Searching for CP Violation in the $B_s^0 \to \phi\phi$ decay at LHCb

Sean Benson
University of Edinburgh

A dissertation submitted to the University of Edinburgh
for the degree of Doctor of Philosophy
January 2014

Abstract

The study of flavour physics allows for the Standard Model (SM) to be tested to higher energies than can be accessed through direct searches. The SM is known not to provide enough of a difference between matter and anti-matter, termed *CP* violation, to explain the dominance of matter in our universe. One of the main purposes of the LHCb experiment is to search for new sources of *CP* violation in the decays of B mesons. Flavour changing neutral current (FCNC) interactions are forbidden at tree level in the SM, and can therefore only be accessed through quantum loops. In New Physics scenarios such as Supersymmetry, new particles could appear in those loops introducing new sources of *CP* violation. The $B_s^0 \to \phi\phi$ decay proceeds via the $b \to s\bar{s}s$ FCNC transition. Triple products provide a method of exploiting the angular distributions of $P \to VV$ decays to create T-odd observables. Asymmetries of these T-odd observables, averaged over the initial flavour of the B_s^0 meson provide a measure of T violation. Assuming *CPT* conservation, violation of time reversal infers *CP* violation. The *CP*-violating weak phase in the interference between B_s^0 mixing and the decay to two ϕ mesons is predicted to be close to zero in the SM. The measurements of the triple product asymmetries and the *CP*-violating weak phase have been performed using $1.0\,\text{fb}^{-1}$ of LHCb data. Events where kaon pairs originate from a spin-0 or non-resonant state are accounted for with the associated angular distributions. Triple product asymmetries are measured to be $A_U = -0.055 \pm 0.036(\text{stat}) \pm 0.018(\text{syst})$ and $A_V = 0.010 \pm 0.036(\text{stat}) \pm 0.018(\text{syst})$. The *CP*-violating phase is found to be in the interval $[-2.46, -0.76]\,\text{rad}$ at $68\,\%$ confidence level. The p-value for the hypothesis of zero radians is found to be $16\,\%$. These results represent the most accurate measurements of the triple product asymmetries and the first measurement of the *CP*-violating weak phase.

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 post-graduate assessment regulations for research degrees.

Sean Benson

"To my Father"

Acknowledgements

Firstly, a big thank you needs to be given to my supervisor Franz Muheim, who has always found time to offer advice, even when there is precious little time to spare, and who has undertaken vast quantities of reading. My Dad, Joseph, my Nan, Ann, and my better half, Becky have had to endure my time shared between them and my laptop, and have always offerred encouragement. I am indebted to the Edinburgh Particle Physics Group, which contains a broad wealth of knowledge and experience, and who never mind someone dropping in to pick brains, with extra thanks to Yuehong, Matt and Franz, who have had to put up with more than most. The Charmless working group provides a fantastic environment to perform challenging analyses and shows all the benefits of working in a collaboration. I owe particular thanks to the convenors, Stephane, Vincenzo, and Angelo, for the opportunities they have given me.

Preface

For the universe we all know to have been created, three conditions, first proposed by Andrei Sakharov, are known to be required. These are that baryon number needs to be violated, there must be an absence of thermal equilibrium, and matter must behave differently to anti-matter. Our current understanding of the laws of physics do not provide the relatively large differences between matter and anti-matter needed to produce our universe. It is the main goal of the LHCb detector, situated on the Large Hadron Collider at CERN, Geneva to find asymmetries between the decays of short lived matter and anti-matter states. The measurements presented in this thesis describe my contribution as part of the LHCb collaboration. This mainly focuses on the study of loop transitions, where such asymmetries have been studied to the highest known precision and in some cases for the first time.

Contents

1. CP Violation and the Phenomenology of B_s^0 mesons **3**
- 1.1. Introduction to the Standard Model . . . 3
  - 1.1.1. Electroweak Interaction . . . 4
  - 1.1.2. Strong Interaction . . . 6
- 1.2. *CP* Violation . . . 7
- 1.3. Phenomenology of B_s^0 mixing . . . 12
- 1.4. $b \to s$ Transitions . . . 15
  - 1.4.1. B_s^0-$\overline{B}_s^0$ Mixing . . . 15
    - 1.4.1.1. Standard Model B_s^0-$\overline{B}_s^0$ Mixing . . . 15
    - 1.4.1.2. Prospects for New Physics in B_s^0-$\overline{B}_s^0$ Mixing . . . 17
  - 1.4.2. The $B_s^0 \to \phi\phi$ Decay . . . 19
    - 1.4.2.1. The $B_s^0 \to \phi\phi$ Decay in the Standard Model . . . 22
    - 1.4.2.2. Prospects for New Physics in Penguin Decays . . . 24
- 1.5. Summary . . . 28

2. The LHCb Experiment **29**
- 2.1. The Large Hadron Collider . . . 29
- 2.2. The LHCb Detector . . . 30
- 2.3. Tracking System . . . 31
  - 2.3.1. Magnet . . . 31
  - 2.3.2. Vertex Location . . . 32
  - 2.3.3. Tracker Turicