vec{x} = \vec{x}_{sig} = \vec{x}_{bkg}$ but with different distributions the PDF becomes:

$$P(\vec{x};[\vec{\alpha}_{sig},\vec{\alpha}_{bkg},f_{sig}]) = f_{sig}P(\vec{x};\vec{\alpha}_{sig}) + (1-f_{sig})P(\vec{x};\vec{\alpha}_{bkg}) \tag{4.6}$$

where f_{sig} is the signal fraction of the sample, $f_{sig} = N_{sig}/N$ with N_{sig} the number of signal candidates and $\vec{\alpha}_{sig}$ denotes the set of parameters that specifies the signal PDF.

4.2.3. Nuisance parameters

Nuisance parameters are the subset $\vec{\alpha}_n$ that are required parameters of the PDF but are not directly of interest: For example, the parameters that describe the background model in a fit must be known in order to permit extraction of the signal parameters. Where possible such parameters are fixed from external sources but it is not always possible or desirable to do so.

4.2.4. Simultaneous fitting and constraints

It is frequently the case that $P(\vec{x};\vec{\alpha})$ depends upon parameters for which the PDF provides no or limited sensitivity. If external information is available for these parameters they can be included in the fit in three ways:

- Fixing the parameter to a known value,
- Introducing a constraint on the known value and its uncertainty which penalises the likelihood, or
- Simultaneously fitting to data that further constrains the parameter.

Fixing to a known value introduces a systematic uncertainty which needs to be quantified. For a parameter α_p belonging to $\vec{\alpha}$ which has been measured elsewhere to be $\alpha'_p \pm \sigma_{\alpha'_p}$ a Gaussian constraint can be used to modify the NLL minimisation like so [87]:

$$-ln\mathcal{L}(\vec{\alpha}) = \sum_i^N -\ln\left[P(\vec{x}_i;\vec{\alpha})\right] + \frac{(\alpha_p - \alpha'_p)^2}{2\sigma^2_{\alpha'_p}} \tag{4.7}$$

This method is desirable to fixing the parameter as it propagates the systematic uncertainly directly into the fit result. If the dataset provides sensitivity to α_p the constraint acts to combine the measurements. In the same way, additional constraints can be added to the NLL, one for each parameter for which information is already available. Such a technique is effective if the uncertainty on α'_p is truly Gaussian. For parameters in which this is not the case, the full Bayesian posterior density should instead be used. This is equivalent to simultaneously fitting for α_p in a combined measurement. More generally a simultaneous fit to two datasets M, N described by PDFs $P_m(\vec{x}^m;\vec{\alpha}^m)$, $P_n(\vec{x}^n;\vec{\alpha}^n)$ in which one or more parameters $\vec{\alpha}_{\text{sim}}$ are in both $\vec{\alpha}^m$ and $\vec{\alpha}^n$ can be fit by minimising:

$$-ln\mathcal{L}(\vec{\alpha}^{m+n}) = \sum_i^M -\ln\left[P(\vec{x}^m_i;\vec{\alpha}^m)\right] + \sum_j^N -\ln\left[P(\vec{x}^n_j;\vec{\alpha}^n)\right] \tag{4.8}$$

4.2.5. Parameter point-estimate uncertainties

Minimisation of the negative log-likelihood is performed using MINUIT [86], which estimates parameter errors based on the curvature of the likelihood minimum. Expanding

about the parameter minimum $\alpha_{\min}$ of the negative log-likelihood:

$$\begin{aligned} -\ln\mathcal{L}(\alpha) = \phi(\alpha) &\approx \phi(\alpha_{\min}) + \left.\frac{\delta\phi}{\delta\alpha}\right|_{\alpha_{\min}} (\alpha-\alpha_{\min}) + \left.\frac{\delta^2\phi}{\delta\alpha^2}\right|_{\alpha_{\min}} \frac{1}{2}(\alpha-\alpha_{\min})^2 \\ &\approx \phi(\alpha_{\min}) + \frac{1}{2}\frac{1}{\beta^2}(\alpha-\alpha_{\min})^2 \end{aligned} \tag{4.9}$$

where the first derivative has vanished by definition of the minimum and we have defined β^2 as the inverse of the second derivative. Exponentiating this to get back to the likelihood we have:

$$\begin{aligned} \mathcal{L}(\alpha) &\approx \exp\left[-\phi(\alpha_{\min})\right] \cdot \exp\left[-\frac{1}{2}\frac{1}{\beta^2}(\alpha-\alpha_{\min})^2\right] \\ &\approx k\exp\left[-\frac{(\alpha-\alpha_{\min})^2}{2\beta^2}\right] \end{aligned} \tag{4.10}$$

This is a Gaussian centered on the parameter minimum $\alpha_{\min}$ with standard deviation $(\delta^2\phi/\delta\alpha^2|_{\alpha_{\min}})^{-2}$. For a Gaussian likelihood function the likelihood minima is parabolic in the limit of large statistics, $k \to 1/\sqrt{2\pi}\sigma$, and so this yields an accurate estimate of the parameter error. This is the technique used to determine point-estimate uncertainties in the ϕ_s analysis in the case where the likelihood is found to be parabolic.

4.2.6. Confidence interval estimation using the Likelihood ratio

The point-estimate technique described in the previous section is suitable for symmetric, parabolic minima arising from a Gaussian likelihood function. Determining if such a technique is suitable requires knowledge about the shape of the minima, and a method of treatment in the case that these criteria are not met. The *profile likelihood* or likelihood scan technique permits both examination of shape of the likelihood minimum and determination of an approximate confidence interval. The profile likelihood is defined as:

$$-\Delta\ln\mathcal{L}(\alpha) = \frac{\mathcal{L}(\alpha)}{\mathcal{L}(\alpha_{\min})} = \ln\mathcal{L}(\alpha_{\min}) - \ln\mathcal{L}(\alpha) \tag{4.11}$$

By determining the negative log-likelihood at fixed values of α with all other nuisance parameters free in the fit, and subtracting off the value at the minimum we construct the shape of the likelihood minimum as a function of α. An approximate confidence

C.L.	σ	$-\Delta \ln \mathcal{L}$ (1D)	$-\Delta \ln \mathcal{L}$ (2D)
68%	1	0.5	1.15
90%		1.35	2.31
95%	2	1.92	3.0
99.7%	3	4.41	5.95

Table 4.7.: Confidence limits as likelihood ratios for one and two-dimensions.

interval can then be constructed noting that for a Gaussian likelihood function

$$\int_{-a}^{a} \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right) dx = \mathrm{erf}(a/\sqrt{2}) \tag{4.12}$$

gives the probability that the true value α lies within the range $[x_0 - a, x_0 + a]$ of the measured parameter x_0. In terms of the profile likelihood, the confidence interval can be determined as the range of x for which:

$$-\Delta \ln \mathcal{L}(\alpha) < a^2/2 \tag{4.13}$$

Table 4.7 lists the values of the profile likelihood for the one, two and three sigma corresponding confidence intervals as likelihood ratio values.

4.2.7. Feldman-cousins coverage correction

In constructing confidence intervals using the profile likelihood technique, an important statistical issue has been overlooked: By providing the likelihood ratio confidence interval we are answering the question "given that we have measured the value x_0, what is the probability that α lies in the range $[x_0 - a, x_0 + a]$?" This is not what we wish to determine: It is more pertinent to ask "given x_0, what is the probability that it lies within the range $[\alpha - a, \alpha + a]$?". The Neyman construction allows the range $[\alpha - a, \alpha + a]$ compatible with a measurement of x_0 to be constructed by way of confidence belts, for which the procedure is to determine the confidence regions for all possible values of α, and then determine the coverage for x_0 as the maximum and minimum values of α for which x_0 intercepts the confidence region. An example of such a confidence belt is shown in Figure 4.1. A problem with the Neyman construction is that it is up to the user to choose whether or not a one or two-sided confidence limit is used: A common

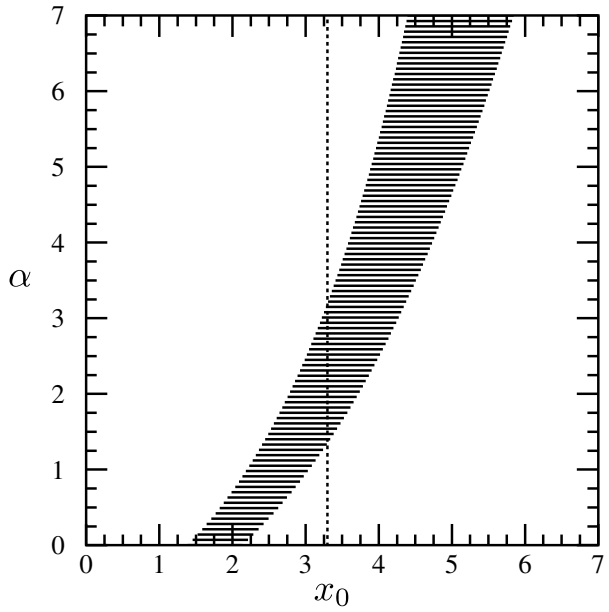


Figure 4.1.: Example of the Neyman construction for determining coverage [88]. For every possible value of α a confidence interval is determined, forming a "confidence belt" in the $\alpha - x$ plane. The measured value of α, x_0, is then compatible with values of α at the specified confidence level for the range of α whose confidence intervals are intercepted by x_0.

error is to decide what limit is used based on the measured value of α, as might be the case for a signal yield in which two-sided confidence limits at low yields may produce unphysical limits. Such a procedure will undercover, artificially inflating the sensitivity of the experiment. Feldman and Cousins [88] solved this problem by considering that confidence regions can be chosen until the correct coverage is reached, and proposed an ordering method based upon the likelihood ratio. The procedure is to choose which points to include in the interval $[x_1, x_2]$ based on order of descending R, where R is the likelihood ratio:

$$R = \ln \mathcal{L}(\alpha) - \ln \mathcal{L}(\alpha_{\text{best}}) \tag{4.14}$$

where α_{best} is the value of α that minimises the NLL. The procedure would then be to determine R for all possible values of α, order them and include values of α in the interval until the correct coverage level is obtained. The implementation of this procedure for the ϕ_s analysis in two dimensions is as follows:

- The x, y plane is divided into discrete points to form an $m \times n$ grid in x_i, y_j.
- The data is fit to with x, y free, yielding NLL_{best}^{data}.

- The data is fit to once at each gridpoint with x, y fixed producing $\mathrm{NLL}^{data}{}_{i,j}$ for all $m \times n$ gridpoints.
- The PDF is used to generate Monte-Carlo "toy' [1] datasets of equivalent size to the selected dataset at each i, j point in the plane, a large number of times. For each toy dataset the other physics and nuisance parameters are generated with the values that minimised $\mathrm{NLL}^{data}{}_{i,j}$.
- Each toy dataset is fit to twice: Once with x, y fixed to their "true" generated values at this point, and once with them left to float to their "best" values. This produces two NLL values for each toy: $\mathrm{NLL}^{toy}_{true\,i,j}$ and $\mathrm{NLL}^{toy}_{best\,i,j}$.
- The likelihood ratios:

$$R^{data}_{i,j} = \mathrm{NLL}^{data}{}_{i,j} - \mathrm{NLL}^{data}_{best} \tag{4.15}$$

$$R^{toy}_{i,j} = \mathrm{NLL}^{toy}_{true\,i,j} - \mathrm{NLL}^{toy}_{best\,i,j} \tag{4.16}$$

  are constructed, leading to a single ratio for data, and one for each toy generated. A p-value is then determined by counting how many toy datasets at this gridpoint have $R^{toy}_{i,j} > R^{data}_{i,j}$:

$$p_{i,j} = (R^{toy}_{i,j} > R^{data}_{i,j})/N_{toys} \tag{4.17}$$

- The maximum confidence interval for which this i, j gridpoint provides coverage is then $1 - p$.

The Feldman-Cousins procedure is extremely computationally intensive, requiring $m \times n \times N_{toys}$ toy datasets to be generated, and $m \times n \times N_{toys} \times 2$ fits to those toy datasets, plus an additional $m \times n$ fits to data.

4.3. Time and angular dependent decay formalism

The parameters of interest, ϕ_s, $\Delta\Gamma_s$ arise from the time-dependent interference between mixing and decay discussed in Section 1.2 and further expanded in Section 1.3. The Probability Density Function describing the decay rate is therefore needed to fit to the

[1] Monte-Carlo datasets generated using only the PDF as a model rather than performing the full LHCb simulation are commonly referred to as "toy" datasets. Repeated generation and fitting of such a dataset is known as a "toy study".

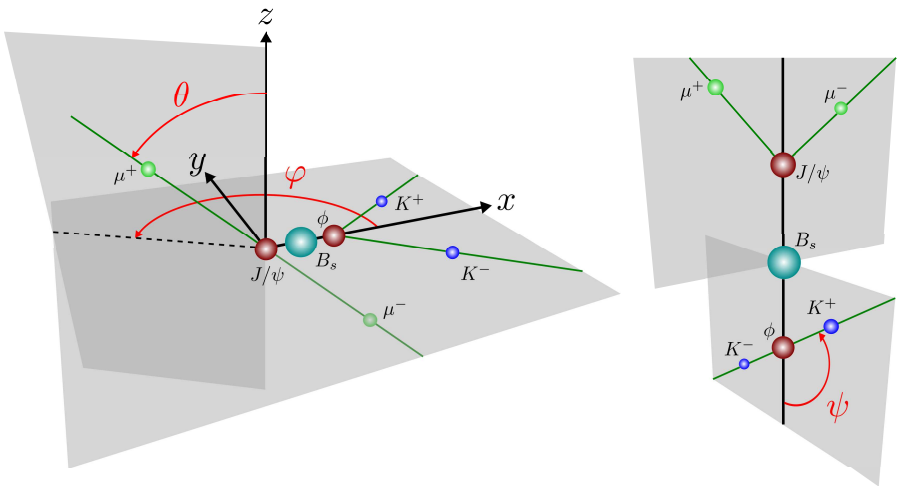


Figure 4.2.: Definition of the transversity angles θ, φ, ψ in the decay $B_s^0 \to J/\psi\,\phi$. φ is the azimuthal angle of the positive muon in the centre-of-mass frame of the J/ψ, where the $x-y$ plane is defined by the direction of the ϕ and the plane in which it decays. θ is the polar angle of the positive muon, and ψ is the angle between the positive kaon and the B_s^0 direction in the ϕ centre-of-mass frame.

k	$h_k(t)$	$f_k(\Omega = \cos\theta, \cos\psi, \varphi)$	$\int f_k \, d\Omega$
1	$\lvert A_0(t)\rvert^2$	$2\cos^2\psi\left(1-\sin^2\theta\cos^2\phi\right)$	$\frac{32}{9}\pi$
2	$\lvert A_\parallel(t)\rvert^2$	$\sin^2\psi\left(1-\sin^2\theta\sin^2\phi\right)$	$\frac{32}{9}\pi$
3	$\lvert A_\perp(t)\rvert^2$	$\sin^2\psi\sin^2\theta$	$\frac{32}{9}\pi$
4	$\Im\mathfrak{m}(A_\parallel(t)\,A_\perp(t))$	$-\sin^2\psi\sin 2\theta\sin\phi$	0
5	$\Re\mathfrak{e}\,(A_0(t)\,A_\parallel(t))$	$\frac{1}{2}\sqrt{2}\,\sin 2\psi\sin^2\theta\sin 2\phi$	0
6	$\Im\mathfrak{m}(A_0(t)\,A_\perp(t))$	$\frac{1}{2}\sqrt{2}\,\sin 2\psi\sin 2\theta\cos\phi$	0

Table 4.8.: Angular amplitude components in the transversity basis for $B_s^0 \to J/\psi\,\phi$ signal terms.

data. As will be seen in the following sections this is dependent upon more than just the proper time.

4.3.1. Differential decay rates

The decay rates for B_s^0 and $\overline{B}_s^0$ to $J/\psi\ \phi$ as a function of proper time t can be written as:

$$\Gamma\left(B_s^0(t) \to J/\psi\phi\right) = \frac{|A_f|^2}{2} e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right)\right. \tag{4.18}$$

$$\left. - \Re\mathfrak{e}\,\lambda_f \sinh\left(\frac{\Delta\Gamma t}{2}\right) - \Im\mathfrak{m}\lambda_f \sin\left(\Delta m t\right)\right] \tag{4.19}$$

$$\Gamma\left(\overline{B}_s^0(t) \to J/\psi\phi\right) = \frac{|\bar{A}_f|^2}{2} e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right)\right. \tag{4.20}$$

$$\left. - \Re\mathfrak{e}\,\lambda_f \sinh\left(\frac{\Delta\Gamma t}{2}\right) + \Im\mathfrak{m}\lambda_f \sin\left(\Delta m t\right)\right] \tag{4.21}$$

Where $A_f, \bar{A}_f$ are the instantaneous decay amplitudes defined in Equations 1.49 and 1.50, and λ_f is the complex quantity defined in equation (1.51). Taking the real and imaginary parts of equation (1.56) allows us to express these in terms of the observable phase difference:

$$\Im\mathfrak{m}\lambda_f = -\eta_f \sin\phi_{\lambda_f} \qquad\qquad \Re\mathfrak{e}\,\lambda_f = \eta_f \cos\phi_{\lambda_f} \tag{4.22}$$

We choose to call the observable phase difference ϕ_s, which may be measured to be either $\phi_{\lambda_f} = -2\beta_s$ as in equation (1.58) or in the presence of new physics ϕ'_{λ_f} in equation (1.61).

$$\Gamma\left(B_s^0(t) \to J/\psi\phi\right) = \frac{|A_f|^2}{2} e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right)\right. \tag{4.23}$$

$$\left. -\eta_f \cos\phi_s \sinh\left(\frac{\Delta\Gamma t}{2}\right) + \eta_f \sin\phi_s \sin\left(\Delta m t\right)\right] \tag{4.24}$$

$$\Gamma\left(\overline{B}_s^0(t) \to J/\psi\phi\right) = \frac{|\bar{A}_f|^2}{2} e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right)\right. \tag{4.25}$$

$$\left. -\eta_f \cos\phi_s \sinh\left(\frac{\Delta\Gamma t}{2}\right) - \eta_f \sin\phi_s \sin\left(\Delta m t\right)\right] \tag{4.26}$$

If the final state was composed of a single $\mathcal{CP}$ eigenstate fitting to the above proper time distributions would be sufficient to determine ϕ_s. $B_s^0 \to J/\psi\,\phi$ is a pseudo-scalar to two vector decay however: Angular momentum conservation allows three possible relative orbital momenta $\ell = 0, 1, 2$ for the vector mesons. The $\mathcal{CP}$ eigenvalue of the final state

is given by:

$$\begin{aligned} \mathcal{CP}|J/\psi\phi\rangle_\ell &= \eta_f |J/\psi\phi\rangle_\ell \\ &= (-1)^\ell |J/\psi\phi\rangle_\ell \end{aligned} \tag{4.27}$$

These $\mathcal{CP}$ eigenstates contribute different polarisations to the final state, with time-dependent complex amplitudes $A_0(t)$, $A_\parallel(t)$ ($\ell = 0, 2$) and $A_\perp(t)$ ($\ell = 1$) and corresponding phases δ_0, $\delta_\parallel$, $\delta_\perp$. The amplitudes are defined at $t = 0$ using equation (1.36) and equation (1.37) as:

$$A_0(0) = \langle J/\psi\phi|_{\ell=0}\mathcal{H}|B_s^0\rangle \quad A_\perp(0) = \langle J/\psi\phi|_{\ell=1}\mathcal{H}|B_s^0\rangle \quad A_\parallel(0) = \langle J/\psi\phi|_{\ell=2}\mathcal{H}|B_s^0\rangle \tag{4.28}$$

The normalised sum of these amplitudes being unity means that only two amplitudes are independent. For the phases where only a phase difference may be measured, it is customary to define $\delta_0 = 0$. The $\mathcal{CP}$-even components contribute to the $\mathcal{CP}$ asymmetry with opposite sign to the $\mathcal{CP}$ odd component, diluting the measurable $\mathcal{CP}$ asymmetry; in order to extract the undiluted $\mathcal{CP}$ asymmetry the magnitude and phase of the three components must be measured. The angular distribution of the final state particles can be used to extract statistically these amplitudes and phases. With four particles in the final state three angles are required to fully specify the coordinate system. Common choices for these angles in pseudo-scalar to two vector decays are the helicity and transversity bases, the latter of which has been chosen for this analysis. The transversity basis is illustrated in Figure 4.2 where the angles θ, φ and ψ are defined. The differential decay rate in the transversity basis for B_s^0 mesons is of the form

$$P(t, \Omega; \vec{\alpha}) = \frac{d^4\Gamma}{dt\ d\cos\theta\ d\varphi\ d\cos\psi} = \frac{d^4\Gamma}{dt\ d\Omega} = \frac{\sum_{k=1}^{6} h_k(t) f_k(\Omega)}{\int\int \sum_{j=1}^{6} h_j(t) f_j(\Omega)\ dt\ d\Omega} \tag{4.29}$$

with $\Omega = (\cos\theta, \varphi, \cos\psi)$. The time-dependent angular functions are $h_k(t)$ and $f_k(\Omega)$ are their corresponding angular functions both of which are listed in Table 4.8. Assuming a signal-only $B_s^0 \to J/\psi\,\phi$ sample fully reconstructed by a perfect detector equation (4.29) would be sufficient to extract parameters of interest. The time-dependent amplitudes

k	$h_k(t)$	$f_k(\Omega = \cos\theta, \cos\psi, \varphi)$	$\int f_k\, d\Omega$
7	$\lvert A_s(t)\rvert^2$	$\frac{2}{3}(1-\sin^2\theta\cos^2\phi)$	$\frac{32}{9}\pi$
8	$\Re\mathfrak{e}\,(A_s(t)A_\parallel(t))$	$\frac{1}{3}\sqrt{6}\sin\psi\sin^2\theta\sin 2\phi$	0
9	$\Im\mathfrak{m}(A_s(t)A_\perp(t))$	$\frac{1}{3}\sqrt{6}\sin\psi\sin 2\theta\cos\phi$	0
10	$\Re\mathfrak{e}\,(A_s(t)A_0(t))$	$\frac{4}{3}\sqrt{3}\cos\psi(1-\sin^2\theta\cos^2\phi)$	0

Table 4.9.: Additional angular amplitude components in the transversity basis for s-wave interference terms arising from nonresonant $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi f_0(980)$ contributions.

for $B_s^0 \to J/\psi\,\phi$ are:

$$|A_0(t)|^2 = |A_0|^2 e^{-\Gamma_s t}\left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right) + \sin\phi_s \sin(\Delta m_s t)\right] \tag{4.30}$$

$$|A_\parallel(t)|^2 = |A_\parallel|^2 e^{-\Gamma_s t}\left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right) + \sin\phi_s \sin(\Delta m_s t)\right] \tag{4.31}$$

$$|A_\perp(t)|^2 = |A_\perp|^2 e^{-\Gamma_s t}\left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right) - \sin\phi_s \sin(\Delta m_s t)\right] \tag{4.32}$$

$$\begin{aligned}\Im\mathfrak{m}(A_\parallel(t)\,A_\perp(t)) = |A_\parallel||A_\perp| e^{-\Gamma_s t}\Bigg[&-\cos(\delta_\perp - \delta_\parallel)\sin\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\\ &- \cos(\delta_\perp - \delta_\parallel)\cos\phi_s \sin(\Delta m_s t) + \sin(\delta_\perp - \delta_\parallel)\cos(\Delta m_s t)\Big]\end{aligned} \tag{4.33}$$

$$\begin{aligned}\Re\mathfrak{e}\,(A_0(t)\,A_\parallel(t)) = |A_0||A_\parallel| e^{-\Gamma_s t}\cos(\delta_\parallel - \delta_0)\Bigg[&\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\\ &+ \sin\phi_s \sin(\Delta m_s t)\Big]\end{aligned} \tag{4.34}$$

$$\begin{aligned}\Im\mathfrak{m}(A_0(t)\,A_\perp(t)) = |A_0||A_\perp| e^{-\Gamma_s t}\Bigg[&-\cos(\delta_\perp - \delta_0)\sin\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\\ &- \cos(\delta_\perp - \delta_0)\cos\phi_s \sin(\Delta m_s t) + \sin(\delta_\perp - \delta_0)\cos(\Delta m_s t)\Big]\end{aligned} \tag{4.35}$$

And the physics parameters $\vec{\alpha}$ consist of the average decay width Γ_s, the decay width difference $\Delta\Gamma_s$, the amplitudes and phases $|A_0|^2, \delta_0,\ |A_\parallel|^2, \delta_\parallel,\ |A_\perp|^2, \delta_\perp$, the B_s^0 mixing frequency Δm_s and the $\mathcal{CP}$ violating phase ϕ_s. For the charge-conjugate decay all mixing terms ($\sin(\Delta m_s t)$ and $\cos(\Delta m_s t)$) change sign in the above equations to provide the necessary rates.

4.3.2. Inclusion of an s-wave component

The $\phi \to K^+K^-$ resonant decay is an $l = 1$ (p-wave) final state. Contributions from $l = 0$ (s-wave) final states such as $B_s^0 \to J/\psi f_0(980)$ and nonresonant $B_s^0 \to J/\psi K^+K^-$ channels are possible around the ϕ invariant mass. These interfere as they cannot be removed by

the application of cuts to the data. As such it is easier, and even desirable, to incorporate an s-wave component to disentangle its effects from that of pure $B_s^0 \to J/\psi\,\phi$ [89]. The addition of an s-wave component introduces a new amplitude $A_s(t)$ and phase δ_s which interferes with the previously defined amplitudes, expanding equation (4.29) to include an additional four terms listed in Table 4.9 with time-dependent amplitude functions of the form:

$$|A_s(t)|^2 = |A_s|^2 e^{-\Gamma_s t}\left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right) - \sin\phi_s \sin(\Delta m_s t)\right] \tag{4.36}$$

$$\begin{aligned}\Re\mathfrak{e}\left(A_s(t)A_\parallel(t)\right) = |A_s||A_\parallel| e^{-\Gamma_s t}\Bigg[&-\sin(\delta_\parallel - \delta_s)\sin\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\\ &-\sin(\delta_\parallel - \delta_s)\ \cos\phi_s \sin(\Delta m_s t) + \cos(\delta_\parallel - \delta_s)\cos(\Delta m_s t)\Big]\end{aligned} \tag{4.37}$$

$$\begin{aligned}\Im\mathfrak{m}(A_s(t)A_\perp(t)) = |A_s||A_\perp| e^{-\Gamma_s t}\sin(\delta_\perp - \delta_s)\Bigg[&\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\\ &-\sin\phi_s \sin(\Delta m_s t)\Big]\end{aligned} \tag{4.38}$$

$$\begin{aligned}\Re\mathfrak{e}\left(A_s(t)A_0(t)\right) = |A_s||A_0| e^{-\Gamma_s t}\Bigg[&-\sin(\delta_0 - \delta_s)\sin\phi_s \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)\\ &-\sin(\delta_0 - \delta_s)\cos\phi_s \sin(\Delta m_s t) + \cos(\delta_0 - \delta_s)\cos(\Delta m_s t)\Big]\end{aligned} \tag{4.39}$$

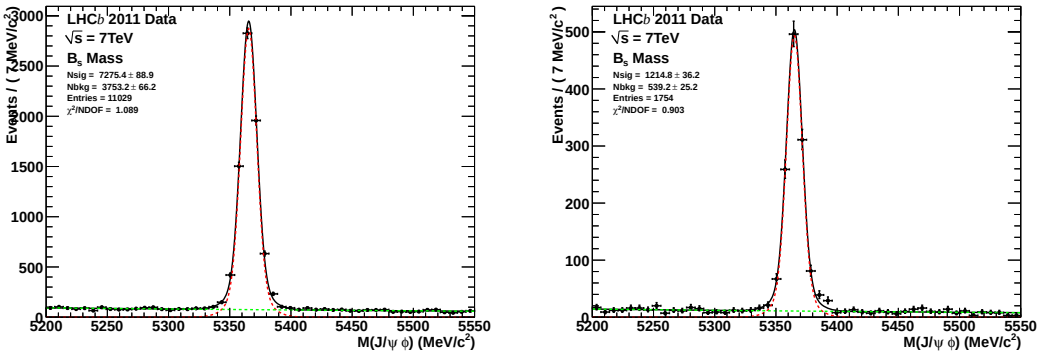
The normalised PDF including the additional s-wave terms is then:

$$P(t, \Omega; \vec{\alpha}) = \frac{\sum_{k=1}^{10} h_k(t) f_k(\Omega)}{\int\int \sum_{j=1}^{10} h_j(t) f_j(\Omega)\ dt\ d\Omega} \tag{4.40}$$

Where $\vec{\alpha}$ now includes as measureable parameters $|A_s|^2$, δ_s. Previous experiments have studied the s-wave contribution [90] [91], from which we expect an s-wave component of $< 6.7\%$ at 95% C.L. Section 6.1 describes the fit validation procedure and associated studies to determine if an s-wave component below this limit can be measured at LHCb.

4.4. The $\mathcal{S}$-fit technique

Fitting to the dataset requires not only evaluation of the signal component, but also some method of handling the background. Typically this would involve construction of a PDF that describes the signal and background in $M(J/\psi\phi), t, \Omega$ and performing a fit of the kind described by equation (4.6). This relies upon knowledge of the shape of the background distribution in five observables many of which may be correlated preventing the construction of a factorised PDF of the form in equation (4.5). The complexity of such a fit increases the possibility of mis-modelling a specific background distribution as well as the computing load. One possible solution to this is an appeal to



(a) proper time unbiased **(b)** proper time biased

Figure 4.3.: $M(J/\psi\phi)$ spectra of proper time biased and proper time unbiased triggered candidates. The overlaid fit is used to determine $\mathcal{S}$-weights for the $\mathcal{S}$-fit, and as the mass model in the fit to signal + background. The signal component is modelled using a double Gaussian while background distribution is modelled by a shallow exponential. The relative ratio of fraction of the two Gaussians and the ratio of their widths are fixed to Monte-Carlo, with the mean and narrower width floated.

the $\mathcal{S}$-plot technique. Fitting to one distribution that offers good discrimination between signal and background results in signal $\mathcal{S}$-weights that when applied to the data cause statistical cancellation of the background component, and thus reconstruction of the signal distribution in other observables. The $\mathcal{S}$-weights can be applied on a per-event basis in the likelihood function so that only the signal PDF is necessary [92]. The likelihood is now:

$$\mathcal{L}(\vec{\alpha}) = \prod_{i}^{N} \left[P(\vec{x}_i;\vec{\alpha})\right]^{w_{\rm sig}(y_i)} \tag{4.41}$$

where $w(y_i)$ is the $\mathcal{S}$-weight applied to the i^{th} event, determined from the observable y which is not correlated to any of the observables in $\vec{x}$. The minimisation then becomes:

$$-\frac{d}{d\vec{\alpha}} \ln \mathcal{L}(\vec{\alpha}) = -\frac{d}{d\vec{\alpha}} \sum_{i}^{N} \left[P(\vec{x}_i;\vec{\alpha})\right] \times w_{\rm sig}(y_i) = 0 \tag{4.42}$$

This is referred to as the $\mathcal{S}$-fit procedure and is a logical extension of the $\mathcal{S}$-plot technique.

Parameter	value
f_1	0.830
r_1	2.14
μ	5365.435 ± 0.082 MeV/c^2
σ_1	6.380 ± 0.069 MeV/c^2
c	$-1.282 \pm 0.14 \times 10^{-3}$ (MeV/c^2)$^{-1}$

Table 4.10.: Parameters used in the fit to $M(J/\psi\phi)$ both in the full signal + background fit and in the fit to extract signal $\mathcal{S}$-weights. Parameters without uncertainties are determined from Monte-Carlo and fixed to these values in the fit.

4.4.1. Fit to extract $\mathcal{S}$-weights

In the $B_s^0 \to J/\psi\,\phi$ analysis the logical choice for a discriminating observable is the $M(J/\psi\phi)$ invariant mass distribution. The mass range of $5200 - 5500$ MeV/c^2 is used to determine the PDF which is of the form:

$$P(M(J/\psi\phi);\vec{\alpha}_{sig}) = \frac{f_1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(M(J/\psi\phi)-\mu)^2}{2\sigma_1}\right] + \frac{1-f_1}{\sqrt{2\pi}r_1\sigma_1}\exp\left[-\frac{(M(J/\psi\phi)-\mu)^2}{2r_1\sigma_1}\right] \tag{4.43}$$

Where the parameters f_1, r_1 define the ratio of the first Gaussian to the second and the ratio of widths respectively. These parameters are found to be the same in data and Monte-Carlo to within statistical uncertainties, and are fixed to those listed in Table 4.10. The background distribution is modelled as a shallow exponential with a single coefficient, c. The $M(J/\psi\phi)$ invariant mass spectra of proper time unbiased and proper time biased triggered datasets are shown in Figure 4.3. In total 7275 ± 89 proper time unbiased signal candidates and 1215 ± 36 proper time biased signal candidates are observed. The simultaneous fit to both species yields the values listed in Table 4.10. Initial studies considered inclusion of the $M(\mu^+\mu^-)$ distribution to permit extraction of specific backgrounds such as those which peak in the B_s^0 mass but do not include a J/ψ candidate. The purity of the selected candidates is such that this yield was found to be statistically insignificant, and so a simpler PDF using only the $M(J/\psi\phi)$ distribution is used.

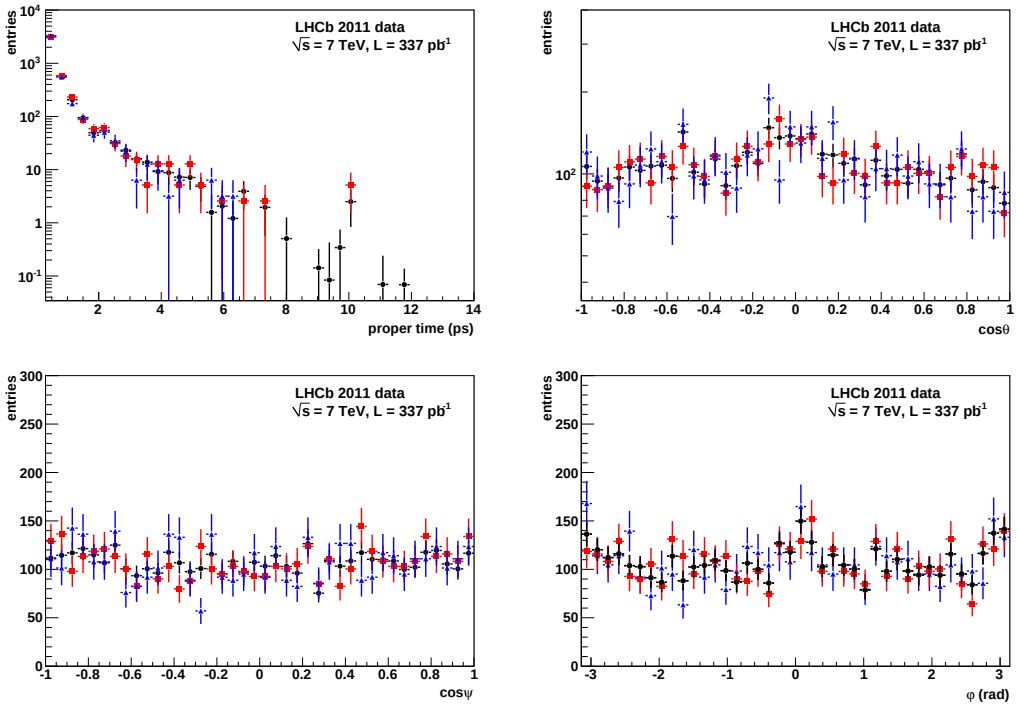


Figure 4.4.: Comparison showing background $\mathcal{S}$-plots over the full B_s^0 mass range (black circles), the $5200 < M(J/\psi\phi) < 5320\text{MeV}/c^2$ sideband (red squares) and $5420 < M(J/\psi\phi) < 5550\text{MeV}/c^2$ sideband (blue triangles) in the observables $t, \cos\theta, \varphi, \cos\psi$ used in the fit normalised to the same area. The $\mathcal{S}$-weighted distributions obtained from the fit in Figure 4.3 are in good agreement with the sidebands.

4.4.2. Verification of the $\mathcal{S}$-fit procedure

The $\mathcal{S}$-fit procedure effectively removes the need to model background proper time and angular distributions in the fit. While this simplification is welcomed, it is important to ensure that the technique correctly reproduces the signal distribution in these observables. The signal distributions are not directly accessible in data by other means, but the background distributions are through the B_s^0 mass sidebands. Figure 4.4 shows these sidebands compared with the background $\mathcal{S}$-plots derived from the fit in Figure 4.3, where it can be seen that the distributions are in excellent agreement. While such a check is reassuring, it is not stringent. For the LHCb ϕ_s measurement three distinct fitting packages are used allowing independent cross-checking of several techniques. The analysis described in this thesis uses the $\mathcal{S}$-fit technique, making it unique: The other two packages use full fits to signal and background in $M(J/\psi\phi), t$ and Ω. The background

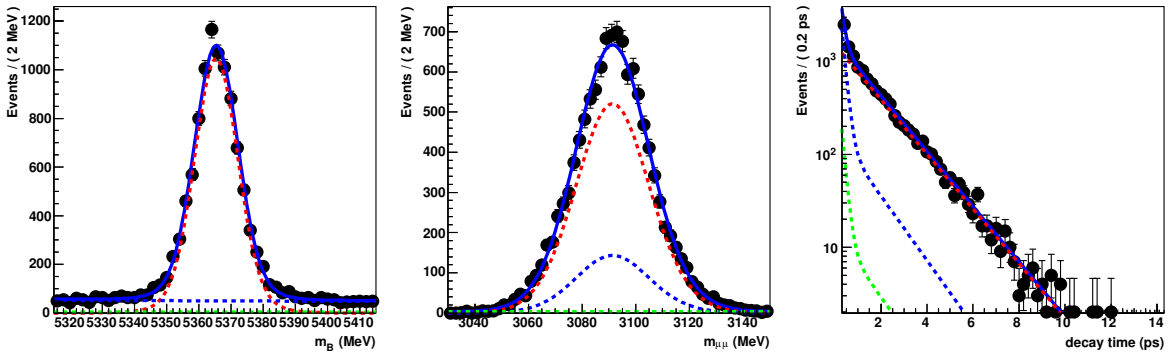


Figure 4.5.: $M(\mu^+\mu^-)$, $M(J/\psi\phi)$ and proper time distributions of the proper time unbiased dataset. A simultaneous fit to the distributions is overlaid, used to determine the background components of the background PDF. Signal $B_s^0 \to J/\psi K^+K^-$ are shown in red, peaking in both $M(\mu^+\mu^-)$ and $M(J/\psi\phi)$. Prompt J/ψ and non-signal $B \to J/\psi X$ are shown in blue making up approximately 32% of candidates. Fully combinatoric background is shown in green and is present at the level of about 2%. Both background types use the same double exponential decay model.

models used in these packages are discussed in the next section, and a direct comparison between the fit results of the three fitters can be found in Chapter 6.

4.5. Background modelling

For the full fits to signal and background, referred to as the Heidelberg and Edinburgh $\mathcal{C}$-fit analyses to distinguish them from the Edinburgh $\mathcal{S}$-fit analysis presented here, the background distributions in $M(J/\psi\phi), t$ and Ω must be modelled and a PDF developed. The $M(J/\psi\phi)$ component of this PDF has already been described in Section 4.4.1.

4.5.1. Sources of background

Several distinct types of background need to be quantified:

- Backgrounds from misreconstructed true $B_s^0 \to J/\psi\,\phi$ candidates.
- Peaking Backgrounds from other B decays.
- Combinatorial backgrounds:
  - With and without a true J/ψ.

The $M(\mu^+\mu^-)$ distribution in data is a good place to start investigating the breakdown of these sources. For peaking J/ψ candidates a bifurcated Gaussian distribution is used, with a simple shallow exponential to describe the nonresonant $\mu^-\mu^+$ background. The $M(\mu^+\mu^-)$ distribution is very pure, with only $1.7 \pm 0.3\%$ of candidates passing the selection not peaking in $M(\mu^+\mu^-)$ as shown in Figure 4.5. There are no candidates that peak in $M(J/\psi\phi)$ and do not peak in $M(\mu^+\mu^-)$, which is why the $\mathcal{S}$-fit only requires the B_s^0 mass fit to produce $\mathcal{S}$-weights: Candidates peaking in B_s^0 in the data are exclusively $B_s^0 \to J/\psi X$, where the nonresonant K^+K^- and pollution from $f_0(980)$ are already accounted for in the s-wave component of the fit.

4.5.2. Background proper time model

For candidates that peak in $M(\mu^+\mu^-)$ but not $M(J/\psi\phi)$ there are two possible sources: Prompt J/ψ and those coming from displaced vertices as daughters of non-signal $B \to J/\psi X$ decays. The prompt J/ψ component is almost entirely removed by the proper time cut at > 0.3ps, leaving only the small fraction of long-lived combinatoric $\mu^+\mu^-$ and J/ψ components requiring modelling. The background parameterisation of the proper time for these candidates is empirically determined from examining the proper time distribution of the $M(J/\psi\phi)$ sidebands and the proper time $\mathcal{S}$-plot in Figure 4.4, and is of the form:

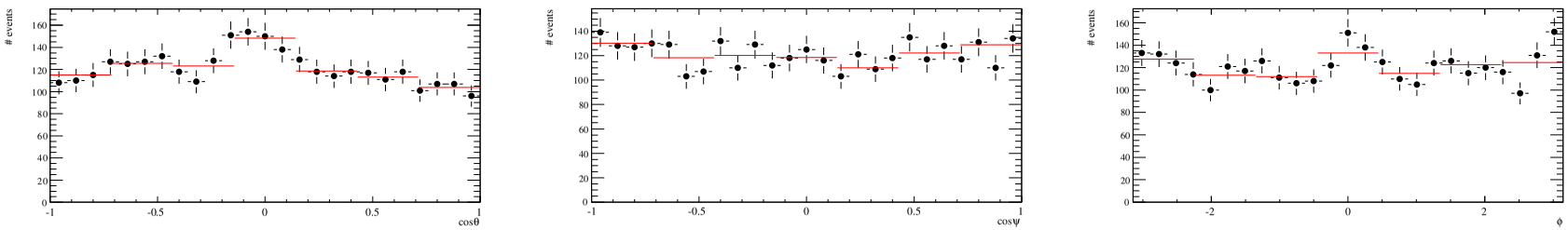
$$P(t, \vec{\alpha}_{\text{bkg}}) = f^1_{\text{bkg}} \tau^1_{\text{bkg}} \exp\left(-t/\tau^1_{\text{bkg}}\right) + (1 - f^1_{\text{bkg}}) \tau^2_{\text{bkg}} \exp(-t/\tau^2_{\text{bkg}}) \tag{4.44}$$

where f^1_{bkg} denotes the fraction of background with effective proper time τ^1_{bkg}, the remainder having an effective proper time τ^2_{bkg}. For the proper time biased dataset a single exponential decay with effective proper time τ^B_{bkg} is sufficient:

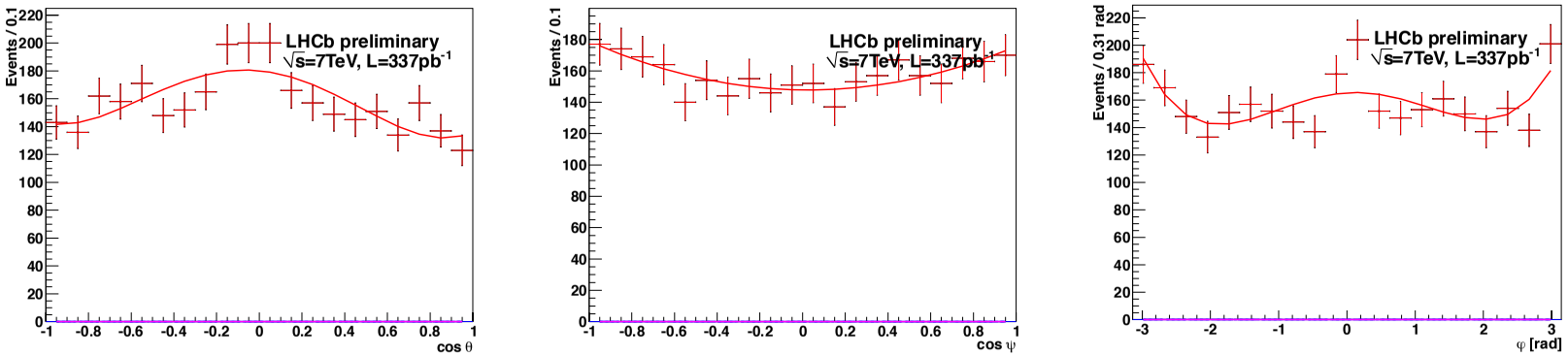
$$P(t, \vec{\alpha}_{\text{bkg}}) = \tau^B_{\text{bkg}} \exp(-t/\tau^B_{\text{bkg}}) \tag{4.45}$$

4.5.3. Background angular model

The Edinburgh $\mathcal{C}$-fit analysis uses a normalised 3D histogram in Ω developed from the mass sidebands to model the angular distribution in the fit. The binning scheme is $7 \times 7 \times 7$ projections of which are shown in Figure 4.6a, with other choices of binning used to estimate systematic uncertainties. The Heidelberg $\mathcal{C}$-fit parameterizes the background by fitting the same sideband distributions in a $10 \times 10 \times 10$ binning scheme to Legendre



(a) Edinburgh $\mathcal{C}$-fit histogram projections



(b) Heidelberg $\mathcal{C}$-fit Legendre polynomial projections

Figure 4.6.: Sideband distributions of $\cos\theta, \cos\psi$ and φ showing projections of the 3D angular background histogram (top) used by the Edinburgh $\mathcal{C}$-fit and projections of the Legendre polynomial used by the Heidelberg $\mathcal{C}$-fit (bottom).

polynomials shown in Figure 4.6b. The polynomials are of the form:

$$P(\Omega, \vec{\alpha}_{\text{bkg}}) = \sum_{i,j,k} \alpha_{ijk} P_\ell(\cos\theta, i) P_\ell(\varphi, j) P_\ell(\cos\psi, k) \tag{4.46}$$

Chapter 5.

Auxiliary studies for the ϕ_s measurement

"... we are in a jungle and find our way by trial and error, building our road behind us as we proceed. "

— Max Born

Chapter 4 has described the likelihood fit necessary to make a measurement of ϕ_s and $\Delta\Gamma_s$ in the ideal case of a perfect detector capable of reconstructing, triggering and selecting $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ candidates with 100% efficiency and infinite resolution. While LHCb has been designed to be as close to ideal as can be, it is still subject to the same technological, budgetary and practical considerations as any other experiment, meaning that inefficiencies, finite resolutions and imperfect understanding of nuisance parameters must be accounted for. With the 2011 dataset the ϕ_s analysis depends upon inputs from a number of auxiliary analyses to model and correct for these imperfections. This chapter describes in detail the modifications and inputs required for the likelihood fit to account for the proper time and angular acceptances, the proper time resolution, the imperfect knowledge of the $\mathrm{B_s^0}$ flavour at production and the $\mathrm{B_s^0}$ mixing frequency in order to maximise sensitivity and minimise bias on the measurement of physics parameters in the fit.

5.1. Proper time and angular acceptance

The proper time and angular observables are subject to inefficiencies related to the physical acceptance of the detector as well as trigger and selection biases. Typically these inefficiencies or *acceptance effects* are determined by fitting to binned efficiency histograms in Monte-Carlo, where the generated and accepted distributions are known. In the four dimensional case as would be required here, the number of events needed to ensure bins are sufficiently populated would take a prohibitively long time to generate. There is also a large computational penalty in using such an acceptance term in the PDF as it must be normalised, requiring numeric integration unless an analytical form of the acceptance integral can be determined. An elegant solution to this problem arises upon closer inspection of the PDF expressed in equation (4.40). Here the signal PDF is specified by the observables $\vec{x} = t, \cos\theta, \psi, \varphi = t, \Omega$ and the physics parameters, $\vec{\alpha}$. If an acceptance $\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon)$ is included the PDF originally introduced with explicit normalisation in equation (4.2) has the form:

$$P(\vec{x}; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon) = \frac{p(\vec{x}; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon)}{\int p(\vec{x}; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon) d\vec{x}} \tag{5.1}$$

If the physics parameters $\vec{\alpha}$ are independent of the acceptance parameterisation $\vec{\alpha_\epsilon}$, the minimisation of the negative log-likelihood becomes:

$$\begin{aligned}
&\frac{d}{d\vec{\alpha}} \sum_i^N - \ln\left[\frac{p(\vec{x}_i; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}_i; \vec{\alpha}_\epsilon)}{\int p(\vec{x}; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon) d\vec{x}}\right] = \\
&\frac{d}{d\vec{\alpha}} \sum_i^N - \left\{\ln\left[\frac{p(\vec{x}_i; \vec{\alpha})}{\int p(\vec{x}; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon) d\vec{x}}\right] + \ln\left[\varepsilon_{\text{acc}}(\vec{x}_i; \vec{\alpha}_\epsilon)\right]\right\} \Rightarrow \\
&\frac{d}{d\vec{\alpha}} \sum_i^N - \ln\left[\frac{p(\vec{x}_i; \vec{\alpha})}{\int p(\vec{x}; \vec{\alpha})\varepsilon_{\text{acc}}(\vec{x}; \vec{\alpha}_\epsilon) d\vec{x}}\right] = 0
\end{aligned} \tag{5.2}$$

where the acceptance term drops out of the log-likelihood numerator and only remains in the normalisation integral in the denominator. The denominator in equation (4.40) now reads:

$$\int\int \sum_{j=1}^{10} h_j(t) f_j(\Omega) \varepsilon_{\text{acc}}(t, \Omega)\ dt\ d\Omega \tag{5.3}$$

The acceptance and angular functions do not depend on any of the physics parameters, so they can be factorised out into an integral that only needs to be calculated once at the

start of the likelihood minimisation. We are left with a set of ten acceptance functions, one for each of the 10 time and angular terms in Tables 4.8 and 4.9:

$$\xi_j(t) = \int f_j(\Omega)\varepsilon_{\text{acc}}(t,\Omega)\ d\Omega \tag{5.4}$$

If the proper time acceptance and angular acceptance factorise such that$\varepsilon_{\text{acc}}(t,\Omega) = \varepsilon_{\text{acc}}(t) \times \varepsilon_{\text{acc}}(\Omega)$, A powerful result becomes apparent: *the functional form of the angular acceptance is unnecessary.* The angular acceptance terms reduce to numeric weights ξ_j which can be computed once [93]. In the absence of any acceptance it can be seen that $\xi_j = \int f_j(\Omega)\ d\Omega$ is just the integral over all angles of the f_j terms in Table 4.8 and Table 4.9. The likelihood minimisation for signal including proper time and angular acceptance now becomes:

$$-\frac{d}{d\vec{\alpha}} \ln \mathcal{L} = -\frac{d}{d\vec{\alpha}} \sum_i^N \ln \frac{\sum_k^{10} h_k(t_i;\vec{\alpha}) f_k(\Omega_i)}{\int \sum_{j=1}^{10} h_j(t)\varepsilon_{\text{acc}}(t)\xi_j\ dt} \tag{5.5}$$

The acceptance now simplifies to a parameterised proper time acceptance and a set of 10 angular acceptance weights.

5.1.1. Determination of the angular acceptance weights

The four-dimensional efficiency in equation (5.4) describes the probability that an event generated with a given t_i and Ω_i will be available after reconstruction, triggering, stripping and selection. This efficiency depends on more than just the proper time and transversity angles: There will be a dependence on the momenta of the final state particles, track multiplicity, impact parameter, etc. If we include these additional dependencies as the set of parameters $\vec{z}$ the efficiency in only t, Ω arises from integrating out the $\vec{z}$ dependence:

$$\varepsilon_{\text{acc}}(t,\Omega) = \frac{\int \varepsilon_{\text{acc}}(t,\Omega,\vec{z}) P(t,\Omega,\vec{z};\vec{\alpha})\ d\vec{z}}{P(t,\Omega;\vec{\alpha})} \tag{5.6}$$

The PDF describing generated events prior to reconstruction, triggering, etc, is independent of $\vec{z}$: $\int P(t,\Omega,\vec{z};\vec{\alpha})\ dz = P(t,\Omega;\vec{\alpha})$, but it is dependent upon t, Ω. Substituting

equation (5.6) into equation (5.4) we have:

$$\xi_j(t) = \int \frac{f_j(\Omega)}{P(t,\Omega;\vec{\alpha})}\varepsilon_{\rm acc}(t,\Omega,\vec{z})P(t,\Omega,\vec{z};\vec{\alpha})\ d\vec{z}\ d\Omega \tag{5.7}$$

$$= \int \frac{f_j(\Omega)}{P(\Omega;t,\vec{\alpha})P(t;\vec{\alpha})}\varepsilon_{\rm acc}(t,\Omega,\vec{z})P(\Omega,\vec{z};t,\vec{\alpha})P(t;\vec{\alpha})\ d\vec{z}\ d\Omega \tag{5.8}$$

$$= \int \frac{f_j(\Omega)}{P(\Omega;t,\vec{\alpha})}\varepsilon_{\rm acc}(t,\Omega,\vec{z})P(\Omega,\vec{z};t,\vec{\alpha})\ d\vec{z}\ d\Omega \tag{5.9}$$

The term $P(t,\Omega,\vec{z};\vec{\alpha})\ dzd\Omega$ is the probability that an event is generated in the infinitesimal phase space at $\vec{z},\Omega$. Discretising this it can be seen that:

$$\xi_j = \frac{1}{N_{\rm gen}}\sum_i^{N_{\rm gen}} \frac{f_j(\Omega_i)}{P(\Omega_i;t_i,\vec{\alpha})}\varepsilon_{\rm acc}(t,\Omega,\vec{z}) \tag{5.10}$$

Where the sum is over all generated events, and $\varepsilon_{\rm acc}(t,\Omega,\vec{z})$ acting on all generated events serves to limit the sum to only those events passing the acceptance. The sum over only the accepted events is then:

$$\xi_j = \frac{1}{N_{\rm gen}}\sum_i^{N_{\rm acc}} \frac{f_j(\Omega_i)}{P(\Omega_i;t_i,\vec{\alpha})} \tag{5.11}$$

The factor $1/N_{\rm gen}$ is a constant which can be ignored during the minimisation of the negative log-likelihood. The procedure is then to take a Monte-Carlo sample for which $\vec{\alpha}$ is known, determine the values of $f_j(\Omega_i)$ for each accepted Monte-Carlo event, and divide it by the value of the PDF at $P(\Omega_i;t_i,\vec{\alpha})$, summing over all events to determine ξ_j. This procedure has been verified by splitting the Monte-Carlo sample in two in order to extract the weights from one sample and fit to the physics parameters in the other. A number of additional checks have been performed using an assortment of acceptance parameterisations yielding results consistent within the expected uncertainties of the generated physics parameters [94]. Table 5.1 presents the extracted acceptance weights used in this analysis as derived from fully simulated Monte-Carlo.

j	$\xi_j \times \frac{9}{32\pi}$
1	0.9804 ± 0.002
2	1.0290 ± 0.002
3	1.0282 ± 0.002
4	-0.0025 ± 0.002
5	-0.0006 ± 0.002
6	-0.0002 ± 0.002
7	0.9936 ± 0.002
8	-0.0028 ± 0.002
9	-0.0030 ± 0.002
10	-0.0031 ± 0.002

Table 5.1.: Angular acceptance weights determined from Monte-Carlo.

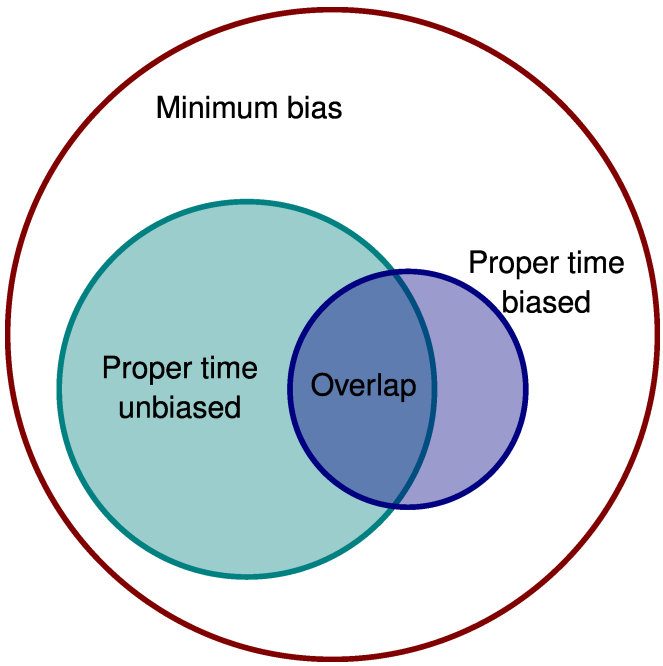
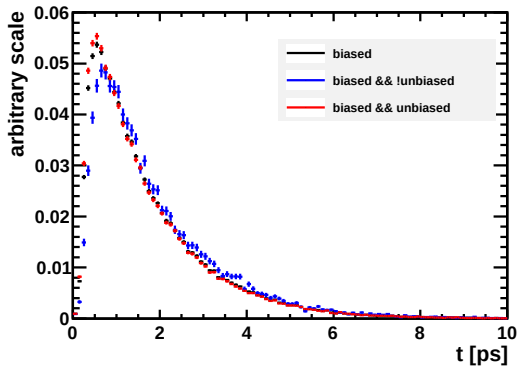
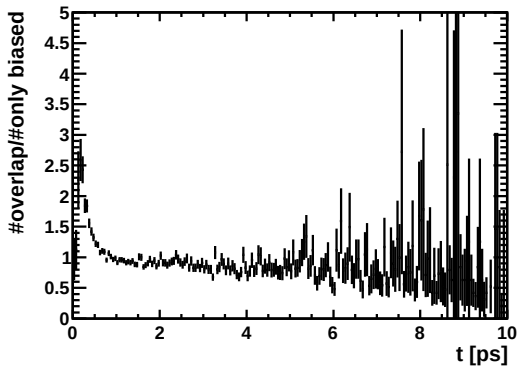


Figure 5.1.: Schematic depiction of the overlap between proper time unbiased and proper time biased triggers. The overlap region can be used to determine the acceptance of one trigger with respect to the other, but it assumes that candidates passing both triggers are representative of those passing only one.

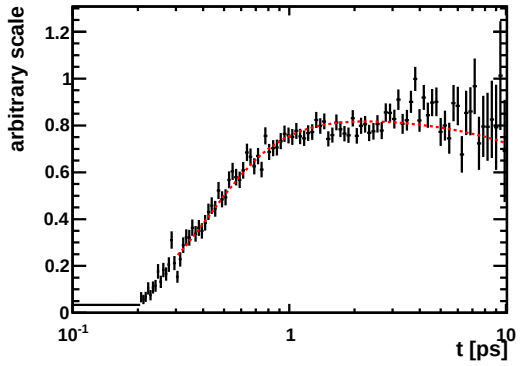


(a) normalised proper time distributions

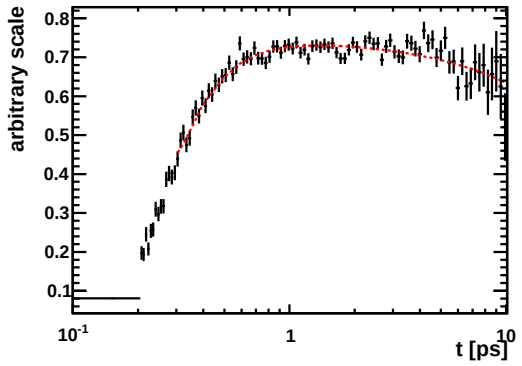


(b) ratio of overlap to proper time biased candidates

Figure 5.2.: (a) Normalised proper time distributions of $B_s^0 \to J/\psi\,\phi$ candidates passing the proper time biased trigger in Monte-Carlo simulation in black, with those passing also the proper time unbiased trigger normalised to the same area in red, and only those passing the proper time biased trigger in blue. (b) The ratio of candidates passing proper time unbiased + biased triggers with respect to only proper time biased candidates as a function of proper time.



(a) proper time biased

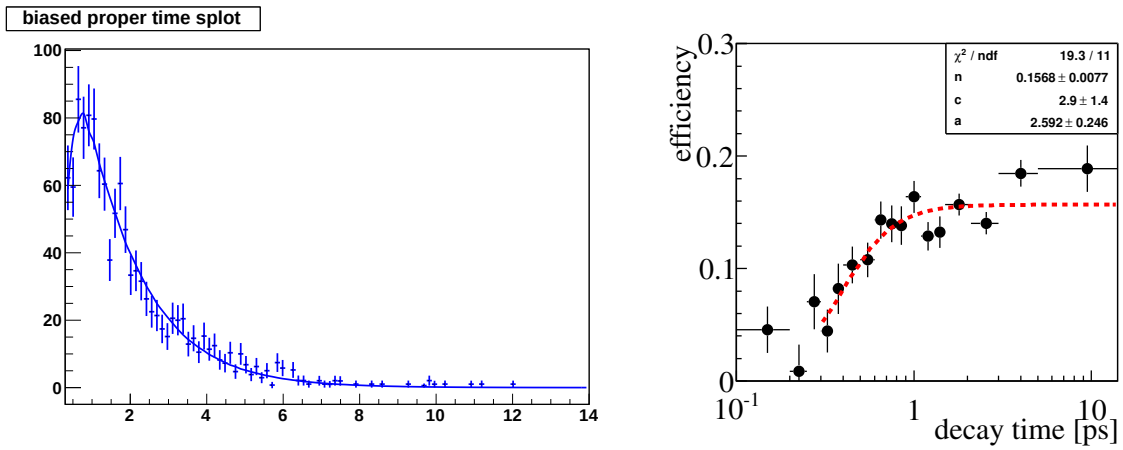


(b) overlap

Figure 5.3.: (a) The ratio of simulated $B_s^0 \to J/\psi\,\phi$ candidates passing only the proper time biased trigger with respect to all candidates in bins of proper time. (b) The same ratio for candidates passing both trigger configurations with respect to all candidates. In both cases a fit to extract the acceptance parameters are overlaid.

parameter	overlap	proper time biased only
c	3.13 ± 0.2	2.54 ± 0.1
a [ps^{-1}]	3.75 ± 0.09	2.33 ± 0.04
β [ps^{-1}]	-0.0157	-0.0157

Table 5.2.: Results of the fits to Monte-Carlo data shown in Figure 5.3. The overlap result is shown to highlight the discrepancy which prevents its use in determining the acceptance in data.



(a) proper time $\mathcal{S}$-plot **(b)** yield ratio

Figure 5.4.: (a) proper time distribution $\mathcal{S}$-plot for $B_s^0 \to J/\psi\,\phi$ signal candidates passing the proper time biased trigger and (b) ratio of extracted yields for exclusively proper time biased candidates over all proper time unbiased candidates determined from fits in bins of proper time.

parameter	Monte-Carlo	data $\mathcal{S}$-plot	overlap
c	2.54 ± 0.1	1.45 ± 0.47	2.9 ± 1.4
a [ps^{-1}]	2.33 ± 0.04	2.37 ± 0.37	2.59 ± 0.25
β [ps^{-1}]	-0.0157	-0.0157	-

Table 5.3.: Summary of acceptance parameter results for the Monte-Carlo simulated data and two data-driven techniques.

5.1.2. Parameterising the proper time acceptance

The VELO pattern recognition algorithm is optimised for tracks pointing back to the beam axis. Longer-lived B_s^0 mesons tend to have vertices further from the beam axis than shorter-lived ones, incurring a loss in reconstruction efficiency as a function of proper time which is of the form:

$$\varepsilon_{\text{reco}}(t) \propto 1 + \beta t \approx e^{\beta t} \tag{5.12}$$

This affects all candidates passing both trigger categories, and results in a bias on the measurement of Γ_s unless corrected for in the fit. The proper time biased triggered candidates are subject to an additional acceptance which must be modelled. Ideally this would be performed with data using a proper time unbiased trigger that passes a superset of the proper time biased trigger selection, but such a trigger is unavailable for the 2011 data-taking period. A possible solution is to look at events that pass both the proper time unbiased and proper time biased trigger lines, using the overlap region described in Figure 5.1 to determine a trigger efficiency. This method assumes that candidates which pass the proper time unbiased trigger *and* the proper time biased trigger are representative of candidates which pass *only* the proper time biased trigger. This assumption is tested on Monte-Carlo and found to be invalid as shown in Figure 5.2, where the ratio of events within the overlap region to those outside of it is not flat, particularly at low lifetimes. As a result the approach chosen to determine the acceptance parameterisation uses a combination of modelling in Monte-Carlo and parameter extraction from data. The functional form is empirical and determined to be:

$$\varepsilon_{\text{biased}}(t) \propto \frac{1}{1 + (at)^{-c}} \tag{5.13}$$

For the proper time biased data the acceptance is the product of $\varepsilon_{\text{biased}}(t)$ and $\varepsilon_{\text{reco}}(t)$. The Monte-Carlo-based parameter extraction uses as input the ratio of reconstructed, triggered events to the full simulated proper time distribution. Figure 5.3 shows these distributions and the fit to the function $\varepsilon_{\text{acc}}(t) = \varepsilon_{\text{reco}}(t) \times \varepsilon_{\text{biased}}(t)$. The fit results are listed in Table 5.2. In data, the acceptance parameters are extracted in two ways:

- The data passing the proper time biased trigger is $\mathcal{S}$-weighted using the B_s^0 mass as the discriminating variable. The full lifetime fit is applied to the signal $\mathcal{S}$-plot with acceptance parameters floated, as shown in Figure 5.4a.

- The proper time biased-only and proper time unbiased mass distributions are fit in bins of proper time to extract their respective yields. The ratio of proper time biased-only to proper time unbiased yield is used to fit the acceptance parameters as shown in Figure 5.4b.

The results of these fits are presented alongside the fit to Monte-Carlo simulated data in Table 5.3 where there is broad agreement between Monte-Carlo data and the two data-driven techniques. The proper time acceptance used in the fit to extract physics parameters uses the $\mathcal{S}$-plot results, and systematic uncertainties are obtained for physics parameters by examining the difference between the results obtained from this and the ratio results. Implementation of the proper time acceptance in the PDF differs slightly from that of the angular weights for cosmetic reasons: The efficiency in proper time is kept in both the numerator and denominator in order to allow projections in t to visibly match the data. This requires the integral in the denominator of equation (5.5) to be evaluable over finite ranges of t:

$$\int \sum_{j=1}^{10} h_j(t)\varepsilon_{\text{acc}}(t)\xi_j \; dt \tag{5.14}$$

While the $h_j(t)$ terms are analytically integrable, $\sum_{j=1}^{10} h_j(t)\varepsilon_{\text{acc}}(t)$ is not, and numerical integration is time-consuming. Instead the functional form of $\varepsilon_{\text{acc}}(t)$ is converted to a histogram in bins of t: $H_{acc}(q)$ with bin height equal to average efficiency on the range $\varepsilon_{\text{acc}}(t+\delta t)$. The denominator is then:

$$\sum_{q}^{t_{max}/\delta t} \int_{q\delta t}^{(q+1)\delta t} \sum_{j=1}^{10} h_j(t) H_{acc}(q)\xi_j \; dt \tag{5.15}$$

This implementation is cross-checked using the full numeric integral.

5.2. Proper time resolution

The proper time resolution due to finite tracking and vertex resolutions has the effect of diluting the oscillation amplitudes, thus reducing sensitivity to ϕ_s. The parameterisation of this resolution is determined in two ways, using both Monte-Carlo and data-driven

techniques. We adopt a triple Gaussian resolution model of the form:

$$R(\delta_t) = \sum_{i=0}^{3} f_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(\delta_t - \mu)^2}{2\sigma_i^2}\right] \tag{5.16}$$

Where a common mean, μ is used across all Gaussians of widths σ_i. δ_t is the difference between the measured and actual proper time: $\delta_t = t - t'$. In data t' is inaccessible while in Monte-Carlo $t' = t_{\text{gen}}$, the generated proper time. The three parameters f_i denote the fraction of each Gaussian and are normalised such that $\sum_{i=1}^{3} f_i = 1$. The dilution is obtained by taking the Fourier transform of the proper time resolution [95]:

$$D(\nu) = \frac{2}{\pi} \int_0^\infty R(\delta_t) \cos(-\nu\delta_t) d\delta_t \tag{5.17}$$

In which ν denotes the mixing frequency. For a Gaussian $R(\delta t)$ and the mixing frequency Δm_s this equates to:

$$D = \exp\left(-{\Delta m_s}^2 \sigma_i^2/2\right) \tag{5.18}$$

The dilution of the triple Gaussian model in equation (5.16) is the sum of the three individual dilutions:

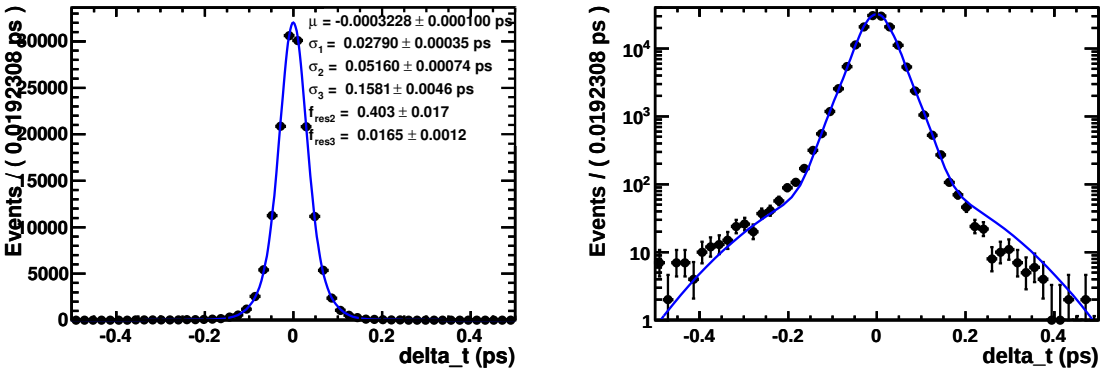
$$D = \sum_{i=1}^{3} f_i \exp\left(-\Delta m_s^2 \sigma_{t;i}^2/2\right) \tag{5.19}$$

It is also convenient to define an effective proper time resolution, $\sigma(D)$ corresponding to a single Gaussian resolution with a dilution that is the same as that of the triple Gaussian model:

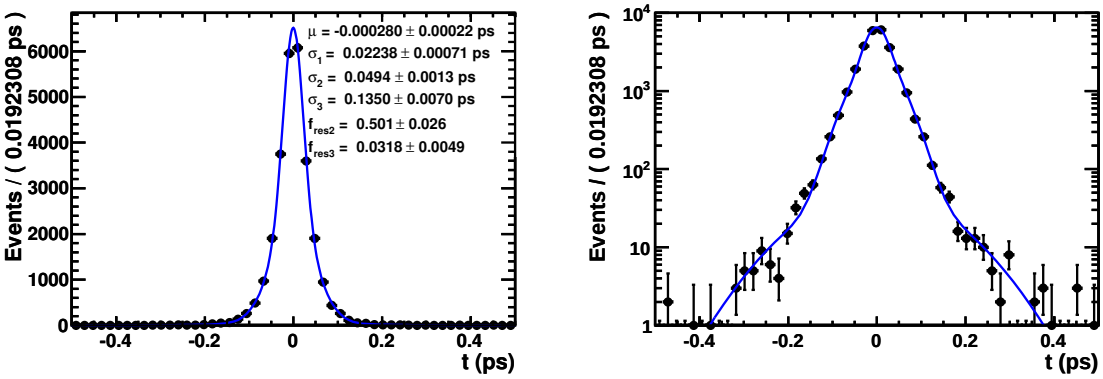
$$\exp\left(-\Delta m_s^2 \sigma_t(D)^2/2\right) = \sum_i f_i \exp\left(-\Delta m_s^2 \sigma_{t;i}^2/2\right) \tag{5.20}$$

5.2.1. Determination of Monte-Carlo proper time resolution

A straightforward technique to determine the proper time resolution is to study the proper time distribution in prompt decays, where any deviation from $\delta t = 0$ must be as a result of the resolution alone. The Monte-Carlo sample of inclusive J/ψ candidates is subjected to the same selection and stripping cuts as that of data but without the proper time cut of > 0.3 ps, providing a copious source of prompt J/ψ candidates further



(a) Simulated $B_s^0 \to J/\psi\,\phi$, $t > 0.3$ ps on linear and logarithmic scales



(b) Simulated prompt J/ψ on linear and logarithmic scales

Figure 5.5.: Monte-Carlo simulated δt and proper time distributions for (a) $B_s^0 \to J/\psi\,\phi$ candidates and prompt (b) J/ψ candidates. A triple-Gaussian fit is overlaid.

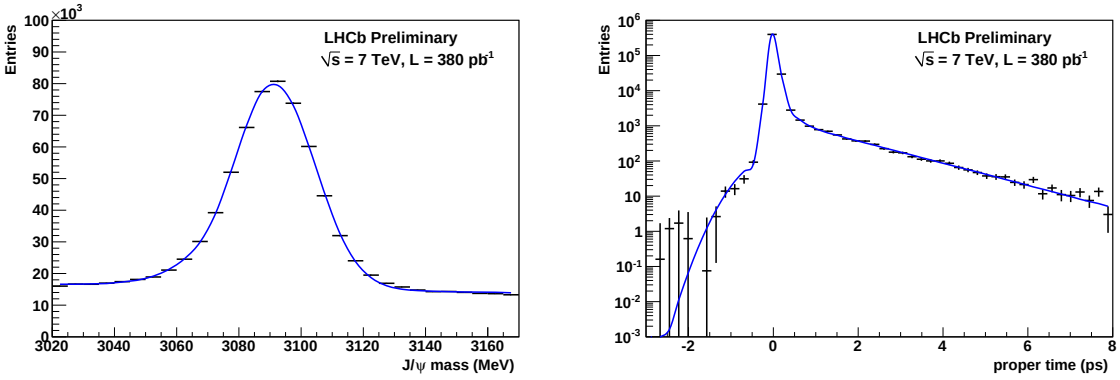
Parameter	Signal	Prompt
μ (fs)	-0.323 ± 0.10	-0.28 ± 0.2
σ_1 (fs)	27.9 ± 0.4	22.4 ± 0.7
σ_2 (fs)	51.6 ± 0.7	49.0 ± 1.0
σ_3 (fs)	158 ± 5	135 ± 7
f_2	0.40 ± 0.02	0.50 ± 0.03
f_3	0.017 ± 0.001	0.032 ± 0.005
D (Eq. 5.19)	0.780	0.776
D from FT (Eq. 5.17)	0.779	0.774
$\sigma(D)$ (fs)	39.9	40.4

Table 5.4.: Fit parameters of the resolution model for $B_s^0 \to J/\psi\,\phi$ and prompt J/ψ candidates. The dilutions are computed using a mixing frequency of 17.7 ps^{-1}. The last three rows show the dilutions and effective resolution, $\sigma(D)$.

enhanced by the exclusion of the selection level ϕ mass window. In total 29000 simulated B_s^0 candidates are selected and pass the proper time unbiased trigger, of which 27000 are found by truth-matching to be prompt J/ψ, from which the resolution may be extracted. As a cross-check to ensure that the procedure is sensible the $B_s^0 \to J/\psi\,\phi$ Monte-Carlo sample is also used to determine the resolution from a fit to $\delta t = t - t_{\text{gen}}$, the difference between the reconstructed and generated proper time. Figure 5.5a presents the δt distribution in Monte-Carlo $B_s^0 \to J/\psi\,\phi$ candidates. Superimposed is a fit to equation (5.16). In Figure 5.5b the proper time distribution of prompt J/ψ Monte-Carlo is shown with the same fit function overlaid. The results of both fits in Figure 5.5 are presented in Table 5.4. Also presented are the dilutions determined from equation (5.19), by integrating the histogram as per equation (5.17) and the effective resolution from equation (5.20). The similarity between the result obtained by integrating the histogram and the computed dilution is a powerful test: The Fourier transform method does not make use of the triple-Gaussian lineshape. The similarity between the two computed dilutions implies that the triple-Gaussian is an accurate and sufficient estimator of the resolution. The fits to δt in the signal channel and prompt J/ψ background channel have different parameterisations but result in identical effective resolutions and dilutions. This confirms that the resolution can be extracted from the background in data, with the differences due to the model taken as a systematic from Monte-Carlo.

5.2.2. Data-driven determination of proper time resolution

The data-driven proper time resolution is determined by fitting the resolution model to the proper time $\mathcal{S}$-plot for events passing the selection and proper time unbiased trigger on the range $-3 < t < 8$ ps. As a cross-check and to aid in determining systematic uncertainties the fit is applied also to the range $-1 < t < 8$ ps. The $\mathcal{S}$-weights are determined from the $M(\mu^-\mu^+)$ distribution so that both signal $B_s^0 \to J/\psi\,\phi$ and prompt J/ψ candidates form the proper time distribution. The distribution is modelled as the sum of two exponentials used to describe the long-lived $B_s^0 \to J/\psi\,\phi$ component and the triple-Gaussian resolution model in equation (5.16) to describe the prompt component from which the resolution parameters can be extracted. The $\mathcal{S}$-weight fit to select the J/ψ component is shown in Figure 5.6a and the fit to extract the proper time resolution from these J/ψ candidates is shown in Figure 5.6b. The results of the fit to both proper time ranges are presented in Table 5.5. It is clear from a comparison of Table 5.5 with



(a) $M(\mu^+\mu^-)$ for selected $B_s^0 \to J/\psi\,\phi$ data **(b)** B_s^0 proper time $\mathcal{S}$-plot

Figure 5.6.: Distributions used to extract the proper time resolution in data. (a) The $M(\mu^+\mu^-)$ distribution of selected $B_s^0 \to J/\psi\,\phi$ candidates on the proper time range $-3 < t < 8$ ps. Overlaid is a fit from which the J/ψ signal $\mathcal{S}$-weights are determined. (b) The proper time $\mathcal{S}$-plot determined from these $\mathcal{S}$-weights, to which a proper time model is fit consisting of the sum of the resolution model describing prompt J/ψ candidates and a double exponential describing the $B_s^0 \to J/\psi\,\phi$ signal distribution.

Parameter	$-3 < t < 8$ ps	$-1 < t < 8$ ps
μ (fs)	-2.7 ± 0.1	-2.6 ± 0.1
σ_1 (fs)	43.4 ± 0.2	35.5 ± 0.6
σ_2 (fs)	85.3 ± 1.1	64.4 ± 0.9
σ_3 (fs)	513 ± 39	236 ± 10
f_2	0.165 ± 0.007	0.494 ± 0.021
f_3	0.0017 ± 0.0002	0.009 ± 0.001
D (Eq. 5.19)	0.673	0.666
D from FT (Eq. 5.17)	0.667	0.667
$\sigma(D)$ (fs)	50.3	50.9

Table 5.5.: Fit parameters of the resolution model for prompt J/ψ candidates. The dilutions are computed using a mixing frequency of 17.7 ps^{-1}. The last three rows show the dilutions and the corresponding effective resolution, $\sigma(D)$.

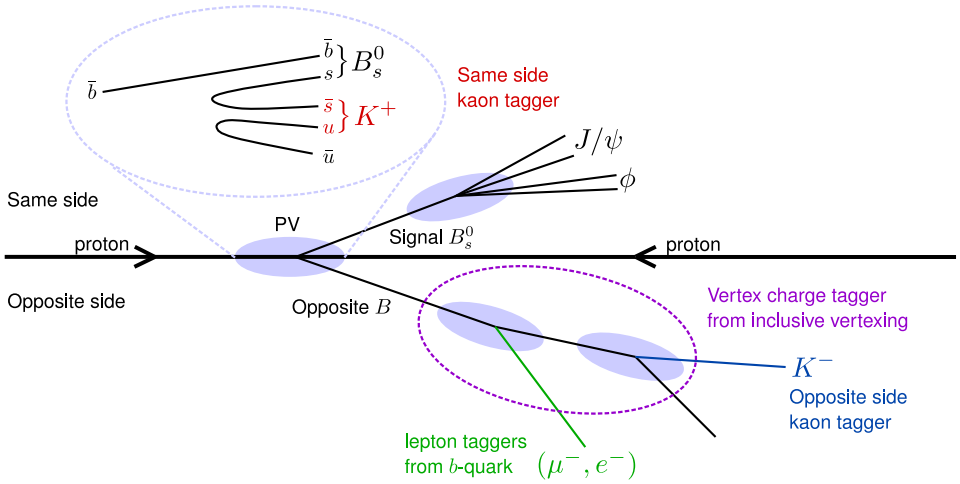


Figure 5.7.: Illustration depicting the flavour tagging algorithms used at LHCb to determine the initial B flavour. Shown are the Same Side Kaon (SSK) tagger which exploits the products of the fragmentation of the signal $\overline{\text{b}}$ quark, the Opposite Side lepton and kaon single particle taggers which look for charged daughters of the associated B decay, and the vertex charge tagger which reconstructs the opposite side decay and sums the charge back to the B vertex.

Table 5.4 that the effective resolution in data of about 50 fs is worse than that of 40 fs in Monte-Carlo simulation. This is consistent with previous studies on 2010 data [96].

5.3. Flavour tagging

The differential decay rates listed in Section 4.3 transform between B_s^0 and $\overline{\text{B}}_\text{s}^0$ by changing the sign of the mixing terms:

$$\sin(\Delta m_s t) \to -\sin(\Delta m_s t) \tag{5.21}$$

$$\cos(\Delta m_s t) \to -\cos(\Delta m_s t) \tag{5.22}$$

Knowledge of the initial B_s^0 meson flavour is a challenge as it cannot be inferred from the $\text{B}_\text{s}^0 \to \text{J}/\psi\,\phi$ decay chain. Instead we must tag the B_s^0 flavour through additional information about the event. LHCb uses a standard approach to flavour tagging. Tagging algorithms are quantified by their tagging efficiency, ϵ_{tag}, and the mistag rate, ω which defines the rate of B_s^0 candidates incorrectly tagged as $\overline{\text{B}}_\text{s}^0$ or vice-versa. The effect of

imperfect tagging leads to a dilution on tag-dependent terms of the form:

$$\sin(\Delta m_s t) \to qD \sin(\Delta m_s t) \tag{5.23}$$

$$\cos(\Delta m_s t) \to qD \cos(\Delta m_s t) \tag{5.24}$$

Where $D = (1-2\omega)$ and $q = \pm 1$ is the result of the flavour tag: $+1$ for candidates tagged as $\mathrm{B_s^0}$, -1 for $\overline{\mathrm{B}}{}_\mathrm{s}^0$. The effective statistical reduction in sensitivity to $\mathcal{CP}$ asymmetries as a result of imperfect tagging, ϵ_{eff} can be determined with knowledge of the tagging efficiency and mistag rate by:

$$\epsilon_{\mathrm{eff}} = \epsilon_{\mathrm{tag}}(1-\omega)^2 = \epsilon_{\mathrm{tag}} D^2 \tag{5.25}$$

Due to the small additional sensitivity from untagged candidates we keep them with $\omega = 0.5$ for the purposes of this analysis, meaning that the tagging dilution slightly overestimates the reduction in sensitivity to ϕ_s. Several tagging algorithms are used at LHCb, which can be grouped into two categories:

- The Same Side (SS) tagger looks for kaons produced during the fragmentation process of the $\overline{\mathrm{b}}$ quark in the signal $\mathrm{B_s^0}$.
- The Opposite Side (OS) taggers use information from the decay containing the other b quark produced in the proton-proton collision.

An additional distinction between the tagging algorithms is that the OS tagger exhibits an intrinsic dilution due to the finite probability that the opposite side B meson mixes prior to decay. The SS tagger is not subject to this dilution. Calibration and optimisation of the SS tagger would ideally require a large sample of self-tagging decay mode $\mathrm{B_s^0}$ decays which at the time of this analysis was not available. Initial optimisation studies were performed using $\mathrm{D_s^\pm} \to \phi\pi^\pm$, and preliminary studies using $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ indicate a $\sim 10\%$ improvement in the statistical uncertainty on ϕ_s if the SS tagger is included but calibration of the SS tagger was not available in time. As a result only the combined OS tag decision is used in this analysis. The selection and tuning of the OS taggers is an iterative procedure to maximise the tagging power, initially performed on Monte-Carlo [97], and reoptimised on data. The criteria for the individual taggers are outlined in the following subsections.

Tagger	p_{T} (GeV/c)	p (GeV/c)	$\mathrm{IP}/\sigma_{\mathrm{IP}}$	Particle ID	$\mathrm{IP}_{\mathrm{PU}}/\sigma_{\mathrm{IP}_{\mathrm{PU}}}$
$\mu^{\pm}$	> 1.2	> 2	-	$\Delta_{LL}(\mu^{\pm} - \pi^{\pm}) > 2.5$	> 3.0
$\mathrm{e}^{\pm}$	> 1.0	> 2	> 2.0	$\Delta_{LL}(\mathrm{e}^{\pm} - \pi^{\pm}) > 4$	> 3.0
$\mathrm{K}^{\pm}$	> 0.8	> 5.9	> 4.0	$\Delta_{LL}(\mathrm{K}^{\pm} - \pi^{\pm}) > 6.5$, $\Delta_{LL}(\mathrm{K}^{\pm} - \mathrm{p}) > -3.5$	> 4.7

Table 5.6.: Single particle tagger selection criteria. $\mathrm{IP}/\sigma_{\mathrm{IP}}$ denotes the impact parameter significance with respect to the Primary Vertex, with $\mathrm{IP}_{\mathrm{PU}}/\sigma_{\mathrm{IP}_{\mathrm{PU}}}$ denoting the same significance on any pile-up vertices.

5.3.1. $\mu^{\pm}$, $\mathrm{e}^{\pm}$, $\mathrm{K}^{\pm}$ taggers

The single particle taggers exploit semileptonic decays and kaons from the $\mathrm{b} \to \mathrm{c} \to \mathrm{s}$ decay chain to determine the flavour of the opposite side B meson at decay. Only tracks with a good quality of fit are used, and are required to be isolated from any decay products of the signal $\mathrm{B}_{\mathrm{s}}^{0}$. A series of cuts are applied specific to each particle type designed to ensure that the track is correctly associated to a secondary (or tertiary in the case of the kaon) vertex consistent with a B decay, namely a large Impact Parameter significance with respect to both the Primary Vertex (PV) and any Pile-Up (PU) vertices, large transverse momentum and particle ID cuts in order to ensure the track is correctly identified for the algorithm being used. These are listed in Table 5.6. Additional cuts are applied specific to the leptons: For the $\mathrm{e}^{\pm}$ candidates a maximum ionization charge deposited in the silicon layers of the VELO is required in order to reduce the number of candidates coming from photon conversions close to the interaction point. An additional cut on $E/p > 0.6$ is applied based on the ECAL energy measurement and the momentum determined by the tracking system. For $\mu^{\pm}$ candidates a track isolation cut is applied so that muon chamber hits cannot be shared by more than one track. In the case of multiple candidates passing a single particle tagger selection the candidate with the highest p_{T} is chosen.

5.3.2. Vertex charge tagging

The Vertex charge tagger reconstructs the B decay vertex and sums the charges of the tracks. Two isolated, good quality tracks of $p_{\mathrm{T}} > 0.15$ GeV/c and $2.5 < \mathrm{IP}/\sigma_{\mathrm{IP}} < 100$ are combined to form a seed vertex with a vertex $\chi^2/DoF < 10$. A likelihood is constructed

based on the kinematics and geometry of the vertex as well as the quality of the fit. The seed vertex with the greatest likelihood is then chosen. Additional tracks are added subject to the requirement that they have a large impact parameter χ^2 with respect to the primary vertex and are compatible with the seed vertex (a distance of closest approach less than 0.2 mm, impact parameter on the seed vertex less than 0.7 mm). The resulting secondary vertex must have $p_{\mathrm{T}} > 10$ GeV/c, an invariant mass greater than 0.5 GeV/c^2 and an impact parameter significance with respect to the PV of all tracks greater than 10. The vertex charge is then calculated as the sum of all track charges Q_i weighted by their transverse momentum to the power $\kappa = 0.4$ which is empirically determined to maximise the tagging power.

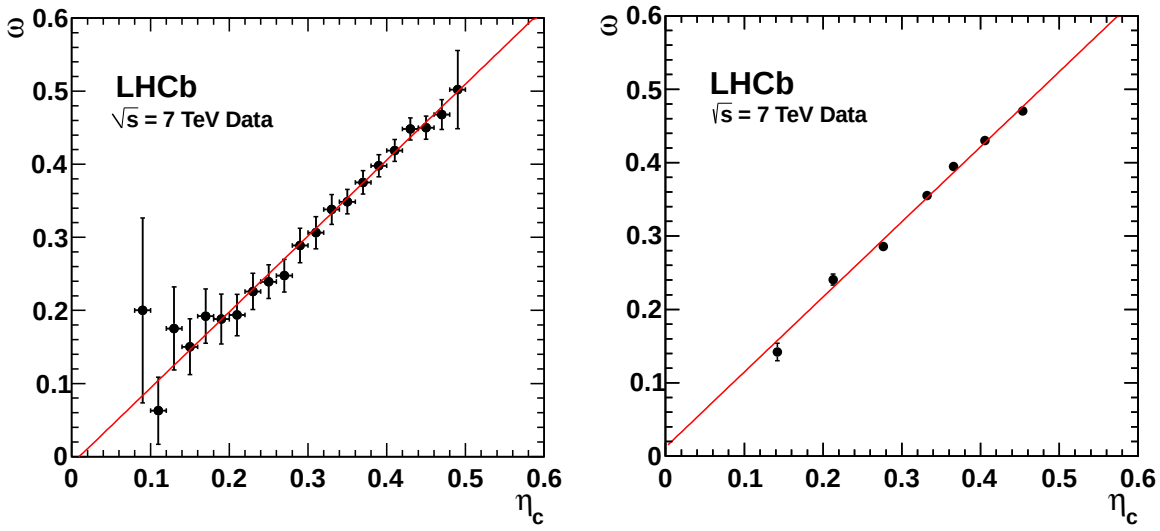
$$Q_{vtx} = \sum_i \frac{Q_i {p_{\mathrm{T}}}_i^{\kappa}}{\sum_j {p_{\mathrm{T}}}_j^{\kappa}} \tag{5.26}$$

Events for which $|Q_{vtx}| < 0.275$ are not given a vertex charge tag.

5.3.3. Mistag probability and combined flavour tag

In addition to the flavour tag, each tagger returns a per-event mistag probability which is estimated from the properties of both the tagging algorithm and the event being tagged. The mistag probability is determined using a neural network trained on $\mathrm{B}^+ \to \mathrm{J}/\psi\mathrm{K}^+$ Monte-Carlo. The neural network inputs consist of the B momentum, the number of pileup vertices the number of tracks preselected as tagging candidates and geometrical and kinematic properties of the tagging particle or of the secondary vertex. For the single particle taggers the momentum, transverse momentum and impact parameter of the tagging particle is used. For the vertex charge tagger the total number of tracks in the event, the number of tracks associated to the secondary vertex, the seed likelihood, vertex charge and lifetime of the tagging B are used in addition to the average p_{T}, impact parameter and distance of closest approach for all tracks associated to the secondary vertex. For events in which more than one tagger provides a decision, the combined tagging probability is determined:

$$P(b) = \frac{p(b)}{p(b) + p(\bar{b})}, \qquad P(\bar{b}) = 1 - P(b), \tag{5.27}$$



(a) $B^{\pm} \to J/\psi K^{\pm}$ **(b)** $B \to D^{*}\mu^{+}\nu_{\mu}$

Figure 5.8.: Measured mistag fraction (ω) versus calculated mistag probability η_c for (a) background subtracted $B^{+} \to J/\psi K^{+}$ signal candidates and (b) $B \to D^{*}\mu^{+}\nu_{\mu}$. Datapoints are shown in black. In red is shown the mistag calibration fit [98].

where

$$p(b) = \prod_i \left(\frac{1+d_i}{2} - d_i(1-\eta_i) \right), \qquad p(\bar{b}) = \prod_i \left(\frac{1-d_i}{2} + d_i(1-\eta_i) \right). \tag{5.28}$$

Here, d_i is the decision taken by the i-th tagger based on the charge of the particle with the convention $d_i = 1(-1)$ for the signal B containing a $\bar{b}(b)$ quark and η_i the corresponding predicted mistag probability. The combined tagging decision and the corresponding mistag probability are $d = -1$ and $\eta = 1 - P(b)$ if $P(b) > P(\bar{b})$, otherwise $d = +1$ and $\eta = 1 - P(\bar{b})$.

5.3.4. Tagging calibration

The neural network has been trained on Monte-Carlo. The output per-event mistag on data requires calibration before it can be used in analyses. This calibration is performed first on a per-tagger basis and then on the combined tagger output using the self-tagging $B^{\pm} \to J/\psi K^{\pm}$ where the charge of the kaon dictates the B flavour permitting a direct measurement of ω. The measured mistag (ω) as a function of neural net per-event calibrated mistag output (η_c) is shown in Figure 5.8a [98]. The calibration to be applied

Channel	Yield	p_0	p_1	$\langle\eta_c\rangle$	$p_0 - p_1\langle\eta_c\rangle$	$\rho(p_0, p_1)$
$B^+ \to J/\psi K^+$	~ 84k	0.384 ± 0.003	1.037 ± 0.038	0.379	-0.009 ± 0.014	0.14
$B \to J/\psi K^*$	~ 39k	0.399 ± 0.008	1.016 ± 0.102	0.378	0.015 ± 0.039	0.05
$B \to D^* \mu^+ \nu_\mu$	~ 380k	0.395 ± 0.002	1.022 ± 0.026	0.375	0.008 ± 0.010	0.14

Table 5.7.: Results of the tagging calibration parameters measured in the listed background subtracted signal samples after calibration on the decay mode $B^+ \to J/\psi K^+$. The uncertainty is statistical only. The linear correlation coefficient between parameters p_0 and p_1 is listed in the last column.

is modelled as:

$$\omega(\eta_c) = p_0 + p_1(\eta - \langle\eta\rangle) \tag{5.29}$$

Where p_0 and p_1 are free parameters which define the calibration and $\langle\eta\rangle$ is the mean calculated mistag probability. The calibrated per-event mistag probability is η_c. For a per-event mistag that is fully calibrated $p_0 = \langle\eta\rangle$ and $p_1 = 1$. There is some correlation between individual tagger outputs resulting in an overestimation of the tagging power. The calibration to the combined mistag accounts for this, leading to a slight increase in the calibrated per-event mistag probability, $\langle\eta_c\rangle$ Figure 5.8 shows the results of the tagging calibration determined from $B^+ \to J/\psi K^+$ data and applied to the same, in addition to a cross-check on $B \to D^* \mu^+ \nu_\mu$. The results of the calibration for both of these decay modes in addition to the mode $B \to J/\psi K^*$ are listed in Table 5.7. The final

tagging calibration parameters used for this analysis are those taken from the fit to the calibrated $B^+ \to J/\psi K^+$ mistag:

$$p_0 = 0.384 \pm 0.003 \pm 0.009 \quad p_1 = 1.037 \pm 0.04 \pm 0.07 \quad (< \eta_c >= 0.379) \tag{5.30}$$

Where the first uncertainty is statistical and the second is systematic.

5.3.5. Tagging performance

In order to exploit maximally the tagging information we make use of the per-event calibrated dilution $D_i = (1 - 2\omega(\eta_c^i))$ in the likelihood fit to extract ϕ_s, where $\omega(\eta_c^i)$ is

the calibrated mistag fraction of the ith event. The effective dilution is then

$$D_{\text{eff}} = \frac{1}{N}\sum_{i}^{N}(1 - 2\omega(\eta_c^i)) \tag{5.31}$$

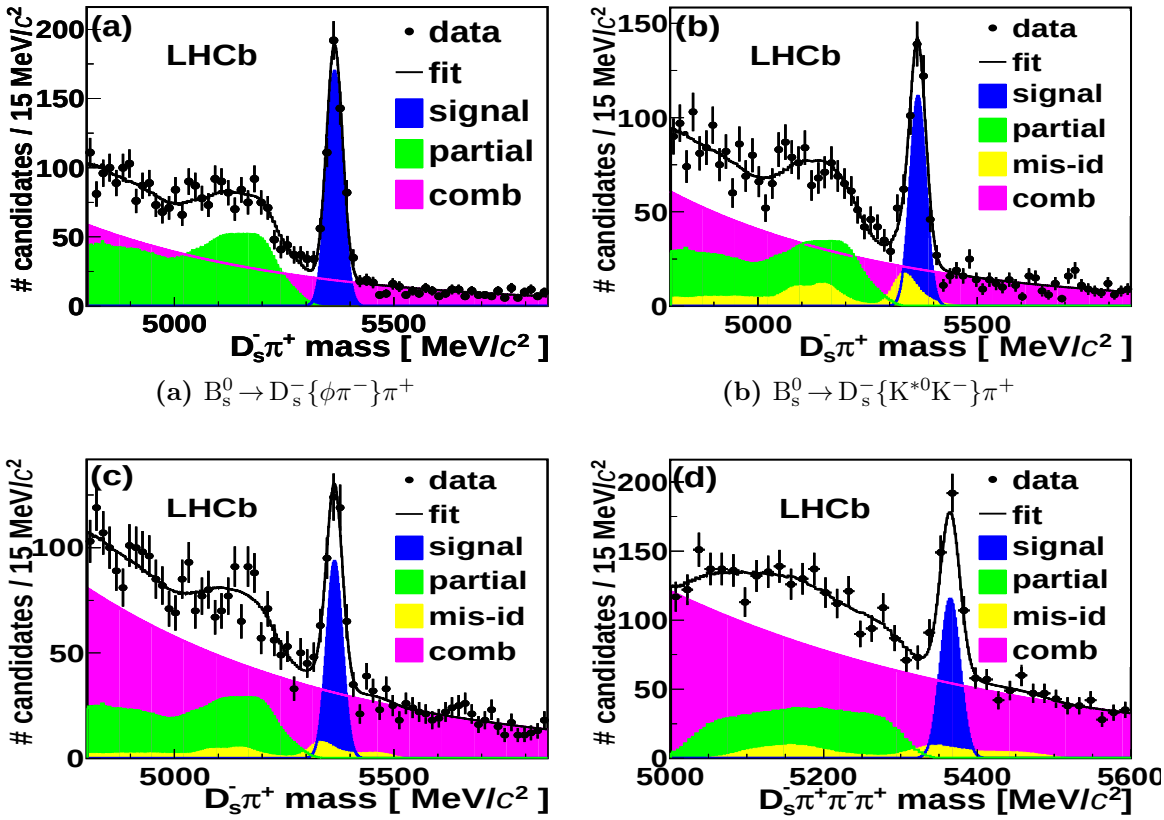
Using the η_c distribution of selected $\mathcal{S}$-weighted $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ signal candidates the effective dilution is determined to be $D_{\text{eff}} = 0.277 \pm 0.006 \pm 0.016$ where the uncertainties are propagated from the calibration parameters in equation (5.30). This corresponds to an effective mistag rate of $\omega_{\text{eff}} = (36.1 \pm 0.3 \pm 0.8)\%$. The number of tagged candidates in the signal $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ sample corresponds to a tagging efficiency of $\epsilon_{\text{tag}} = (24.9 \pm 0.5)\%$. Using equation (5.25) the effective per-event tagging power is then $\epsilon_{\text{eff}} = (1.91 \pm 0.08 \pm 0.22)\%$.

5.4. Determination of Δm_s

$\mathrm{B_s^0}$-$\overline{\mathrm{B}}{}_\mathrm{s}^0$ mixing occurs at a frequency which is 35 times faster than that of the $\mathrm{B_d^0}$-$\overline{\mathrm{B}}{}_\mathrm{d}^0$ system. A precise measurement of Δm_s is not only important for the ϕ_s measurement, but also as a standalone LHCb result. With the 36 pb^{-1} of data collected in 2010, LHCb measured Δm_s in the flavour-specific modes $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ and $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+\pi^-\pi^+$ where the $\mathrm{D_s^-}$ decays to $\mathrm{K^+K^-}\pi^-$ [99]. The intermediate resonances $\mathrm{D_s^-} \to \mathrm{K^{*0}K^-}$ and $\mathrm{D_s^-} \to \phi\pi^-$ are selected separately in the $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ mode exploiting additional cuts on the resonance mass. To prevent double counting, events that are selected in one channel are prevented from being selected by another. In order to extract Δm_s an unbinned likelihood fit is performed to the $\mathrm{B_s^0}$ mass and proper time distribution simultaneously across all decay modes.

5.4.1. Mass fit description and signal channel yields

For the signal $M(\mathrm{D_s^\pm}\pi^\mp\{\pi^+\pi^-\})$ distribution a Gaussian function is used in which the mean is kept as a single parameter across all modes in the fit but the widths are allowed to float separately for $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+\pi^-\pi^+$ and the combined $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ sample. The fits to the mass distribution incorporate distributions for sources of background, namely: Partially reconstructed $\mathrm{B_s^0}$ candidates, $\mathrm{B_d^0}$ and Λ_b decays in which one daughter particle is misidentified and a combinatoric background. Background shapes in the $M(\mathrm{D_s^\pm}\pi^\mp\{\pi^+\pi^-\})$ distributions for each mode were determined from a high-statistics



(a) $B_s^0 \to D_s^-\{\phi\pi^-\}\pi^+$ (b) $B_s^0 \to D_s^-\{K^{*0}K^-\}\pi^+$

(c) $B_s^0 \to D_s^-\{K^+K^-\pi^-\}\pi^+$ (d) $B_s^0 \to D_s^-\{K^+K^-\pi^-\}\pi^+\pi^-\pi^+$

Figure 5.9.: Mass distributions and fits to the selected B_s^0 candidates in flavour-specific decay modes. In the legend "partial" denotes background from partially reconstructed B_s^0 decays, "mis-id" refers to B_d^0 and Λ_b candidates with one mis-identified daughter and "comb" denotes the combinatoric background component [99].

Channel	Signal yield
$B_s^0 \to D_s^-\{\phi\pi^-\}\pi^+$	515 ± 25
$B_s^0 \to D_s^-\{K^{*0}K^-\}\pi^+$	338 ± 27
$B_s^0 \to D_s^-\{K^+K^-\pi^-\}\pi^+$	283 ± 27
$B_s^0 \to D_s^-\{K^+K^-\pi^-\}\pi^+\pi^-\pi^+$	245 ± 46
Total	1381 ± 65

Table 5.8.: Signal yields for each channel determined from the $M(D_s^\pm\pi^\pm\{\pi^+\pi^-\})$ distributions shown in Figure 5.9.

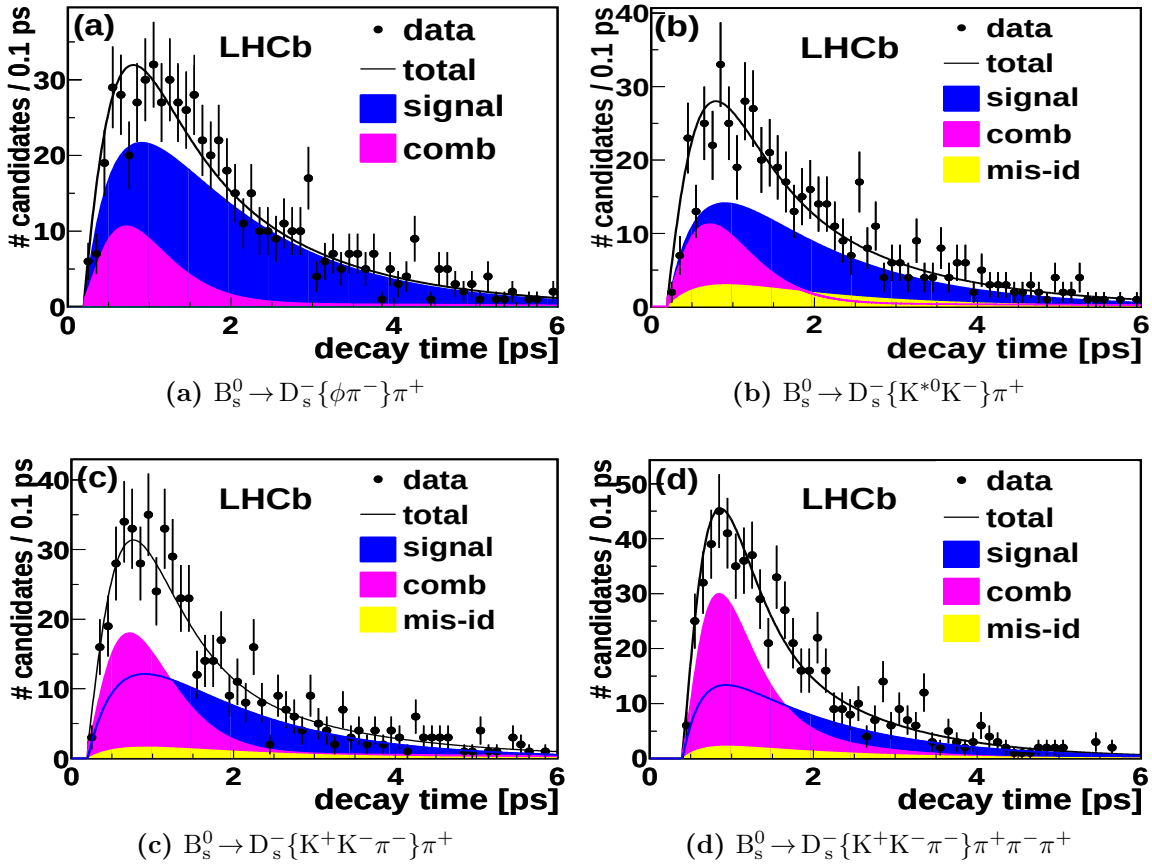
generator level study, smeared with a Gaussian detector resolution and verified with a fully simulated dataset. The background fractions and combinatoric exponential parameterisation are allowed to float separately for each decay mode. The mass fit is shown in Figure 5.9 for each of the decay modes and the resulting signal yields are listed in Table 5.8 [99]. The fit returns $\sigma_m = 12.7\text{MeV}/c$ for the $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ mode and $\sigma_m = 18.1\text{MeV}/c$ for the $B_s^0 \to D_s^-\pi^+$ modes. For the simultaneous fit to extract Δm_s these widths are fixed and the mass windows are restricted to $\pm 3\sigma$ in order to remove the partially reconstructed B_s^0 background.

5.4.2. Propertime fit description

The signal proper time PDF is of the form:

$$\begin{aligned}\mathcal{P}_t(t,q|\sigma_t,\eta_c) \;\propto\; & \epsilon(t)\Big[\epsilon_{\text{tag}}\Big\{\Gamma_s e^{-\Gamma_s t}\,\frac{1}{2}\big[\cosh(\frac{\Delta\Gamma_s}{2}t) + q(1-2\omega(\eta_c))\cos(\Delta m_s t)\big]\Big\} \\ & +(1-\epsilon_{\text{tag}})\Big\{\Gamma_s e^{-\Gamma_s t}\,\frac{1}{2}\big[\cosh(\frac{\Delta\Gamma_s}{2}t)\big]\Big\}\Big] \otimes G(t,\sigma_t). \qquad (5.32)\end{aligned}$$

Where the function $\omega(\eta_c)$ is the per-event calibrated mistag and q the flavour tag as described in Section 5.3, using the 2010 data calibration. The tagging efficiency ϵ_{tag} is left as a free parameter in the fit and kept separate for the $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ and $B_s^0 \to D_s^-\pi^+$ modes. The proper time acceptance function $\epsilon(t)$ is of the form described in Section 5.1.2 with acceptance parameters derived from Monte-Carlo. The proper time resolution $G(t,\sigma_t)$ consists of the per-event estimated proper time uncertainty σ_t as determined by the decay fitting algorithm multiplied by a scale factor to account for spatial misalignment and imperfect understanding of detector material. The scale



(a) $B_s^0 \to D_s^-\{\phi\pi^-\}\pi^+$ (b) $B_s^0 \to D_s^-\{K^{*0}K^-\}\pi^+$

(c) $B_s^0 \to D_s^-\{K^+K^-\pi^-\}\pi^+$ (d) $B_s^0 \to D_s^-\{K^+K^-\pi^-\}\pi^+\pi^-\pi^+$

Figure 5.10.: Proper time distributions and fits to the selected B_s^0 candidates in flavour-specific decay modes. In the legend "mis-id" refers to B_d^0 and Λ_b candidates with one mis-identified daughter and "comb" denotes the combinatoric background component [99].

factor is extracted from the lifetime distribution of fake B_s^0 candidates constructed from a prompt D_s^- and a π^+ from the primary vertex. The proper time distribution of these candidates divided by their per-event proper time uncertainty is fit to a Gaussian function where the width is used as the scale-factor. This scale factor results in a nominal proper time resolution of 36 fs for $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ candidates and 44 fs for $B_s^0 \to D_s^-\pi^+$ candidates. In the fit $\Delta\Gamma_s$ is fixed to the nominal PDG value of $0.09 \cdot \Gamma_s$ with Γ_s initially left to float. The background proper time distribution for Λ_b and B_d^0 mis-id candidates is constructed in the same way as that of signal, using their PDG values for lifetimes and with $\Delta\Gamma = 0$. For the combinatoric background a double exponential function is used, multiplied by a polynomial to account for acceptance. Γ_s is found to be consistent within 1σ of the PDG value for all modes. Figure 5.10 presents the proper time fits for each channel [99].

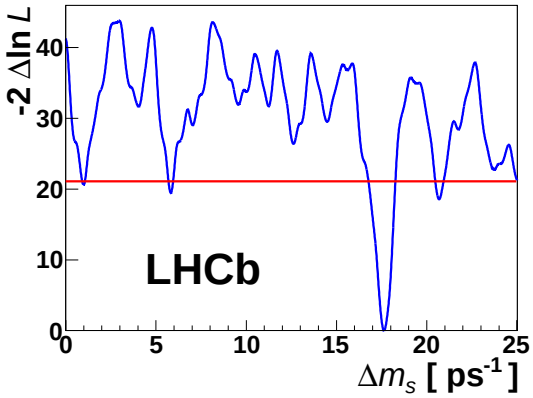


Figure 5.11.: Profile likelihood for Δm_s in the range $[0.0, 25.0]$ ps^{-1}. The line at $-2\Delta \ln L = 20.9$ indicates the value in the limit $\Delta m_s = \infty$ [99].

5.4.3. Results

The full mass and proper time fit is performed simultaneously to each of the four decay modes with Gaussian constraints applied to the mistag calibration parameters p_0 and p_1. The lifetime and mass parameters are fixed to their fitted values, allowing only Δm_s and tagging calibratrion parameters to float. The fit returns the value $\Delta m_s = 17.63 \pm 0.11 \pm 0.02$ ps^{-1} where the first uncertainty is statistical. The second uncertainty is the quadratic sum of systematic uncertainties the largest of which is due to the z scale. The profile likelihood of the fit is shown in Figure 5.11, where it can be seen that the statistical significance of the signal is 4.6σ with respect to the NLL at $\Delta m_s = \infty$. In the fit to $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ to extract physics parameters, this value of Δm_s is used as a Gaussian constraint.

Chapter 6.

A measurement of ϕ_s with 0.37 fb^{-1}

"The story of a theory's failure often strikes readers as sad and unsatisfying. Since science thrives on self-correction, we who practice this most challenging of human arts do not share such a feeling. We may be unhappy if a favored hypothesis loses or chagrined if theories that we proposed prove inadequate. But refutation almost always contains positive lessons that overwhelm disappointment, even when [...] no new and comprehensive theory has yet filled the void."

— Stephen Jay Gould

Chapters 4 and 5 have described in detail the dataset, analysis method and auxiliary inputs to the ϕ_s measurement. The contents of these chapters are brought together in this chapter. The fit is first studied to determine if it is possible to include an s-wave component, and the chapter culminates with an analysis of the decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ with the first two years of $\sqrt{s} = 7$ TeV data at LHCb, presenting a measurement of the $\mathcal{CP}$ violating phase ϕ_s, the decay width difference $\Delta\Gamma_s$ and associated physics parameters. The analysis presented here has been published in reference [100].

6.1. Sensitivity to an s-wave component

In the fit the requirement that $|\mathrm{A}_\perp|^2+|\mathrm{A}_\parallel|^2+|\mathrm{A}_0|^2+|\mathrm{A}_s|^2 = 1$ with all amplitudes greater than or equal to zero means that if any amplitude is small the likelihood will be non-parabolic. This can lead to bias in the fit if not accounted for. While previous analyses

have determined that $|A_\perp|^2, |A_\parallel|^2, |A_0|^2$ are sufficiently large, the s-wave component included in the fit as discussed in Section 4.3.2 is expected to be small. The treatment of the s-wave terms in the fit is therefore dependent upon how sensitive we are to $|A_s|^2$.

6.1.1. Monte-Carlo studies on 0.2 fb^{-1}

The procedure to determine how sensitive we are to the treatment of $|A_s|^2$ is as follows:

- 200 toy Monte-Carlo datasets are generated according to the parameterisation listed in Table 6.1 at values of $|A_s|^2 = 0\% \rightarrow 6\%$ in 1% intervals.
- Each toy dataset is fit to three times: Once with $|A_s|^2$ and the associated strong phase δ_s fixed to their generated values, once with $|A_s|^2$ and δ_s floated and once with $|A_s|^2$ and δ_s fixed to zero regardless of the generated value.
- The systematic shift in parameter values between the fully floated fit and the fixed fit indicates the level of bias present as a function of generated $|A_s|^2$. This is scaled by the systematic uncertainty on the parameter in the fully floated fit.

The procedure is repeated for $\phi_s = 0.0$ and $\phi_s = -0.7$ in order to account for biases that scale with the size of ϕ_s. The toy Monte-Carlo data is an approximation to the full Monte-Carlo, in which we generate only proper time unbiased data equivalent in size to the yield from both proper time unbiased + biased datasets expected in 200 pb^{-1}. Instead of a per-event mistag we use a single average mistag parameter of $\omega = 0.33$. The proper time resolution is modelled as a single Gaussian function of width 50 fs and there is no angular acceptance. With this prescription the statistical uncertainty on the amplitude terms is ~ 0.02. Table 6.2 shows the nominal fit results, sensitivities and biases in the absence of any s-wave component for each of the main physics parameters with $\phi_s = 0.0$. The pull is defined as:

$$\text{Pull}_i = \frac{x_i - x_{\text{gen}}}{\sigma_i} \tag{6.1}$$

Where x_i, σ_i is the fitted value and fit uncertainty of the parameter x determined from the i^{th} toy dataset generated with value x_{gen}. For a well-behaved fit in which the parameter uncertainty estimates are correct and in which no bias exists the pull distribution is Gaussian with $\mu_{\text{pull}} = 0$, $\sigma_{\text{pull}} = 1$.

Parameter	Generated value	Fixed/Floated
Γ_s (ps^{-1})	0.68	Float
$\Delta\Gamma_s$ (ps^{-1})	0.06	Float
$\|A_\perp\|^2$	0.16	Float
$\|A_0\|^2$	0.60	Float
$\|A_s\|^2$	$0.0 \rightarrow 0.06$	-
$\delta_\parallel$	2.50	Float
$\delta_\perp$	-0.17	Float
δ_s	0.5	-
ϕ_s	0.0,-0.7	Float
Δm_s (ps^{-1})	17.77	Constrained ± 0.12
τ resolution	0.05	Fixed
$<\omega>$	0.33	Fixed

Table 6.1.: Values of fit parameters with which toys are generated for the s-wave sensitivity study.

Parameter	μ_{fit}	σ_{fit}	μ_{pull}	σ_{pull}
Γ_s (ps^{-1})	0.7021 ± 0.0003	0.0205 ± 0.0002	0.03 ± 0.01	0.994 ± 0.009
$\Delta\Gamma_s$ (ps^{-1})	0.063 ± 0.001	0.0628 ± 0.0008	0.10 ± 0.02	1.07 ± 0.01
$\|A_\perp\|^2$	0.1614 ± 0.0003	0.0217 ± 0.0002	0.02 ± 0.01	1.04 ± 0.01
$\|A_0\|^2$	0.5992 ± 0.0002	0.0151 ± 0.0002	-0.03 ± 0.01	1.01 ± 0.01
$\delta_\parallel$	2.507 ± 0.002	0.113 ± 0.001	-0.21 ± 0.02	0.78 ± 0.01
$\delta_\perp$	-0.15 ± 0.01	0.76 ± 0.01	0.04 ± 0.02	1.10 ± 0.02
ϕ_s	-0.008 ± 0.005	0.318 ± 0.004	-0.01 ± 0.01	1.014 ± 0.009

Table 6.2.: Baseline toy dataset results for $|A_s|^2 = 0.0$ ignored, $\phi_s = 0.0$. The parameter values returned by the toys are fit to a Gaussian of mean and width μ_{fit}, σ_{fit}. The uncertainties on these parameters as returned by the fit are also shown. The last two columns are the result of fitting a Gaussian to the pull distribution as described in the text. $\delta_\parallel$ has a non-parabolic likelihood and so the pull is biased even in the baseline fit.

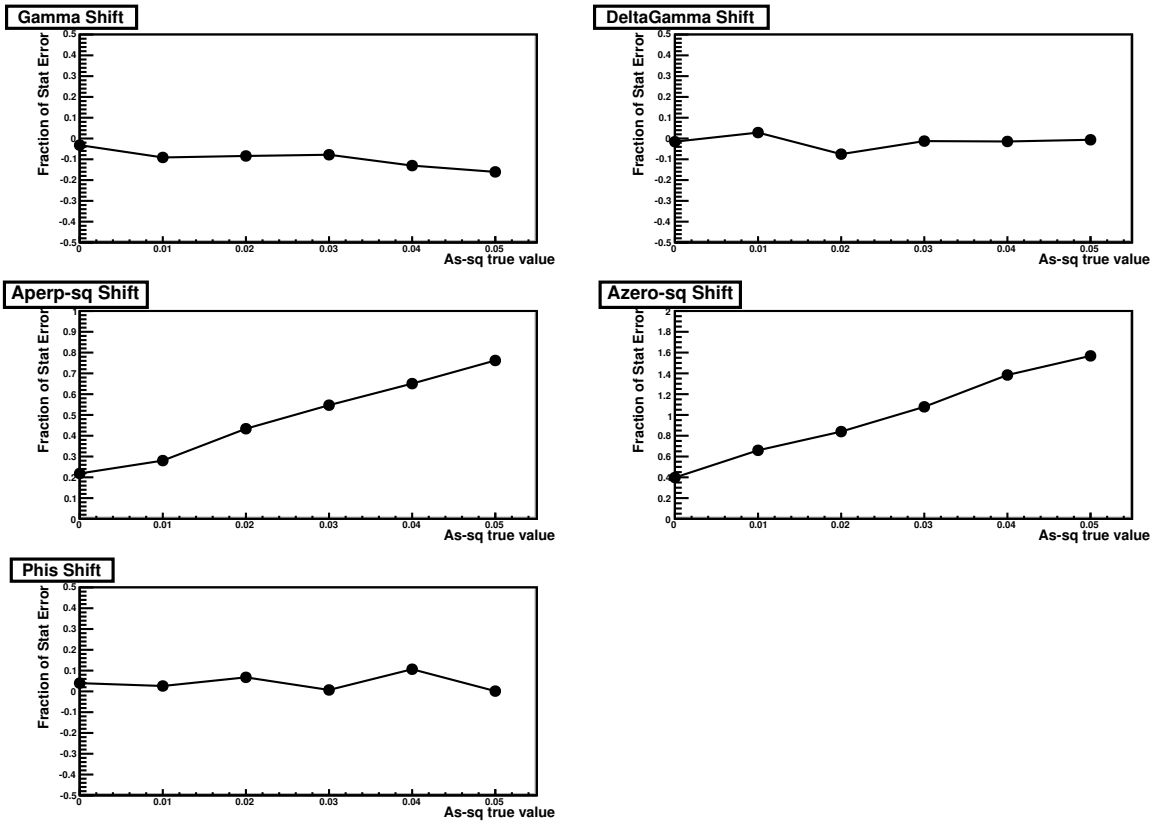


Figure 6.1.: Results with toy datasets for $|A_s|^2 = 0 \to 5\%$ generated, fixed to zero in the fit. The plots indicate the deviation from the fitted value when $|A_s|^2$ is floated to that of the fitted value when $|A_s|^2$ is fixed to zero, for data generated with $\phi_s = 0.0$. The deviations are in units of the statistical uncertainty.

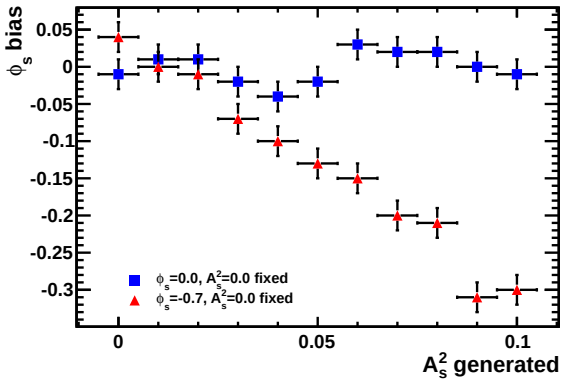


Figure 6.2.: Absolute bias on ϕ_s as a function of generated $|A_s|^2$ when the data is fit without an s-wave component for $\phi_s = 0.0$ and $\phi_s = -0.7$.

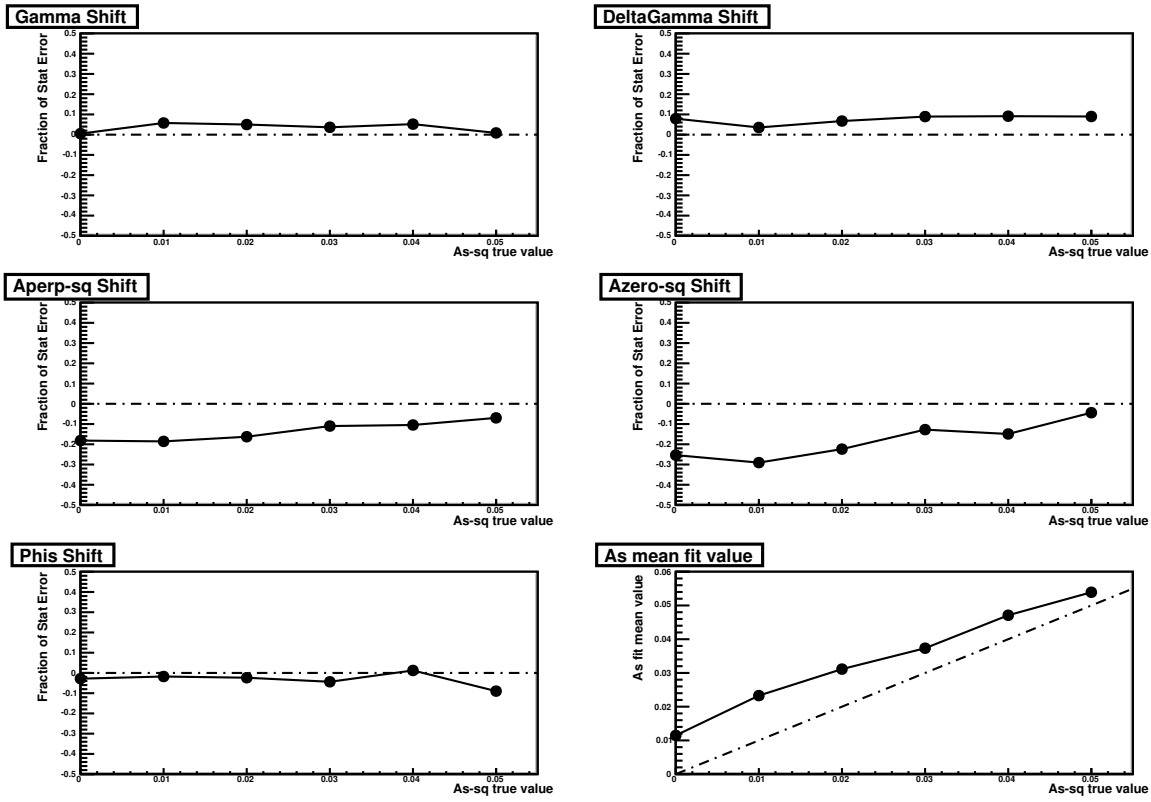


Figure 6.3.: Results with toy datasets for $\phi_s = 0.0$ with $|A_s|^2 = 0 \rightarrow 5\%$. The absolute deviation of the mean from the generated value for $|A_s|^2$ is shown, while the pull is shown as a function of $|A_s|^2$ for selected physics parameters.

6.1.2. Toy study results for $|A_s|^2$ ignored in the fit

Figure 6.1 presents the bias imparted to parameters when a non-zero s-wave component is ignored in the fit. The other amplitudes compensate for the missing component by biasing to larger values as a function of the generated $|A_s|^2$ value. Due to the shift in the amplitudes we would expect a corresponding shift in the measured value of ϕ_s: Figure 6.2 presents the bias imparted to ϕ_s in greater detail and includes an additional study in which ϕ_s is generated at -0.7. Here it can be seen that for nonzero generated values of ϕ_s there is a bias towards zero if $|A_s|^2$ is ignored in the fit.

6.1.3. Toy study results for $|A_s|^2$ floated in the fit

Figure 6.3 Shows the systematic shift divided by the statistical uncertainty averaged over each of the 200 toy datasets in the case where $\phi_s = 0.0$ for each of the pertinent physics parameters. The lower right-hand plot shows the absolute average fitted value

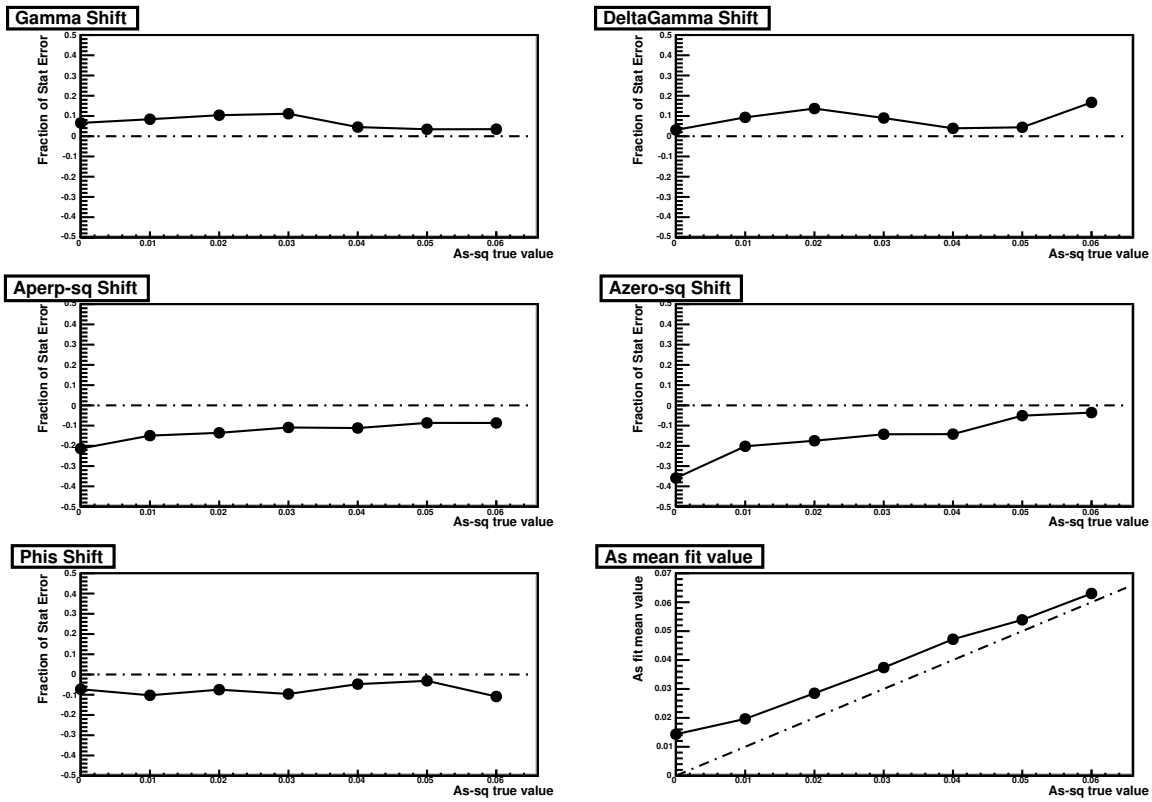


Figure 6.4.: Results with toy datasets for $\phi_s = -0.7$ with $|\mathrm{A}_s|^2 = 0 \rightarrow 5\%$ The absolute deviation of the mean from the generated value for $|\mathrm{A}_s|^2$ is shown, while the pull is shown as a function of $|\mathrm{A}_s|^2$ for selected physics parameters.

of $|A_s|^2$ as a function of the generated value. As $|A_s|^2$ increases the bias on it and the other amplitudes decreases as on would expect. $\phi_s, \Gamma_s, \Delta\Gamma_s$ remain stable as a function of the generated $|A_s|^2$ value. The same study using $\phi_s = -0.7$ is shown in Figure 6.4. Based on this study it would appear that the only parameters affected by the size of the s-wave component are the amplitudes when $|A_s|^2$ is allowed to float, and a systematic uncertainty can be assigned to them based upon the measured value of $|A_s|^2$. For the approximately 200 pb^{-1}sample that these toys represent the bias on the amplitudes is less than 20% of the statistical uncertainty for any s-wave component larger than 1%.

6.2. Fit Strategy and preliminary results

The three fit methods described in Section 4.4 are applied to the $B_s^0 \to J/\psi\,\phi$ data using a simultaneous fit to the proper time unbiased and proper time biased data samples. The data is further split and fit simultaneously to tagged and untagged datasets. For the $\mathcal{C}$-fit analyses the background models in proper time, angles and $M(J/\psi\phi)$ are included in the fit and the signal and background fractions for each of the tagged, untagged, proper time unbiased and proper time biased components are allowed to float. The fit is performed in two stages, initially using only proper time unbiased data to ensure consistency between the three analyses. A simultaneous fit is then performed to both the proper time biased and proper time unbiased datasets. At each of these steps the s-wave component is first fixed to $|A_s|^2 = 0$, then floated to ensure that the different fitter implementations obtain the same result. At each stage the 2D profile likelihood contours are produced in the $\phi_s - \Delta\Gamma_s$ plane and the 1D profile likelihoods are also produced for all physics parameters. Before the final fit results were produced, both ϕ_s and $\Delta\Gamma_s$ were blinded by an unknown offset to prevent the onset of observer bias. The results presented here are after unblinding unless specified otherwise.

6.2.1. Results for proper time unbiased data

The proper time unbiased data sample consists of 11029 $B_s^0 \to J/\psi\,\phi$ candidates passing the selection described in Section 4.1.1, of which 7994 are untagged and 3033 are tagged. The invariant $M(J/\psi\phi)$ mass distribution of these candidates is shown in Figure 4.3a. The time and angular distributions are shown in Figure 6.5, in which the $\mathcal{C}$-fit projections are overlaid. The three fitters are in excellent agreement as indicated by the results for

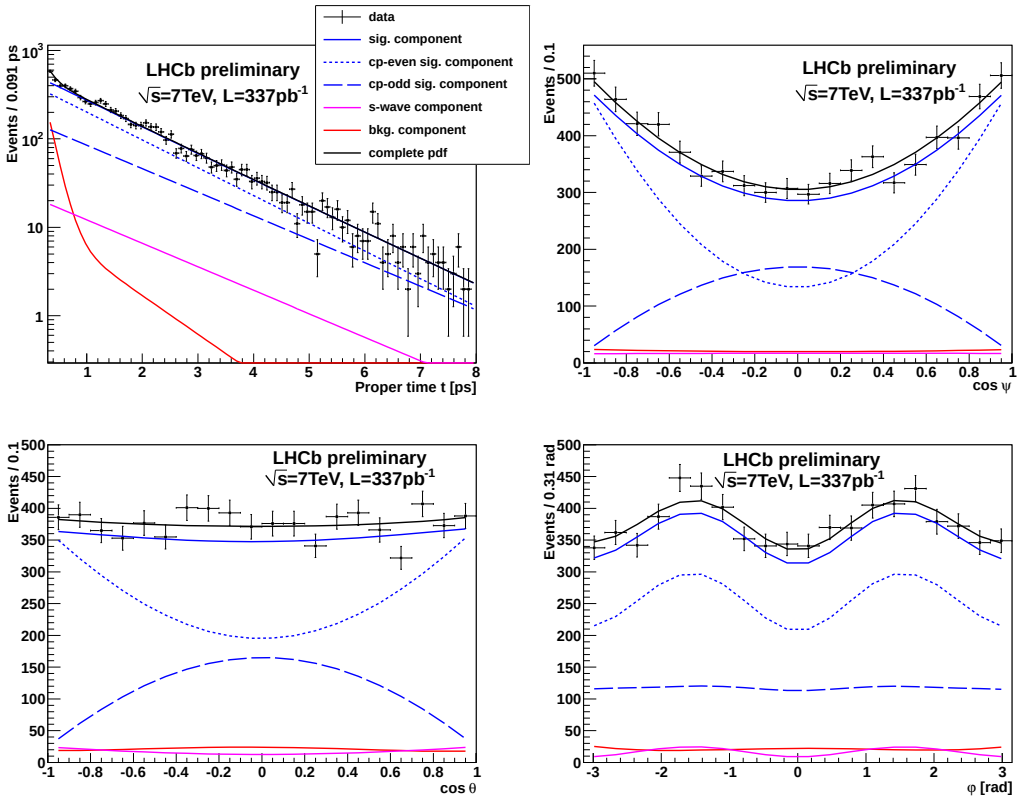
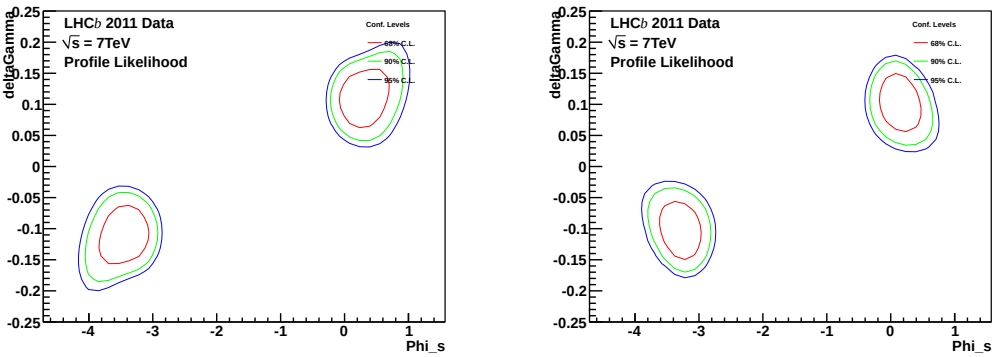


Figure 6.5.: Proper time and angular distributions for $B_s^0 \to J/\psi\,\phi$ candidates passing the proper time unbiased trigger. The projections of the fit are overlaid.



(a) $|A_s|^2 = 0.0$ fixed **(b)** $|A_s|^2$ free

Figure 6.6.: Profile likelihood scans in the $\phi_s - \Delta\Gamma_s$ plane using proper time unbiased data.

Parameter	$\mathcal{S}$-fit	E $\mathcal{C}$-fit	H $\mathcal{C}$-fit
Γ_s (ps^{-1})	0.6608 ± 0.0091	0.6611 ± 0.0093	0.6605 ± 0.0093
$\Delta\Gamma_s$ (ps^{-1})	0.105 ± 0.032	0.111 ± 0.033	0.110 ± 0.033
$\|A_\perp\|^2$	0.236 ± 0.016	0.235 ± 0.017	0.236 ± 0.016
$\|A_0\|^2$	0.495 ± 0.014	0.496 ± 0.015	0.496 ± 0.014
$\|A_s\|^2$	0.044 ± 0.017	0.042 ± 0.018	0.040 ± 0.017
$\delta_\parallel$	3.27 ± 0.20	3.25 ± 0.21	3.24 ± 0.20
$\delta_\perp$	2.81 ± 0.38	2.83 ± 0.40	2.81 ± 0.41
δ_s	2.93 ± 0.38	2.96 ± 0.40	2.94 ± 0.41
ϕ_s	0.113 ± 0.222	0.102 ± 0.220	0.099 ± 0.218
$m_{B_s^0}$ (MeV/c^2)	-	5365.5 ± 0.093	
σ_{m1} (MeV/c^2)	-	6.48 ± 0.07	
c (MeV/c^2)$^{-1}$	-	-0.00126 ± 0.00016	
f_{bkg1}	-	0.9921 ± 0.0009	
τ_{bkg1} (ps)	-	0.144 ± 0.006	
τ_{bkg2} (ps)	-	0.993 ± 0.060	
$f_{\text{sig}}^{\text{untagged}}$	-	0.683 ± 0.006	
$f_{\text{sig}}^{\text{tagged}}$	-	0.606 ± 0.010	

Table 6.3.: Fit results for the proper time unbiased dataset.

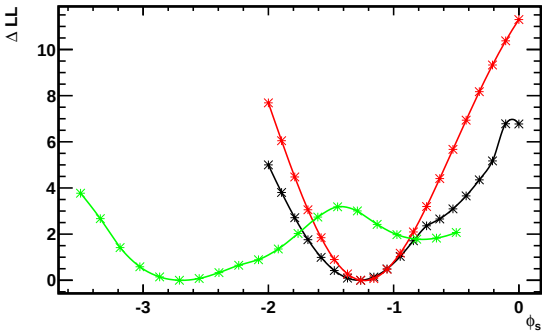


Figure 6.7.: Blinded profile likelihood of ϕ_s for proper time unbiased (Black), proper time biased (Green), and proper time unbiased + biased (Red) datasets. The asymmetry in the proper time unbiased profile is caused by the phase $\delta_{\parallel}$ transitioning from just above π to just below it at (blinded) $\phi_s = -0.7$.

all parameters tabulated in Table 6.3. Because the $\mathcal{S}$-fit does not require a background model the fit is simpler, using only the nine free physics parameters. The Edinburgh $\mathcal{C}$-fit takes 19 minutes to perform the fit, with 14 minutes taken for the $\mathcal{S}$-fit. The 2D profile likelihoods determined from the $\mathcal{S}$-fit are shown in Figure 6.6 for the case of ignoring a possible s-wave component, and allowing it to float. Here it can be seen that permitting the s-wave to float reduces the size of the contours for all confidence levels. The s-wave component is found to $4.4 \pm 1.7\%$.

6.2.2. Studies on the effect of including proper time biased data

The proper time biased dataset consists of 1754 additional candidates passing the selection described in Section 4.1.1, of which 1283 are untagged and 471 are tagged. The invariant $M(\mathrm{J}/\psi\phi)$ mass distribution of these candidates is shown in Figure 4.3b. Prior to unblinding some concern was raised about the effect of including the proper time biased dataset. While a decrease in coverage of the 2D contour was expected, the scale of the reduction appeared by-eye to be larger than the statistical reduction caused by a 15% increase in data would allow. The differences in 1D profile likelihoods for ϕ_s between the proper time unbiased-only and combined proper time unbiased + biased datasets as shown in Figure 6.7 further added to this concern. The proper time unbiased data sample deviates from a parabolic minimum at $-\Delta \ln \mathcal{L} = 2.5$ on the positive edge of the minimum. The proper time biased dataset has a different minimum, but when fit to simultaneously the profile becomes symmetric and close to the original proper time unbiased dataset minimum. The phase $\delta_{\parallel}$ transitions from just above π to just below

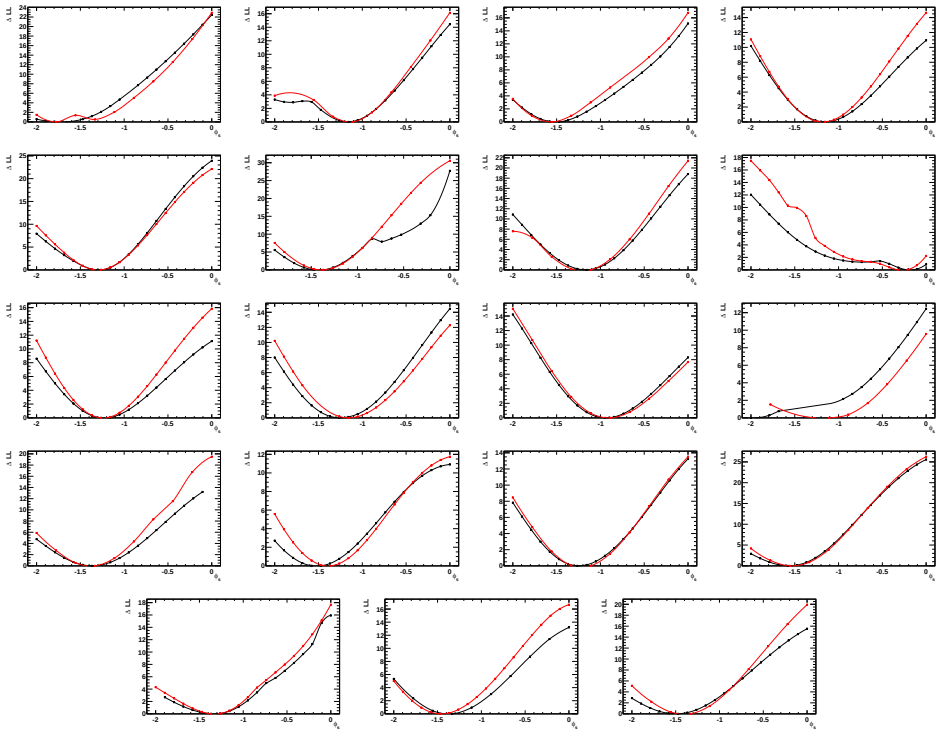


Figure 6.8.: Profile likelihood scans to toy datasets equivalent to the full dataset and proper time unbiased only dataset. The scans in black are to 86% of the data generated in the toy dataset, equivalent to proper time unbiased data only. The scans in red are to 100% of the toy dataset, and are equivalent to the full proper time unbiased + biased dataset.

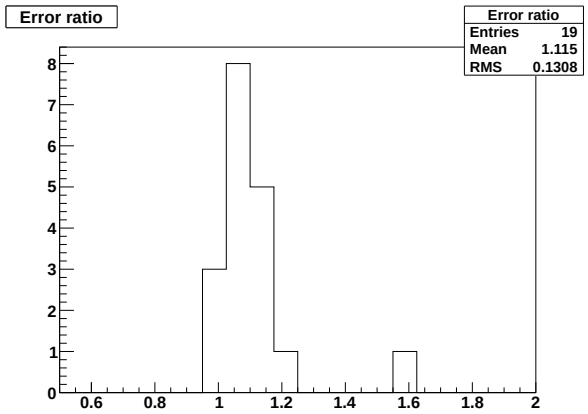


Figure 6.9.: A histogram of the ratio of ϕ_s uncertainties determined from fits to toy datasets of equivalent sizes to that of the proper time unbiased and proper time unbiased + biased datasets. The error ratio in data is 1.2. In this sample of toys 2 have larger error ratios.

it in the fit to the proper time unbiased dataset causing the $-\Delta \ln \mathcal{L}$ to broaden on the positive side of the minimum. In order to study this effect a set of 20 toy datasets equivalent in size to the full proper time unbiased + biased data sample were generated. A number of simplifications are made to the generation in order to speed up the procedure:

- Per-event mistag is replaced with a fixed average mistag of 36%.
- Events are generated without a proper time acceptance but the total number of events generated correspond to the total events in the combined proper time unbiased + biased dataset.
- Events are generated without angular acceptance.

The toy datasets are then fit to twice, and a profile likelihood scan of ϕ_s is made for both fits. The first fit is made to the full toy dataset and the second is made to 86% of it, simulating the effect of fitting to the full proper time unbiased + biased dataset and the proper time unbiased only dataset without having to include the proper time biased acceptance parameterisation. Comparing the results of these scans and the ratio of the uncertainties on ϕ_s can give some insight into whether or not the reduction in contour coverage is consistent with the addition of proper time biased data. One toy fails to fit, leaving 19 profile likelihood scan pairs which are presented in Figure 6.8. We also determine the ratio of uncertainties on the central value of ϕ_s returned by the fit. For data this ratio is $0.22/0.18 = 1.2$. By naïve scaling we expect this to be equivalent to the square-root of the ratio of sample sizes, eg: $\sqrt{1.0/0.86} = 1.1$. We determine this ratio for the 19 toys, the result of which is shown in Figure 6.9. Two of the 19 fits return an error ratio equal to or larger than that seen in data, with a mean error ratio of 1.1 ± 0.13. One of the 19 profile likelihood pairs in Figure 6.8 is particularly interesting as it exhibits very similar behaviour to that of data. Figure 6.10 shows this $-\Delta \ln \mathcal{L}$ pair and the value of $\delta_{\parallel}$ returned by the fit at each point on the $-\Delta \ln \mathcal{L}$ profile. The strong phase makes the same transition as observed in the fit to proper time unbiased data at the same value of ϕ_s. As with data, including the additional 14% of events from the toy dataset prevents this transition, leading to a parabolic minima once more. From this we conclude that the decrease in coverage of the likelihood contour was not purely statistical, but caused by the additional proper time biased data constraining $\delta_{\parallel}$ to remain above π.

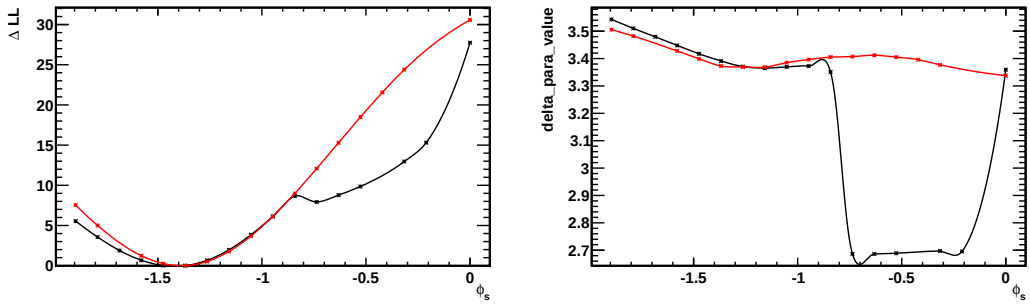
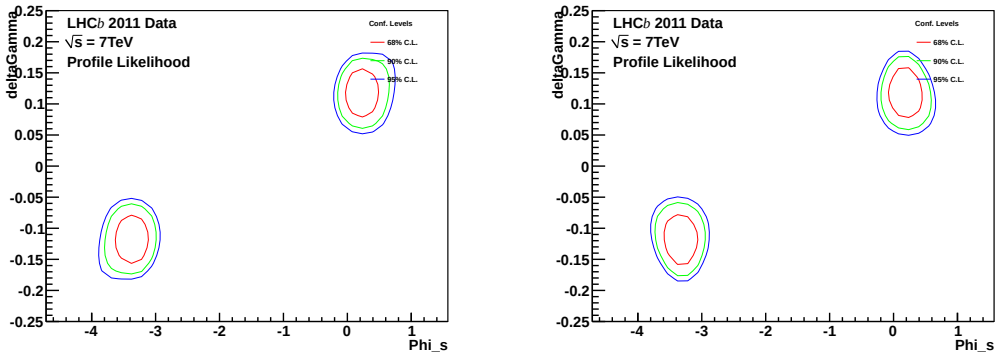


Figure 6.10.: Profile likelihoods and values of $\delta_{||}$ as a function of ϕ_s for one of the toy datasets described in the text. The fit to the full dataset is shown in red, and to the subset in black. $\delta_{||}$ transitions in the same way as observed in data, leading to a very similar scan to that of Figure 6.7.



(a) $|A_s|^2 = 0.0$ fixed **(b)** $|A_s|^2$ free

Figure 6.11.: Profile likelihood scans in the $\phi_s - \Delta\Gamma_s$ plane for the full fit to proper time unbiased + biased data.

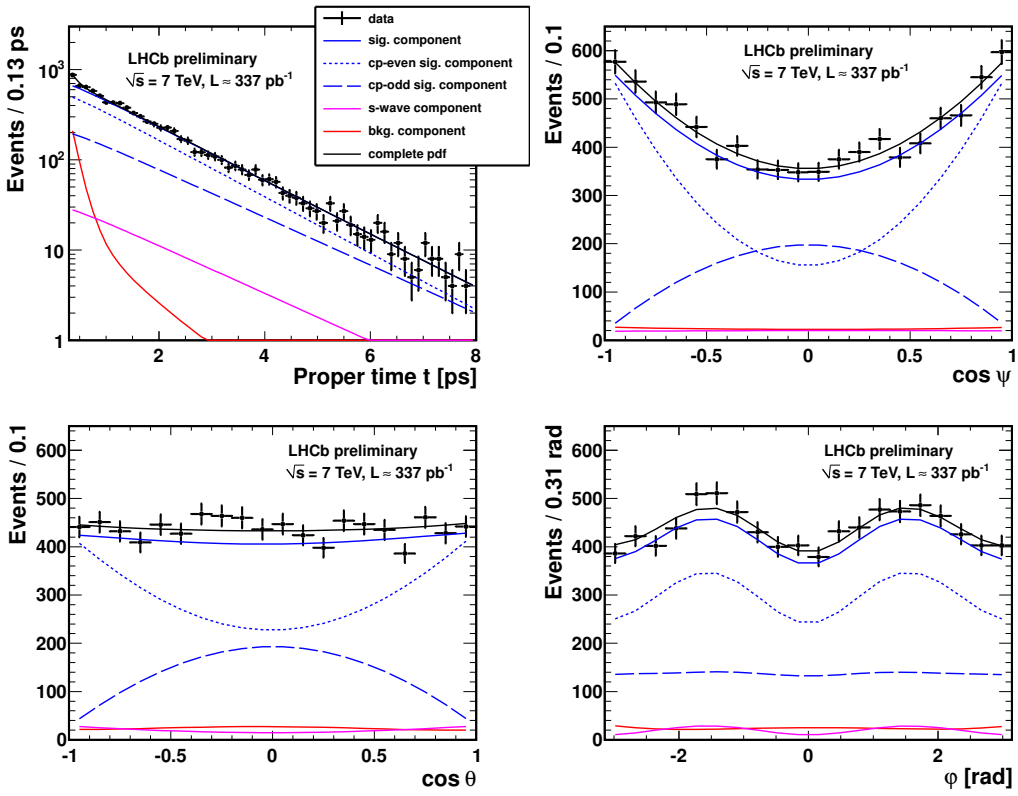


Figure 6.12.: Distributions of proper time and angles for the full fit to proper time unbiased + biased $B_s^0 \to J/\psi\,\phi$ data. The $\mathcal{C}$-fit projections are overlaid.

Parameter	$\mathcal{S}$-fit	E $\mathcal{C}$-fit	H $\mathcal{C}$-fit
Γ_s (ps^{-1})	0.6541 ± 0.0083	0.6535 ± 0.0086	0.6530 ± 0.0085
$\Delta\Gamma_s$ (ps^{-1})	0.118 ± 0.027	0.121 ± 0.028	0.120 ± 0.028
$\lvert A_\perp\rvert^2$	0.251 ± 0.015	0.250 ± 0.015	0.251 ± 0.015
$\lvert A_0\rvert^2$	0.518 ± 0.010	0.517 ± 0.011	0.516 ± 0.010
$\lvert A_s\rvert^2$	0.0	0.0	0.0
$\delta_\parallel$	3.32 ± 0.25	3.31 ± 0.27	3.28 ± 0.28
$\delta_\perp$	2.79 ± 0.60	2.86 ± 0.50	2.78 ± 0.66
ϕ_s (rad)	0.189 ± 0.177	0.179 ± 0.177	0.179 ± 0.178
Δ NLL	+7.39	+7.39	+6.40

Table 6.4.: Final fit results for the full proper time unbiased + biased dataset. These results are for the fit in which the s-wave contribution is ignored. The change in likelihood with respect to the results in which the s-wave is included is shown.

6.2.3. The full fit to proper time unbiased + biased data

After confirmation that the effect of including the proper time unbiased data was understood, the fits were checked for consistency and unblinded. The results presented in Table 6.4 and Table 6.5 are the final unblinded results with statistical uncertainties only for the case where an s-wave contribution is ignored and included respectively. Also shown is the difference in NLL between the two fits. The profile likelihood contours in the $\phi_s - \Delta\Gamma_s$ plane are shown in Figure 6.11 along with proper time and angular distributions in Figure 6.12, where a fit is overlaid. Consistency between the three analyses is excellent, with the largest difference in a single parameter being 25% of the statistical uncertainty, as indicated by the spanning difference in Table 6.5. The correlation matrix for the full fit is presented in Table 6.6, where linear correlations greater than 50% are shown in bold. It can be seen that the central physics parameters $\Delta\Gamma_s$, Γ_s and ϕ_s are not strongly correlated with each other, and that the amplitudes are correlated to $\Delta\Gamma_s$ as is expected. Figure 6.14 shows the Feldman-Cousins coverage corrected contours compared to the profile likelihood method for one of the two ambiguous solutions. It can be seen that the profile likelihood undercovers in comparison, but the overall effect is small. This is consistent with the 1D profile likelihoods being parabolic, which is true for all physics parameters except for $\delta_\parallel$ as evidenced in Figure 6.13. Because of the parabolic

Parameter	$\mathcal{S}$-fit	E $\mathcal{C}$-fit	H $\mathcal{C}$-fit	Sp.Diff.[%]
Γ_s (ps^{-1})	0.6581 ± 0.0081	0.6573 ± 0.0083	0.6566 ± 0.0083	18%
$\Delta\Gamma_s$ (ps^{-1})	0.119 ± 0.029	0.124 ± 0.029	0.123 ± 0.029	17%
$\|A_\perp\|^2$	0.238 ± 0.014	0.236 ± 0.015	0.237 ± 0.015	13%
$\|A_0\|^2$	0.498 ± 0.013	0.498 ± 0.013	0.497 ± 0.013	8%
$\|A_s\|^2$	0.046 ± 0.015	0.044 ± 0.016	0.042 ± 0.015	25%
$\delta_\parallel$	3.26 ± 0.21	3.24 ± 0.22	3.24 ± 0.21	10%
$\delta_\perp$	2.97 ± 0.34	2.97 ± 0.36	2.95 ± 0.37	6%
δ_s	3.00 ± 0.34	3.00 ± 0.36	2.98 ± 0.36	6%
ϕ_s (rad)	0.160 ± 0.183	0.151 ± 0.181	0.148 ± 0.180	7%
$m_{B_s^0}$ (MeV/c^2)	-	5365.4 ± 0.085		
σ_{m1} (MeV/c^2)	-	6.45 ± 0.07		
c (MeV/c^2)$^{-1}$	-	-0.00126 ± 0.00016		
$f_{\rm bkg1}$	-	0.9921 ± 0.0009		
$\tau_{\rm bkg,biased}$ (ps)	-	0.365 ± 0.018		
$\tau_{\rm bkg1}$ (ps)	-	0.145 ± 0.006		
$\tau_{\rm bkg2}$ (ps)	-	0.993 ± 0.060		
$f^{\rm untagged}_{\rm sig,unbiased}$	-	0.682 ± 0.006		
$f^{\rm tagged}_{\rm sig,unbiased}$	-	0.605 ± 0.010		
$f^{\rm untagged}_{\rm sig,biased}$	-	0.729 ± 0.014		
$f^{\rm tagged}_{\rm sig,biased}$	-	0.620 ± 0.024		

Table 6.5.: Fit results for the full simultaneous fit to proper time unbiased + biased data including an s-wave component. The "Spanning difference" is presented in the last column, defined as the largest difference between the three fitters as a percentage of the statistical uncertainty.

	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\lvert A_s\rvert^2$	$\delta_\parallel$	$\delta_\perp$	δ_s	Δm_s	ϕ_s
Γ_s	1	-0.29	0.27	-0.3	0.21	-0.066	-0.022	-0.024	-0.036	0.15
$\Delta\Gamma_s$		1	**-0.66**	**0.56**	-0.097	0.025	0.033	0.035	0.024	-0.093
$\lvert A_\perp\rvert^2$			1	-0.31	-0.22	-0.13	-0.02	-0.023	-0.048	0.14
$\lvert A_0\rvert^2$				1	**-0.62**	0.068	0.058	0.065	0.024	-0.11
$\lvert A_s\rvert^2$					1	-0.13	-0.083	-0.092	-0.0014	0.0049
$\delta_\parallel$						1	0.18	0.17	0.019	0.0034
$\delta_\perp$							1	**0.95**	0.41	0.069
δ_s								1	0.42	0.076
Δm_s									1	-0.2
ϕ_s										1

Table 6.6.: Correlation matrix for physics parameters in the full fit to proper time unbiased + biased data.

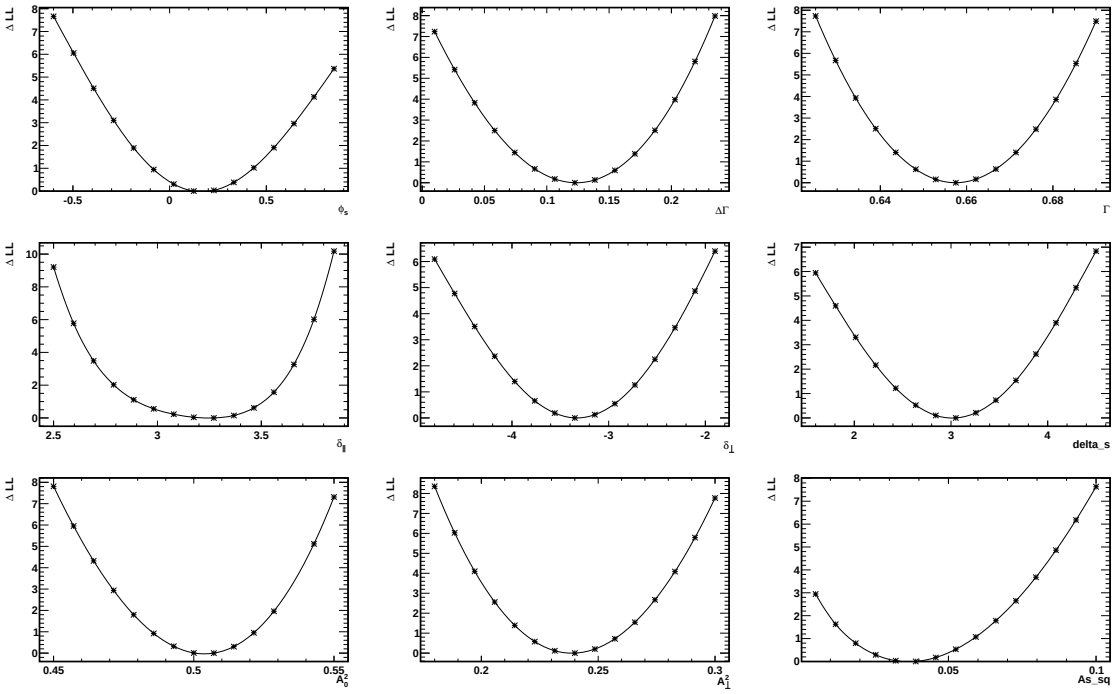


Figure 6.13.: Profile likelihoods as a function of each of the floated physics parameters in the full fit to proper time unbiased + biased data. Only one parameter, $\delta_\parallel$, exhibits non-parabolic behaviour about the minimum.

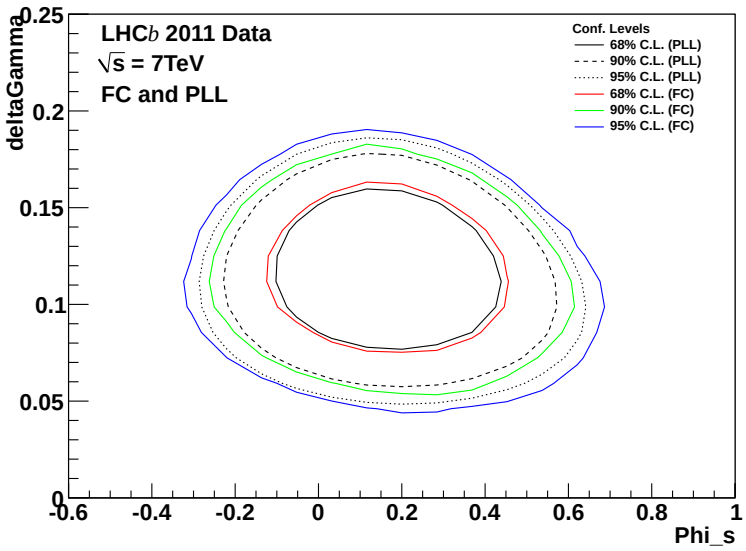


Figure 6.14.: Feldman-cousins coverage-corrected confidence limit contours for the $\Delta\Gamma_s > 0$ result. A 20×20 grid is used with 2000 toys generated per gridpoint. The contours in colour are the result of the Feldman-Cousins procedure, while those in black are the profile likelihood contours, which undercover.

nature of the other physics parameters the minos error estimates returned by MINUIT are considered to be good estimators of the statistical uncertainty on all parameters except for $\delta_\parallel$, for which we state a $\Delta \ln \mathcal{L} = 0.5$ interval of $\delta_\parallel \in [3.01, 3.42]$.

6.3. Systematic uncertainties

A number of systematic uncertainties are already accounted for in the fit by including Gaussian constraints, namely the B_s^0 mixing frequency Δm_s, the tagging calibration parameters p_0, p_1 and the proper time resolution parameters $\sigma_1, \sigma_2, \sigma_3, f_1, f_2$. The remaining systematic uncertainties are determined from a combination of toy studies and refitting the data with different parameterisations.

6.3.1. Uncertainty due to treatment of background

The systematic uncertainty due to treatment of background is determined from the differences between the fit results of the three analyses: The $\mathcal{S}$-fit does not model background explicitly, while the two $\mathcal{C}$-fit analyses treat background differently as described

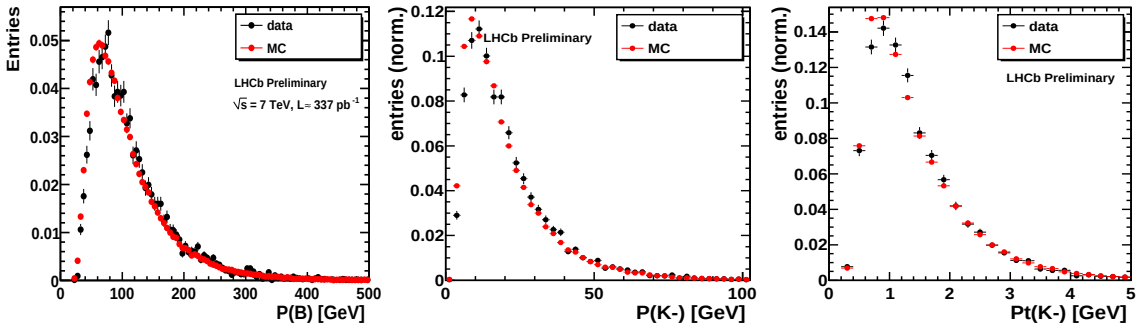


Figure 6.15.: B_s^0 and $K^\pm$ Momentum differences between sideband-subtracted data (black) and truth-matched Monte-Carlo (red).

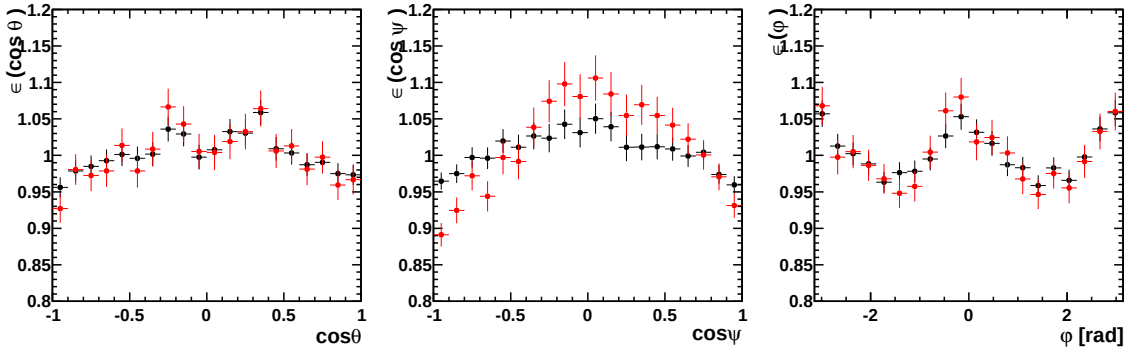


Figure 6.16.: Angular acceptance efficiency projections before and after reweighting with the $K^\pm$ momentum difference between data and Monte-Carlo. The nominal Monte-Carlo angular efficiencies are in black. In red are those after reweighting.

in Section 4.5. The uncertainty is taken as the largest difference in parameter value between the three fitters, divided by the statistical uncertainty on that parameter. These are listed in the last column of Table 6.5.

6.3.2. Uncertainty due to the determination of angular acceptance

The distributions of the angular observables are dependent upon the B_s^0 momentum and that of the final state particles, and there is some discrepancy between the momentum of candidates in data compared to those in Monte-Carlo. The data has harder momentum spectra than Monte-Carlo as is shown in Figure 6.15. To account for this discrepancy the Monte-Carlo sample is reweighted as a function of the $K^\pm$ and the B_s^0 momentum and the angular acceptance weights are recalculated. The acceptance efficiency projections before and after reweighting are shown in Figure 6.16. Toy studies are then performed in which 500 toys of 200,000 events each are generated with the momentum-corrected

Parameter	reweighting	MC stats
Γ_s	0.0010	0.00048
$\Delta\Gamma_s$	0.0030	0.00080
$\|A_\perp\|^2$	0.011	0.0028
$\|A_0\|^2$	0.030	0.0022
$\|A_s\|^2$	0.014	0.0033
$\delta_\parallel$	0.11	0.02
$\delta_\perp$	0.12	0.023
δ_s	0.12	0.017
ϕ_s	0.024	0.025

Table 6.7.: Systematic uncertainties due to the determination of the angular acceptance weights. The first column is due to Monte-Carlo-data momentum differences while the second is due to the Monte-Carlo sample size.

acceptance weights but fit using the nominal acceptance weights. The bias imparted to physics parameters in the toy datasets as a result of neglecting the reweighting is then taken as a systematic uncertainty. Determination of the acceptance weights also imparts a statistical uncertainty based on the Monte-Carlo sample size used. This uncertainty is propagated through to the physics parameters by refitting to data using acceptance weights randomly varied within their uncertainties a large number of times. The width of the variation on the physics parameter is then taken as the systematic due to the statistical uncertainty on the weights. Both the systematic and statistical uncertainties arising from the acceptance weighting procedure are presented in Table 6.7.

6.3.3. Systematic uncertainties due to the proper time acceptance

The systematic uncertainty due to the calculation of the proper time acceptance for proper time biased candidates is assigned by refitting to the data using the parameters determined from the ratio method rather than the nominal $\mathcal{S}$-plot method. The differences in parameterisation are shown in Table 5.3. For the upper lifetime acceptance we assign an uncertainty to Γ_s equal to half the value of the correction applied, $\beta = -0.00157$ ps^{-1}, $\sigma(\beta) = 0.0008$ ps^{-1}.

6.3.4. z-scale and momentum scale

The uncertainty on the z and momentum scales is at most 0.001 each, which will affect $\Delta\Gamma_s$ and Γ_s by 0.1% [99]. For $\Delta\Gamma_s$ this is equivalent to about 2 fs and for Γ_s about 1 fs, or 1% and 12% of the statistical uncertainty.

6.3.5. Systematics due to the treatment of the s-wave component

The fraction of s-wave measured by the fit is 4.6%, with a statistical significance of 3.1σ. Repeating the study in Section 6.1 using as input to the toys the measured physics parameters and signal yields we find the bias on the amplitudes which we use as systematics to the measured values. These are: $\sigma(|\mathrm{A}_\perp|^2) = 0.002$, $\sigma(|\mathrm{A}_0|^2) = 0.003$ and $\sigma(|\mathrm{A}_s|^2) = 0.010$.

6.3.6. Systematic uncertainty due to nuisance $\mathcal{CP}$ asymmetries

A number of detector-imposed "fake" and physics-imposed "real" $\mathcal{CP}$ asymmetries can enter into the measurement of ϕ_s that cannot be directly measured with the present statistics. The asymmetries considered are:

- $\mathrm{B_s^0}/\overline{\mathrm{B}}_\mathrm{s}^0$ production asymmetry.
- Differences in $\mathrm{B_s^0}/\overline{\mathrm{B}}_\mathrm{s}^0$ tagging efficiency and mistag.
- Nonzero $\mathcal{CP}$ violation in mixing.
- Nonzero $\mathcal{CP}$ violation in decay.

Detailed studies of these asymmetries are performed using toys described in detail in [101], a summary of which is presented here. The recipe for estimating systematic effects is the same for all studies: High statistics data samples are generated using a model that includes the nuisance asymmetries for a range of values of ϕ_s. The input value for $\Delta\Gamma_s$ is always 0.05 ps^{-1}. Generated data are fitted with a model that does not include any nuisance asymmetries. Systematic deviations of fit parameters from their input values are used as estimates of the systematic uncertainty. The effect of a production asymmetry is estimated by generating events at a conservative asymmetry of $\pm 10\%$. The systematic effect depends on the input value of ϕ_s, but the bias on ϕ_s does not vanish at zero. The maximum deviations in the $\pm\ 3\sigma$ range of ϕ_s are 0.01

in ϕ_s, 0.001 ps^{-1} in $\Delta\Gamma_s$ and 0.0002 ps^{-1} in Γ_s. Systematic errors from non-zero $\mathcal{CP}$ violation in mixing and/or decay are larger than for the production asymmetry. Events were generated with the squared magnitude of the $\mathcal{CP}$ violation parameter λ set to 0.95 and to 1.05. $|\lambda|^2$ goes to one in the absence of $\mathcal{CP}$ violation in mixing and/or decay. The maximum systematic effect in ϕ_s is 0.03. The deviations in $\Delta\Gamma_s$ and in Γ_s depend strongly on the value of ϕ_s. For $\Delta\Gamma_s$, the effect varies from 0 at $\phi_s = 0$ to 0.006 ps^{-1} at a 3 sigma upward fluctuation of $\phi_s = 0.01$. For Γ_s the effect oscillates between approximately ± 0.001 ps^{-1} for nonzero ϕ_s and vanishes for $\phi_s = 0$. Both effects (production asymmetry and $\mathcal{CP}$ in mixing & decay) show dependencies on the ϕ_s value. As these uncertainties are small compared to the other uncertainties we quote the maximum values in a $\pm\ 3\sigma$ range as absolute systematic uncertainty.

6.3.7. Combined systematic uncertainties

The total systematic uncertainties are listed in Table 6.8 along with the statistical uncertainty. The largest systematic uncertainty on ϕ_s and $\Delta\Gamma_s$ is due to the angular acceptance weighting procedure, but the total uncertainty is still dominated by the size of the data sample. The 2D profile likelihood contour for the full dataset adjusted to include the systematic uncertainties is shown in Figure 6.17.

6.4. Final result

The final results with 0.37 fb^{-1} of LHCb data using both proper time unbiased + biased data is summarised in Table 6.9. The central physics parameter $\mathcal{S}$-fit point-estimates for the positive $\Delta\Gamma_s$ solution are:

$$\phi_s = 0.16\ \pm 0.18\ (\text{stat}) \pm 0.06\ (\text{syst})\ \text{rad}, \tag{6.2}$$

$$\Gamma_s = 0.658\ \pm 0.008\ (\text{stat}) \pm 0.008\ (\text{syst})\ \text{ps}^{-1}, \tag{6.3}$$

$$\Delta\Gamma_s = 0.119\ \pm 0.029\ (\text{stat}) \pm 0.011\ (\text{syst})\ \text{ps}^{-1}, \tag{6.4}$$

The second solution is related to this one by the transform $(\phi_s, \Delta\Gamma_s) \mapsto (\pi - \phi_s, -\Delta\Gamma_s)$, and both solutions are shown including systematic uncertainties in the $\phi_s - \Delta\Gamma_s$ plane along with the Standard Model prediction in Figure 6.17. The $\Delta\Gamma_s > 0$ results are in

Source	Γ_s [ps^{-1}]	$\Delta\Gamma_s$ [ps^{-1}]	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\lvert A_s\rvert^2$	$\delta_\parallel$	$\delta_\perp$	δ_s	ϕ_s
Stat. uncertainty	0.0083	0.029	0.015	0.013	0.015	0.23	0.37	0.36	0.18
Bkg., Ang., Time. method	0.0015	0.005	0.002	0.001	0.004	0.02	0.02	0.02	0.012
(of which) Bkg. model	0.0003	0.002	0.00	0.002	0.002	0.01	0.01	0.01	0.006
angular acc. re-weighting	0.0010	0.008	0.011	0.030	0.014	0.11	0.12	0.12	0.043
angular acc. statistical	0.0005	0.001	0.003	0.003	0.003	0.02	0.02	0.02	0.025
t acc. model	0.0007	0.002	0.001	0.001	0.000	0.00	0.00	0.00	0.001
β	0.008	-	-	-	-	-	-	-	-
z + p scale	0.001	0.000	-	-	-	-	-	-	-
s-wave	-	-	0.002	0.003	0.010	-	-	-	-
prod. asym	< 0.0002	< 0.001	-	-	-	-	-	-	< 0.01
CPV mix+decay	< 0.001	< 0.006	-	-	-	-	-	-	< 0.03
quad. sum of syst.	0.0084	0.011	0.012	0.030	0.018	0.11	0.12	0.12	0.06
tot. uncertainties	0.0118	0.032	0.019	0.033	0.023	0.25	0.39	0.38	0.19

Table 6.8.: Summary of systematic uncertainties for the final result. The different contributions are described in the text. The penultimate column lists the quadrature sum of all systematic uncertainties where for upper limits the maximum value is used in the sum.

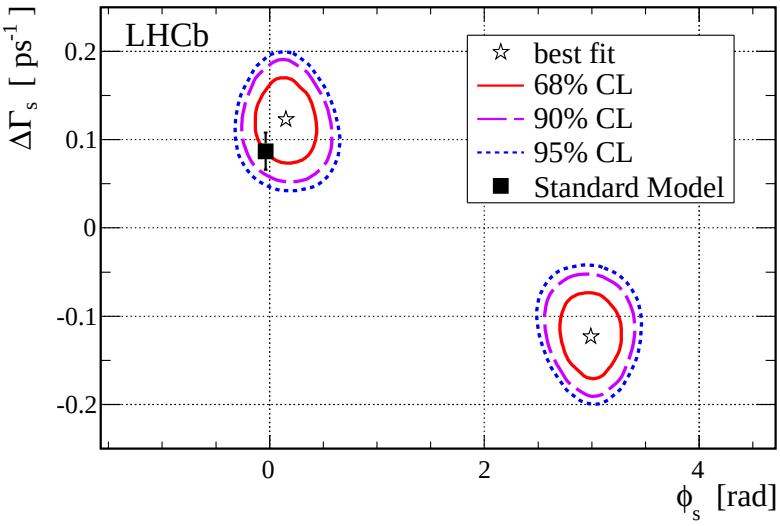


Figure 6.17.: Profile likelihood including systematic uncertainties in the $\Delta\Gamma_s - \phi_s$ plane.

parameter	value	$\sigma_{\text{stat.}}$	$\sigma_{\text{syst.}}$
Γ_s [ps^{-1}]	0.658	0.008	0.008
$\Delta\Gamma_s$ [ps^{-1}]	0.119	0.029	0.011
$\lvert A_\perp \rvert^2$	0.238	0.014	0.012
$\lvert A_0 \rvert^2$	0.498	0.013	0.030
$\lvert A_s \rvert^2$	0.046	0.015	0.018
$\delta_\perp$ [rad]	2.97	0.34	0.12
$\delta_\parallel$ [rad]	$\in [3.01, 3.42]$		0.12
δ_s [rad]	3.00	0.34	0.12
ϕ_s [rad]	0.16	0.18	0.06

Table 6.9.: Final results for the physics parameters with their statistical and systematic uncertainties.

good agreement with the Standard Model predictions [102, 103, 104, 105]:

$$\phi_s^{\mathrm{SM}} = -0.036 \pm 0.002 \text{ rad}$$
$$\Delta\Gamma_s{}^{\mathrm{SM}} = 0.082 \pm 0.021 \text{ ps}^{-1}$$

This is the most precise single measurement of ϕ_s, $\Delta\Gamma_s$ and Γ_s, and provides the first direct evidence of a nonzero $\Delta\Gamma_s$. We additionally measure the s-wave fraction in the $M(\mathrm{K^+K^-})$ range of $M(\phi) \pm 12$ MeV/c^2 to be $|\mathrm{A}_s|^2 = 4.6 \pm 1.5 \pm 1.8\%$. The $\mathcal{S}$-fit technique, which does not require modelling of the background component of the fit, produces results that are in excellent agreement with two different fit strategies that explicitly model the background.

6.5. Interpretation

The ϕ_s measurement presented in this thesis has implications for a number of New Physics (NP) scenarios. Figure 6.18 shows the complex plane defined by the parameter Δ_s which parameterises the deviation from a Standard Model value of the diagonal terms of the mass matrix M_{12} responsible for mixing [106]. In the absence of New Physics entering the mixing diagram $\Re\mathfrak{e}\,\Delta_s = 1$ and $\Im\mathfrak{m}\Delta_s = 0$. The LHCb constraint on ϕ_s greatly reduces the range over which a NP scenario can enter quark mixing, such as a new particle with mass $M > M_W$. A number of Minimally Supersymmetric Standard Model (MSSM) variants are also constrained as shown in Figure 6.19 [107] in combination with the LHCb BR($\mathrm{B_s^0} \to \mu^+\mu^-$) measurement [108]. While at the present sensitivity it is not possible to exclude physics beyond the Standard Model, any new physics entering the mixing of $\mathrm{B_s^0}$ mesons must be small in order to be compatible with the measurement presented in this thesis.

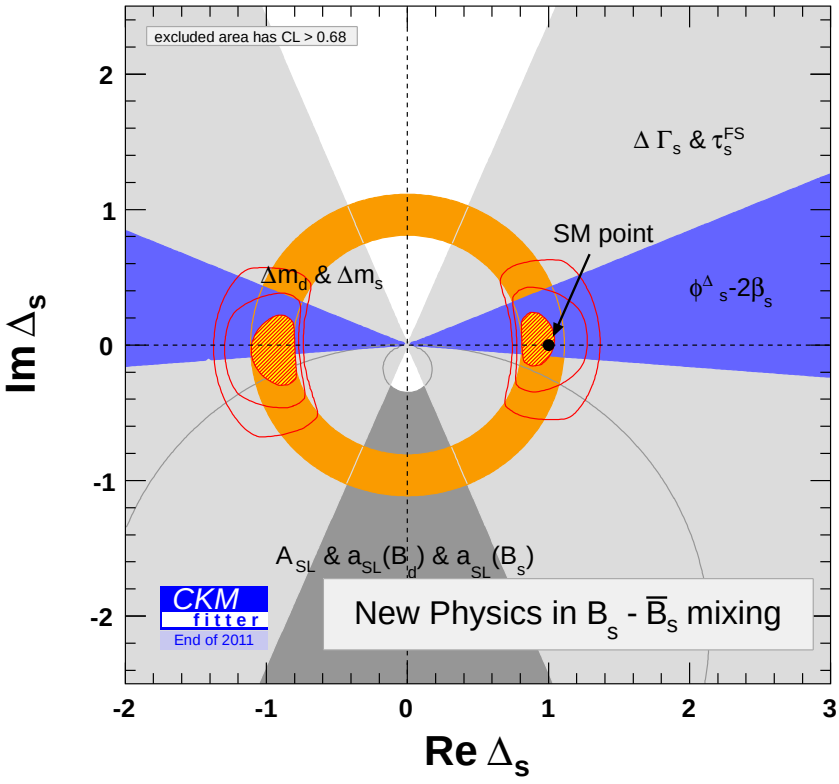


Figure 6.18.: Constraints on new physics in mixing including the LHCb ϕ_s measurement. The constraint is shown in the $\Re\mathfrak{e}\,\Delta_s - \Im\mathfrak{m}\Delta_s$ plane where Δ_s is defined in equation (1.60) [106].

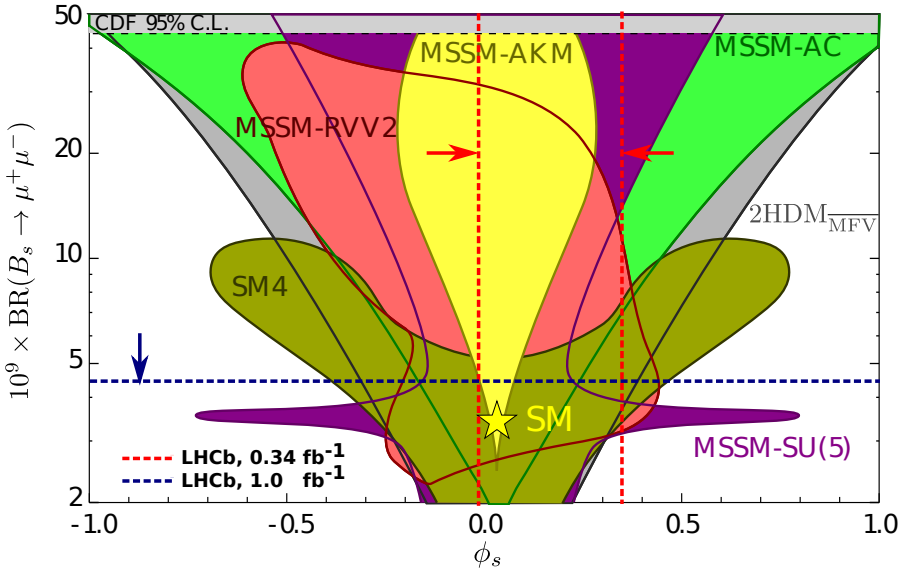


Figure 6.19.: Constraints on several MSSM models imposed by LHCb ϕ_s and $B_s^0 \to \mu^+\mu^-$ measurements [107]. The blue dashed line indicates the upper limit on $BR(B_s^0 \to \mu^+\mu^-)$ with 1 fb^{-1} [108]. The red dashed lines indicate the 1σ uncertainty on the ϕ_s measurement presented in this thesis.

Chapter 7.

Conclusions

"There is nothing new to be discovered in physics now. All that remains is more and more precise measurement"

— Disputed, commonly attributed to William Thomson, 1st Baron Kelvin

This thesis has presented three contributions to the body of work undertaken by LHCb. In Chapter 2.5 the time alignment procedure for the RICH detectors of LHCb was presented. Chapter 3 presents a measurement of the $D_s^{\pm}$ and $D^{\pm}$ cross-sections and their ratios with 1.81 nb^{-1} of proton-proton collisions at the LHC, and I present in Chapters 4, 5 and 6 a measurement of the $\mathcal{CP}$-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\,\phi$ with 0.37 fb^{-1} of proton-proton collisions at the LHC. Here I will comment upon each of these contributions, progress since the writing of this thesis, and possible further studies.

7.1. Time Alignment

The time alignment analysis described in this thesis is used at the start of each new running period of the LHCb experiment after shutdown periods and after hardware is changed. It is used both as a diagnostic to determine the inter-alignment of HPDs on an L0 board and to apply and confirm alignment of the RICH detectors to the global timing of LHCb. It consistently allows alignment to 1 ns resolution. The alignment software is expected to continue being used throughout the lifetime of the LHCb detector, and will

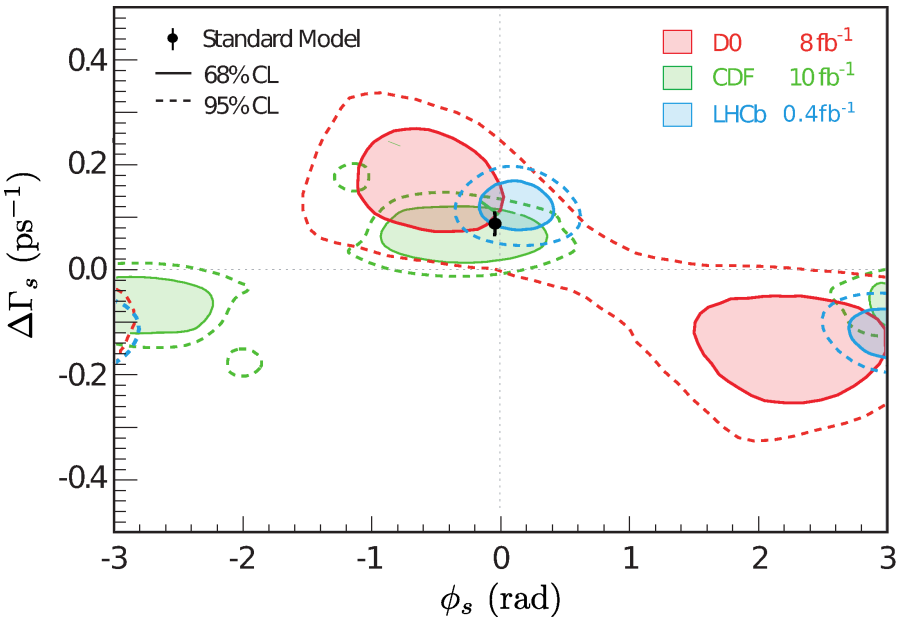


Figure 7.1.: Comparison of the current ϕ_s measurements. The LHCb result described in this thesis and [100] is shown in blue. The CDF [15] and DØ [14] results with their full datasets are shown in green and red respectively.

be modified if necessary for use after the LHCb upgrade in which the present photon detectors will be replaced.

7.2. Charm Cross-sections

The Charm cross-sections and cross-section ratio presented in this thesis were presented as one of the first LHCb results at ICHEP 2010 [109]. They are presently being updated for publication with 14 nb^{-1}. In combination with cross-sections of $\text{D}^{*\pm}$ and D^0 mesons and the $\Lambda_c^{\pm}$ baryon a prompt $\text{pp} \to \text{c}X$ cross-section is intended.

7.3. ϕ_s in $B_s^0 \to J/\psi\,\phi$

The measurement of ϕ_s presented in this thesis is now published [100]. It is the most sensitive measurement of ϕ_s. In Figure 7.1 this measurement is compared to the final results published by CDF [15] and DØ [14] previously shown in Figure 1.7. While this result marks a significant improvement to the status quo, the point-estimate uncertainty of $\phi_s = 0.16 \pm 0.18\ (\text{stat}) \pm 0.06\ (\text{syst})$ rad is still two orders of magnitude short of the

uncertainty on the Standard Model prediction of $\phi_s^{\rm SM} = -0.036 \pm 0.002$ rad. A number of improvements to the analysis are foreseen: In addition to updating the measurement with a larger dataset, the inclusion of the same-side Kaon tagger described in Section 5.3 will result in an improvement in tagging performance. There are also plans to simultaneously fit to additional $\mathrm{b} \to \mathrm{c\bar{c}s}$ decays such as $\mathrm{B_s^0} \to \mathrm{J/\psi f_0}(980)$ and $\mathrm{B_s^0} \to \mathrm{J/\psi}\eta, \eta'$.

An additional analysis using the same 0.37 fb^{-1} has been performed during the writing of this thesis that determines the sign of $\Delta\Gamma_s$, completely resolving the two fold ambiguity [110]. This method uses a similar analysis technique to the one presented in this thesis and measures the phase difference $\delta_s - \delta_\perp$ in bins of $M(\mathrm{K^+K^-})$ over the ϕ resonance to obtain a physical solution. The analysis finds that $\Delta\Gamma_s$ is positive, concluding that the mass hierarchy in the $\mathrm{B_s^0}$ system is the same as that of the $\mathrm{K^0}$ system.

The ϕ_s result presented in this thesis has been updated to the full 1 fb^{-1} dataset collected by LHCb throughout the 2011 running period [111]. This preliminary result for the Winter 2011 conference period implements only minor changes from those described in this thesis, and will be updated to include some of the modifications described here prior to publication.

Appendix A.

Charm Cross-Section Efficiency tables

This appendix contains the tables in y, p_{T} for $\mathrm{D_s^{\pm}}$, $\mathrm{D^{\pm}}$ candidates used to calculate efficiencies for extrapolation of prompt yields back to a cross-section.

A.1. Acceptance Efficiencies

p_{T} (MeV/c)	Generated $\mathrm{D_s^{\pm}}$ N_{gen}	$\mathrm{D_s^{\pm}}$ in acceptance N_{acc}	Acceptance efficiency $\epsilon_{acc} = \frac{N_{acc}}{N_{gen}}$
$(0, 1000)$	89683 ± 299	72022 ± 268	0.8031 ± 0.0013
$(1000, 2000)$	105545 ± 325	92489 ± 304	0.8763 ± 0.0010
$(2000, 3000)$	59635 ± 244	56015 ± 237	0.9393 ± 0.0010
$(3000, 4000)$	31137 ± 176	30075 ± 173	0.9659 ± 0.0010
$(4000, 5000)$	16165 ± 127	15791 ± 126	0.9769 ± 0.0012
$(5000, 6000)$	8476 ± 92	8344 ± 91	0.9844 ± 0.0013
$(6000, 7000)$	4780 ± 69	4736 ± 69	0.9908 ± 0.0014
$(7000, 8000)$	2768 ± 53	2750 ± 52	0.9935 ± 0.0015

Table A.1.: Prompt $\mathrm{D_s^{\pm}}$, Monte-Carlo yields at generator level and acceptance efficiency in bins of p_{T}, integrated over y on the range (2.0,4.5).

y	Generated $\mathrm{D_s^{\pm}}$ N_{gen}	$\mathrm{D_s^{\pm}}$ in acceptance N_{acc}	Acceptance efficiency $\epsilon_{acc} = \frac{N_{acc}}{N_{gen}}$
$(2, 2.5)$	87248 ± 295	75014 ± 274	0.8598 ± 0.0012
$(2.5, 3)$	77374 ± 278	73184 ± 271	0.9458 ± 0.0008
$(3, 3.5)$	65539 ± 256	62614 ± 250	0.9554 ± 0.0008
$(3.5, 4)$	52004 ± 228	46381 ± 215	0.8919 ± 0.0014
$(4, 4.5)$	39720 ± 199	28707 ± 169	0.7227 ± 0.0022

Table A.2.: Prompt $\mathrm{D_s^{\pm}}$, Monte-Carlo yields at generator level and acceptance efficiency in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c.

p_{T} (MeV/c)	Generated $\mathrm{D^{\pm}}$ N_{gen}	$\mathrm{D^{\pm}}$ in acceptance N_{acc}	Acceptance efficiency $\epsilon_{acc} = \frac{N_{acc}}{N_{gen}}$
$(0, 1000)$	96089 ± 310	76358 ± 276	0.7947 ± 0.0013
$(1000, 2000)$	107262 ± 328	95076 ± 308	0.8864 ± 0.0010
$(2000, 3000)$	57510 ± 240	54590 ± 234	0.9492 ± 0.0009
$(3000, 4000)$	28717 ± 169	27885 ± 167	0.9710 ± 0.0010
$(4000, 5000)$	14904 ± 122	14620 ± 121	0.9809 ± 0.0011
$(5000, 6000)$	7815 ± 88	7721 ± 88	0.9880 ± 0.0012
$(6000, 7000)$	4297 ± 66	4253 ± 65	0.9898 ± 0.0015
$(7000, 8000)$	2509 ± 50	2495 ± 50	0.9944 ± 0.0015

Table A.3.: Prompt $\mathrm{D^{\pm}}$, Monte-Carlo yields at generator level and acceptance efficiency in bins of p_{T}, integrated over y on the range (2.0,4.5).

y	Generated $\mathrm{D^{\pm}}$ N_{gen}	$\mathrm{D^{\pm}}$ in acceptance N_{acc}	Acceptance efficiency $\epsilon_{acc} = \frac{N_{acc}}{N_{gen}}$
$(2, 2.5)$	87155 ± 295	75043 ± 274	0.8610 ± 0.0012
$(2.5, 3)$	77220 ± 278	73469 ± 271	0.9514 ± 0.0008
$(3, 3.5)$	64891 ± 255	61932 ± 249	0.9544 ± 0.0008
$(3.5, 4)$	53031 ± 230	47033 ± 217	0.8869 ± 0.0014
$(4, 4.5)$	40108 ± 200	28810 ± 170	0.7183 ± 0.0022

Table A.4.: Prompt $\mathrm{D^{\pm}}$, Monte-Carlo yields at generator level and acceptance efficiency in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c.

A.2. Selection Efficiency without PID

p_{T} (MeV/c)	$\mathrm{D_s^{\pm}}$ in acceptance N_{acc}	Selected $\mathrm{D_s^{\pm}}$ $N_{sel'}$	Selection Efficiency $\epsilon_{sel'} = \frac{N_{sel'}}{N_{acc}}$
$(0, 1000)$	39739 ± 199	638 ± 25	0.0161 ± 0.0006
$(1000, 2000)$	49824 ± 223	1513 ± 39	0.0304 ± 0.0008
$(2000, 3000)$	29789 ± 172	1778 ± 42	0.0597 ± 0.0014
$(3000, 4000)$	15471 ± 124	1407 ± 38	0.0909 ± 0.0023
$(4000, 5000)$	8209 ± 91	977 ± 31	0.1190 ± 0.0036
$(5000, 6000)$	4357 ± 66	590 ± 24	0.1354 ± 0.0052
$(6000, 7000)$	2404 ± 49	372 ± 19	0.1547 ± 0.0074
$(7000, 8000)$	1422 ± 38	239 ± 15	0.1681 ± 0.0099

Table A.5.: Prompt $\mathrm{D_s^{\pm}}$, Monte-Carlo yields at selection level and selection efficiency in bins of p_{T}, integrated over y on the range (2.0,4.5), excluding particle ID efficiency.

y	$\mathrm{D_s^{\pm}}$ in acceptance N_{acc}	Selected $\mathrm{D_s^{\pm}}$ $N_{sel'}$	Selection efficiency, ϵ_{sel} $\epsilon_{sel'} = \frac{N_{sel'}}{N_{acc}}$
$(2, 2.5)$	38857 ± 197	1038 ± 32	0.0267 ± 0.0008
$(2.5, 3)$	38265 ± 196	2289 ± 48	0.0598 ± 0.0012
$(3, 3.5)$	33213 ± 182	2275 ± 48	0.0685 ± 0.0014
$(3.5, 4)$	25131 ± 159	1429 ± 38	0.0569 ± 0.0015
$(4, 4.5)$	15749 ± 125	483 ± 22	0.0307 ± 0.0014

Table A.6.: Prompt $\mathrm{D_s^{\pm}}$, Monte-Carlo yields at selection level and selection efficiency in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c, excluding particle ID efficiency.

p_{T} (MeV/c)	$\mathrm{D}^{\pm}$ in acceptance N_{acc}	Selected $\mathrm{D}^{\pm}$ $N_{sel'}$	Selection Efficiency, ϵ_{sel} $\epsilon_{sel'} = \frac{N_{sel'}}{N_{acc}}$
$(0, 1000)$	19450 ± 139	816 ± 29	0.0420 ± 0.0014
$(1000, 2000)$	23631 ± 153	1707 ± 41	0.0722 ± 0.0017
$(2000, 3000)$	13295 ± 116	1607 ± 40	0.1209 ± 0.0028
$(3000, 4000)$	6901 ± 83	1196 ± 35	0.1733 ± 0.0046
$(4000, 5000)$	3445 ± 59	666 ± 26	0.1933 ± 0.0067
$(5000, 6000)$	1838 ± 43	380 ± 19	0.2067 ± 0.0094
$(6000, 7000)$	1044 ± 32	248 ± 16	0.2375 ± 0.0132
$(7000, 8000)$	559 ± 24	146 ± 12	0.2612 ± 0.0186

Table A.7.: Prompt $\mathrm{D}^{\pm}$, Monte-Carlo yields at selection level and selection efficiency in bins of p_{T}, integrated over y on the range (2.0,4.5), excluding particle ID efficiency.

y	$\mathrm{D}^{\pm}$ in acceptance N_{acc}	Selected $\mathrm{D}^{\pm}$ $N_{sel'}$	Selection efficiency, ϵ_{sel} $\epsilon_{sel'} = \frac{N_{sel'}}{N_{acc}}$
$(2, 2.5)$	17983 ± 134	903 ± 30	0.0502 ± 0.0016
$(2.5, 3)$	17885 ± 134	2062 ± 45	0.1153 ± 0.0024
$(3, 3.5)$	15442 ± 124	2072 ± 46	0.1342 ± 0.0027
$(3.5, 4)$	11630 ± 108	1317 ± 36	0.1132 ± 0.0029
$(4, 4.5)$	7223 ± 85	412 ± 21	0.0570 ± 0.0027

Table A.8.: Prompt $\mathrm{D}^{\pm}$, Monte-Carlo yields at selection level and selection efficiency in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c, excluding particle ID efficiency.

A.3. Particle ID calibration and efficiencies

$K^{\pm}\Delta_{LL}(K-\pi)>9$	p_T (MeV/c)		
η	(0,1000)	(1000,2000)	>2000
<2.0	0.0 ± 0.0	0.9543 ± 0.0499	0.9764 ± 0.0544
(2.0,2.5)	0.8202 ± 0.0099	0.9671 ± 0.0084	0.9968 ± 0.0183
(2.5,3.0)	0.8805 ± 0.0055	0.9568 ± 0.0069	0.9614 ± 0.0175
(3.0,3.5)	0.9163 ± 0.0045	0.9764 ± 0.0067	0.9207 ± 0.0181
(3.5,4.0)	0.9622 ± 0.0049	0.9393 ± 0.0072	0.7600 ± 0.0219
(4.0,4.5)	0.7725 ± 0.0056	0.7317 ± 0.0096	0.3711 ± 0.0325
(4.5,5.0)	0.7738 ± 0.0109	0.4045 ± 0.0185	0.0 ± 0.0
>5.0	0.4859 ± 0.1238	0.0 ± 0.0	0.0 ± 0.0

Table A.9.: $K^{\pm}$ PID efficiencies in p_T,η determined from signal $\phi\to K^+K^-$ candidates in data.

$\pi^{\pm}\Delta_{LL}(\pi-K)>-2$	p_T (MeV/c)		
η	(0,1000)	(1000,2000)	>2000
<2.0	0.6322 ± 0.0540	0.0 ± 0.0	0.0 ± 0.0
(2.0,2.5)	0.6793 ± 0.0052	0.9314 ± 0.0132	0.9586 ± 0.0393
(2.5,3.0)	0.7286 ± 0.0055	0.9204 ± 0.0095	0.9748 ± 0.0316
(3.0,3.5)	0.7978 ± 0.0020	0.9622 ± 0.0104	0.9674 ± 0.0380
(3.5,4.0)	0.8779 ± 0.0020	0.9620 ± 0.0099	0.9072 ± 0.0335
(4.0,4.5)	0.9210 ± 0.0028	0.9114 ± 0.0154	0.8061 ± 0.0834
(4.5,5.0)	0.8871 ± 0.0045	0.8444 ± 0.0400	0.0 ± 0.0
>5.0	0.8787 ± 0.0106	0.0 ± 0.0	0.0 ± 0.0

Table A.10.: $K^{\pm}$ PID efficiencies in p_T,η determined from $K^0_S\to\pi^+\pi^-$ candidates in data.

p_{T} (MeV/c)	Selected $\mathrm{D_s^{\pm}}$ $N_{sel'}$	Selected $\mathrm{D_s^{\pm}}$ weighted by PID $N_{PID} = \sum_i^{N_{sel'}} \omega_{PIDi}$	PID Efficiency, ϵ_{pid} $\epsilon_{PID} = \frac{N_{PID}}{N_{sel'}}$
$(0, 1000)$	638 ± 25	396 ± 16	$0.620 \pm 0.019 \pm 0.007$
$(1000, 2000)$	1513 ± 39	998 ± 26	$0.660 \pm 0.012 \pm 0.008$
$(2000, 3000)$	1778 ± 42	1244 ± 30	$0.699 \pm 0.011 \pm 0.010$
$(3000, 4000)$	1407 ± 38	1054 ± 29	$0.749 \pm 0.012 \pm 0.016$
$(4000, 5000)$	977 ± 31	742 ± 24	$0.759 \pm 0.014 \pm 0.021$
$(5000, 6000)$	590 ± 24	452 ± 19	$0.766 \pm 0.017 \pm 0.028$
$(6000, 7000)$	372 ± 19	287 ± 16	$0.773 \pm 0.022 \pm 0.030$
$(7000, 8000)$	239 ± 15	186 ± 12	$0.778 \pm 0.027 \pm 0.032$

Table A.11.: Prompt $\mathrm{D_s^{\pm}}$, Monte-Carlo yields at selection level and PID efficiency in bins of p_{T}, integrated over y on the range (2.0,4.5). The first uncertainty on the PID efficiency is statistical, the second is systematic.

y	Selected $\mathrm{D_s^{\pm}}$ $N_{sel'}$	Selected $\mathrm{D_s^{\pm}}$ weighted by PID $N_{PID} = \sum_i^{N_{sel'}} \omega_{PIDi}$	PID efficiency, ϵ_{pid} $\epsilon_{PID} = \frac{N_{PID}}{N_{sel'}}$
$(2, 2.5)$	1038 ± 32	716 ± 23	$0.690 \pm 0.014 \pm 0.020$
$(2.5, 3)$	2289 ± 48	1707 ± 36	$0.746 \pm 0.009 \pm 0.015$
$(3, 3.5)$	2275 ± 48	1751 ± 37	$0.770 \pm 0.009 \pm 0.014$
$(3.5, 4)$	1429 ± 38	967 ± 26	$0.677 \pm 0.012 \pm 0.014$
$(4, 4.5)$	483 ± 22	217 ± 11	$0.450 \pm 0.023 \pm 0.017$

Table A.12.: Prompt $\mathrm{D_s^{\pm}}$, Monte-Carlo yields at selection level and PID efficiency in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c. The first uncertainty on the PID efficiency is statistical, the second is systematic.

p_{T} (MeV/c)	Selected $\mathrm{D}^{\pm}$ $N_{sel'}$	Selected $\mathrm{D}^{\pm}$ weighted by PID $N_{PID} = \sum_i^{N_{sel'}} \omega_{PIDi}$	Selection Efficiency, ϵ_{pid} $\epsilon_{PID} = \frac{N_{PID}}{N_{sel'}}$
$(0, 1000)$	816 ± 29	506 ± 18	$0.621 \pm 0.017 \pm 0.007$
$(1000, 2000)$	1707 ± 41	1125 ± 27	$0.659 \pm 0.012 \pm 0.008$
$(2000, 3000)$	1607 ± 40	1134 ± 29	$0.706 \pm 0.011 \pm 0.010$
$(3000, 4000)$	1196 ± 35	877 ± 26	$0.733 \pm 0.013 \pm 0.015$
$(4000, 5000)$	666 ± 26	505 ± 20	$0.759 \pm 0.017 \pm 0.021$
$(5000, 6000)$	380 ± 19	290 ± 15	$0.763 \pm 0.022 \pm 0.027$
$(6000, 7000)$	248 ± 16	185 ± 12	$0.747 \pm 0.028 \pm 0.030$
$(7000, 8000)$	146 ± 12	116 ± 10	$0.795 \pm 0.033 \pm 0.032$

Table A.13.: Prompt $\mathrm{D}^{\pm}$, Monte-Carlo yields at selection level and PID efficiency in bins of p_{T}, integrated over y on the range (2.0,4.5). The first uncertainty on the PID efficiency is statistical, the second is systematic.

y	Selected $\mathrm{D}^{\pm}$ $N_{sel'}$	Selected $\mathrm{D}^{\pm}$ weighted by PID $N_{PID} = \sum_i^{N_{sel'}} \omega_{PIDi}$	Selection efficiency, ϵ_{pid} $\epsilon_{PID} = \frac{N_{PID}}{N_{sel'}}$
$(2, 2.5)$	903 ± 30	606 ± 21	$0.672 \pm 0.016 \pm 0.017$
$(2.5, 3)$	2062 ± 45	1489 ± 33	$0.722 \pm 0.010 \pm 0.012$
$(3, 3.5)$	2072 ± 46	1557 ± 35	$0.752 \pm 0.010 \pm 0.012$
$(3.5, 4)$	1317 ± 36	893 ± 25	$0.678 \pm 0.013 \pm 0.013$
$(4, 4.5)$	412 ± 20	193 ± 10	$0.470 \pm 0.025 \pm 0.016$

Table A.14.: Prompt $\mathrm{D}^{\pm}$, Monte-Carlo yields at selection level and PID efficiency in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c. The first uncertainty on the PID efficiency is statistical, the second is systematic.

A.4. Total efficiency and efficiency corrected yields

p_{T} (MeV/c)	Raw Yield N_{raw}	Total Efficiency $\epsilon_{tot} = \epsilon_{acc} \times \epsilon_{sel'} \times \epsilon_{PID}$	Corrected Yield $N_{corr} = \frac{N_{raw}}{\epsilon_{tot}}$
$(0, 1000)$	2.4 ± 5.3	0.0080 ± 0.0004	302 ± 662
$(1000, 2000)$	66.6 ± 9.8	0.0176 ± 0.0006	3796 ± 570
$(2000, 3000)$	66.9 ± 9.7	0.0392 ± 0.0011	1705 ± 251
$(3000, 4000)$	75.6 ± 9.7	0.0658 ± 0.0020	1150 ± 151
$(4000, 5000)$	38.9 ± 7.0	0.0883 ± 0.0031	440 ± 80
$(5000, 6000)$	24.7 ± 5.8	0.1022 ± 0.0046	242 ± 57
$(6000, 7000)$	13.5 ± 4.1	0.1185 ± 0.0066	114 ± 35
$(7000, 8000)$	12.9 ± 3.9	0.1299 ± 0.0089	100 ± 31

Table A.15.: $\mathrm{D_s^{\pm}}$ effective yields in 4π in bins of p_{T}, integrated over y on the range (2.0,4.5).

y	Raw Yield N_{raw}	Total Efficiency $\epsilon_{tot} = \epsilon_{acc} \times \epsilon_{sel'} \times \epsilon_{PID}$	Corrected Yield $N_{corr} = \frac{N_{raw}}{\epsilon_{tot}}$
$(2, 2.5)$	39.5 ± 7.1	0.0159 ± 0.0006	2489 ± 455
$(2.5, 3)$	92.1 ± 10.9	0.0422 ± 0.0010	2182 ± 264
$(3, 3.5)$	122.5 ± 12.4	0.0504 ± 0.0012	2431 ± 252
$(3.5, 4)$	44.1 ± 9.4	0.0343 ± 0.0011	1284 ± 277
$(4, 4.5)$	3.6 ± 4.0	0.0100 ± 0.0007	364 ± 405

Table A.16.: $\mathrm{D_s^{\pm}}$ effective yields in 4π in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c.

p_{T} (MeV/c)	Raw Yield N_{raw}	Total Efficiency $\epsilon_{tot} = \epsilon_{acc} \times \epsilon_{sel'} \times \epsilon_{PID}$	Corrected Yield $N_{corr} = \frac{N_{raw}}{\epsilon_{tot}}$
$(0, 1000)$	13.5 ± 5.2	0.0207 ± 0.0009	650 ± 255
$(1000, 2000)$	34.6 ± 7.3	0.0422 ± 0.0012	821 ± 175
$(2000, 3000)$	55.5 ± 8.5	0.0810 ± 0.0023	685 ± 106
$(3000, 4000)$	43.2 ± 7.2	0.1234 ± 0.0039	351 ± 59
$(4000, 5000)$	24.3 ± 5.6	0.1439 ± 0.0059	169 ± 39
$(5000, 6000)$	8.4 ± 3.4	0.1559 ± 0.0084	54 ± 22
$(6000, 7000)$	8.1 ± 3.1	0.1756 ± 0.0117	46 ± 18
$(7000, 8000)$	3.6 ± 2.2	0.2064 ± 0.0171	17 ± 11

Table A.17.: $\mathrm{D}^{\pm}$ effective yields in 4π in bins of p_{T}, integrated over y on the range (2.0,4.5).

y	Raw Yield N_{raw}	Total Efficiency $\epsilon_{tot} = \epsilon_{acc} \times \epsilon_{sel'} \times \epsilon_{PID}$	Corrected Yield $N_{corr} = \frac{N_{raw}}{\epsilon_{tot}}$
$(2, 2.5)$	34.5 ± 6.6	0.0290 ± 0.0012	1189 ± 231
$(2.5, 3)$	68.2 ± 9.2	0.0792 ± 0.0020	861 ± 119
$(3, 3.5)$	50.1 ± 8.7	0.0963 ± 0.0023	520 ± 91
$(3.5, 4)$	35.4 ± 6.9	0.0681 ± 0.0022	521 ± 103
$(4, 4.5)$	3.1 ± 3.0	0.0192 ± 0.0014	160 ± 157

Table A.18.: $\mathrm{D}^{\pm}$ effective yields in 4π in bins of y, integrated over p_{T} on the range $(0, 8000)$MeV/c.

Colophon

This thesis was made in LaTeX 2_ε using the `hepthesis` class [112]. Diagrams, unless referenced were produced with `inkscape` [113] and plots were produced exclusively within the `ROOT` framework [114].

Bibliography

[1] J. Houghton, Philosophical Transactions **21**, 311 (1699).

[2] F.-J. Lapointe and P. Legendre, Journal of the Royal Statistical Society. Series C (Applied Statistics) **43**, pp. 237 (1994).

[3] C. Quigg, Annual Review of Nuclear and Particle Science **59**, 505 (2009), arXiv:0905.3187.

[4] D. J. Gross, Rev. Mod. Phys. **77**, 837 (2005).

[5] B. R. Webber and P. Weber, *Introduction to QCD*, 2003.

[6] G. Altarelli, A QCD Primer, in *New States of Matter in Hadronic Interactions*, edited by H.-T. Elze, E. M. Ferreira, T. Kodama, J. Letessier, J. Rafelski, & R. L. Thews, , American Institute of Physics Conference Series Vol. 631, pp. 70–111, 2002, arXiv:hep-ph/0204179.

[7] Particle Data Group, C. Amsler *et al.*, Phys. Lett. B **667**, 1 (2008).

[8] E. Noether, Invariant variation problems, 1971, arXiv:physics/0503066.

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Physical Review Letters **13**, 138 (1964).

[10] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[11] I. Dunietz, R. Fleischer, and U. Nierste, Phys. Rev. D **63**, 114015 (2001), arXiv:hep-ph/0012219.

[12] LHCb Collaboration, B. Adeva *et al.*, ArXiv e-prints (2009), arXiv:0912.4179.

[13] CDF/DØ $\Delta\Gamma_s$, β_s Combination Working Group, Combination of DØ and CDF Results on $\Delta\Gamma_s$ and the $\mathcal{CP}$-Violating Phase β_s, 2009, CDF/PHYS/BOTTOM/CDFR/9787,DØ Note 5928-CONF.

[14] S. Burdin, Measurements of $\mathcal{CP}$ violation in the B_s^0 system at DØ, in *EPS-HEP*, 2011, PoS(EPS-HEP2011)148.

[15] CDF Collaboration, Measurement of the B_s^0 mixing phase using $B_s^0 \to J/\psi\,\phi$ decays from the full run II dataset, 2012, CDF note 10778.

[16] SuperB Collaboration, ArXiv e-prints (2007), arXiv:0709.0451.

[17] T. Aushev *et al.*, ArXiv e-prints (2010), arXiv:1002.5012.

[18] HERA-B Collaboration, A. J. Schwartz, Nucl. Inst. Meth. A **446**, 199 (2000), arXiv:hep-ex/9912012.

[19] CLEO Collaboration, D. Andrews *et al.*, Nucl. Inst. Meth. **211**, 47 (1983).

[20] CUSB Collaboration, G. Finocchiaro *et al.*, Phys. Rev. Lett. **45**, 222 (1980).

[21] CLEO Collaboration, D. Andrews *et al.*, Physical Review Letters **45**, 219 (1980).

[22] CLEO Collaboration, C. Bebek *et al.*, Physical Review Letters **46**, 84 (1981).

[23] CLEO Collaboration, K. Chadwick *et al.*, Physical Review Letters **46**, 88 (1981).

[24] CLEO Collaboration, S. Behrends *et al.*, Physical Review Letters **50**, 881 (1983).

[25] ARGUS Collaboration, H. Albrecht *et al.*, Nucl. Inst. Meth. A **275**, 1 (1989).

[26] ARGUS Collaboration, H. Albrecht *et al.*, Physics Letters B **192**, 245 (1987).

[27] UA1 Collaboration, G. Arnison *et al.*, Physics Letters B **126**, 398 (1983).

[28] J. T. Seeman, Annual Review of Nuclear and Particle Science **41**, 389 (1991).

[29] CERN, *LEP design report vol. 1 : The LEP injector chain* (CERN, Geneva, 1983), CERN-LEP-TH-83-29.

[30] CERN, *LEP design report vol. 2 : The LEP main ring* (CERN, Geneva, 1984), CERN-LEP-84-01.

[31] SLD Collaboration, P. C. Rowson, D. Su, and S. Willocq, Annual Review of Nuclear and Particle Science **51**, 345 (2001), arXiv:hep-ph/0110168.

[32] LEP EW Combination Working Group, Nucl. Inst. Meth. A **378**, 101 (1996).

[33] ALEPH Collaboration, D. Decamp *et al.*, Physics Letters B **278**, 209 (1992).

[34] ALEPH Collaboration, D. Buskulic *et al.*, Physics Letters B **294**, 145 (1992).

[35] OPAL Collaboration, P. D. Acton *et al.*, Physics Letters B **295**, 357 (1992).

[36] Y. Rozen, Nuclear Physics B Proceedings Supplements **109**, 99 (2002).

[37] CDF Collaboration, F. Abe *et al.*, Nucl. Inst. Meth. A **271**, 387 (1988).

[38] CDF Collaboration, F. Abe *et al.*, Physical Review Letters **74**, 2626 (1995), arXiv:hep-ex/9503002.

[39] CDF Collaboration, A. Abulencia *et al.*, Physical Review Letters **97**, 242003 (2006), arXiv:hep-ex/0609040.

[40] CDF, DØ Collaborations, TEVNPHWG Working Group, ArXiv e-prints (2011), arXiv:1103.3233.

[41] CDF Collaboration, T. Aaltonen *et al.*, Physical Review Letters **100**, 101802 (2008), arXiv:0712.1708.

[42] P. Oddone, DETECTOR CONSIDERATIONS, in *Workshop on Conceptual Design of a Test Linear Collider: Possibilities for a B anti-B Factory*, UCLA, 1987.

[43] BABAR Collaboration, P. F. Harrison, ed. and H. R. Quinn, ed., Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[44] BELLE Collaboration, M. Cheng *et al.*, A Study of CP violation in B meson decays: Technical design report, 1995.

[45] BABAR Collaboration, P. Biassoni, ArXiv e-prints (2011), arXiv:1102.1370.

[46] BELLE Collaboration, P. Chen *et al.*, Phys. Rev. D **84**, 071501 (2011), arXiv:1108.4271.

[47] CERN AC Team, Diagram of an lhc dipole magnet., CERN-DI-9906025, 1999.

[48] C. Lefevre, The CERN accelerator complex, CERN-DI-0606052, 2006.

[49] G. Brianti, Phys. Rep. **403**, 349 (2004).

[50] O. S. Bruning *et al.*, *LHC Design Report* (CERN, Geneva, 2004), CERN-2004-003-V-1.

[51] A. Cho, Science **321**, 1753 (2008).

[52] CERN Press Office, *The LHC is back* (CERN, Geneva, 2009), PR16.09.

[53] ATLAS Collaboration, W. W. Armstrong *et al.*, (1994), CERN-LHCC-94-43.

[54] CMS Collaboration, M. Della Negra *et al.*, (1994), CERN-LHCC-94-38.

[55] ALICE Collaboration, J. Schukraft *et al.*, (1995), CERN-LHCC-95-71.

[56] LHCb Collaboration, S. Amato *et al.*, (1998), CERN-LHCC-98-004 ; LHCC-P-4.

[57] Lhcb Collaboration, R. Aaij *et al.*, Physics Letters B **694**, 209 (2010), arXiv:1009.2731.

[58] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, European Physical Journal C **63**, 189 (2009), arXiv:0901.0002.

[59] LHCb Collaboration *et al.*, Journal of Instrumentation **3**, 8005 (2008).

[60] LHCb Collaboration, *LHCb TDR 1: Magnet* (CERN, 2000), CERN-LHCC-2000-007.

[61] J. D. Jackson, *Classical electrodynamics*, second ed. (Wiley, 1975).

[62] T. Ypsilantis and J. Seguinot, Nucl. Inst. Meth. A **343**, 30 (1994).

[63] J. Bibby *et al.*, Nucl. Inst. Meth. A **546**, 93 (2005).

[64] L. Somerville, Nucl. Inst. Meth. A **546**, 81 (2005).

[65] E. G. Chesi *et al.*, CERN Report No. LHCb-98-037, 1998 (unpublished).

[66] K. Wyllie, Nucl. Inst. Meth. A **591**, 260 (2008).

[67] A. Sancho, IEEE Transactions on Nuclear Science **43**, 1773 (1996).

[68] N. Styles, *RICH detector time alignment and studies of CP violation in the decay $B_s^0 \to \phi\phi$ at the LHCb experiment.*, PhD thesis, Edinburgh, University of Edinburgh, Edinburgh, 2009, Presented on 21 Oct 2009.

[69] D. Breton and D. Charlet, CERN Report No. LHCb-2003-004, 2003 (unpublished).

[70] O. Holme, M. Gonzlez-Berges, P. Golonka, and S. Schmeling, CERN Report No. CERN-OPEN-2005-027, 2005 (unpublished).

[71] M. Gancarz, *Linux and the Unix philosophy* (Digital Press, 2003).

[72] C. Fitzpatrick, CERN Report No. LHCb-INT-2009-029, 2009 (unpublished).

[73] M. Matsumoto and T. Nishimura, ACM Trans. Model. Comput. Simul. **8**, 3 (1998).

[74] A. Hoecker *et al.*, ArXiv Physics e-prints (2007), arXiv:physics/0703039.

[75] M. Pivk and F. R. Le Diberder, Nucl. Inst. Meth. A **555**, 356 (2005), arXiv:physics/0402083.

[76] CLEO Collaboration, J. P. Alexander *et al.*, Physical Review Letters **100**, 161804 (2008), arXiv:0801.0680.

[77] LHCb Collaboration, R. Aaij *et al.*, Journal of Instrumentation **7**, 1010 (2012).

[78] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, ArXiv e-prints (2012), arXiv:1202.0439.

[79] M. Cacciari, S. Frixione, M. Mangano, M. Nason, and G. Ridolfi, (2010), private communication.

[80] T. Sjöstrand, L. Lönnblad, and S. Mrenna, ArXiv High Energy Physics - Phenomenology e-prints (2001), arXiv:hep-ph/0108264.

[81] M. Calvi, B. Khanji, G. Lanfranchi, O. Leroy, and S. Poss, CERN Report No. LHCb-2009-025. CERN-LHCb-2009-025, 2009 (unpublished).

[82] M. Needham, CERN Report No. LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002, 2008 (unpublished).

[83] W. D. Hulsbergen, Nuclear Instruments and Methods in Physics Research A **552**, 566 (2005), arXiv:physics/0503191.

[84] R. A. Fisher, Royal Society of London Philosophical Transactions Series A **222**, 309 (1922).

[85] A. W. F. Edwards, International Statistical Review / Revue Internationale de Statistique **42**, pp. 9 (1974).

[86] F. James and M. Roos, Computer Physics Communications **10**, 343 (1975).

[87] J. S. Conway, ArXiv e-prints (2011), arXiv:1103.0354.

[88] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998),

arXiv:physics/9711021.

[89] Y. Xie, P. Clarke, G. Cowan, and F. Muheim, Journal of High Energy Physics **9**, 74 (2009), arXiv:0908.3627.

[90] DØ Collaboration, V. M. Abazov *et al.*, Phys. Rev. D **85**, 032006 (2012), arXiv:1109.3166.

[91] CDF Collaboration, T. Aaltonen *et al.*, ArXiv e-prints (2011), arXiv:1112.1726.

[92] Y. Xie, ArXiv e-prints (2009), arXiv:0905.0724.

[93] T. Du Pree and G. Raven, CERN Report No. LHCb-2009-024. CERN-LHCb-2009-024, 2009 (unpublished).

[94] T. Du Pree, *Search for a Strange Phase in Beautiful Oscillations*, PhD thesis, Amsterdam, Vrije Universiteit, Amsterdam, 2010, Presented on 22 Oct 2010.

[95] H.-G. Moser and A. Roussarie, Nucl. Inst. Meth. A **384**, 491 (1997).

[96] LHCb Collaboration, "β_s and mixing Working Group", Report No. LHCb-ANA-2011-006, 2011 (unpublished).

[97] LHCb Collaboration, M. Calvi, O. Leroy, and M. Musy, Report No. LHCb-2007-058, 2007 (unpublished).

[98] LHCb Collaboration, R. Aaij *et al.*, ArXiv e-prints (2012), arXiv:1202.4979.

[99] LHCb Collaboration, R. Aaij *et al.*, Physics Letters B **709**, 177 (2012), arXiv:1112.4311.

[100] LHCb Collaboration, R. Aaij *et al.*, Physical Review Letters **108**, 101803 (2012), arXiv:1112.3183.

[101] LHCb Collaboration, β_s and mixing Working Group, Report No. LHCb-ANA-2011-036, 2011 (unpublished).

[102] A. Lenz and U. Nierste, Journal of High Energy Physics **6**, 72 (2007), arXiv:hep-ph/0612167.

[103] A. Badin, F. Gabbiani, and A. A. Petrov, Phys. Lett. B **653**, 230 (2007), arXiv:0707.0294.

[104] A. Lenz and U. Nierste, ArXiv e-prints (2011), arXiv:1102.4274.

[105] J. Charles *et al.*, Phys. Rev. D **84**, 033005 (2011), arXiv:1106.4041.

[106] A. Lenz *et al.*, ArXiv e-prints (2012), arXiv:1203.0238.

[107] D. M. Straub, ArXiv e-prints (2011), arXiv:1107.0266.

[108] LHCb collaboration, R. Aaij *et al.*, ArXiv e-prints (2012), arXiv:1203.4493.

[109] S. Stone, First physics results from lhcb, in *ICHEP*, 2010, PoS(ICHEP 2010)025.

[110] LHCb Collaboration, R. Aaij *et al.*, ArXiv e-prints (2012), arXiv:1202.4717.

[111] LHCb Collaboration, R. Aaij *et al.*, Report No. LHCb-CONF-2012-002, 2012 (unpublished), Preliminary result for winter conferences.

[112] A. Buckley, The `hepthesis` class, CTAN.org, 2007.

[113] The Inkscape Team, Inkscape, Open Source Scalable Vector Graphics Editor, inkscape.org, 2009.

[114] R. Brun and F. Rademakers, ROOT - An object oriented data analysis framework, 1997.

List of figures

List of tables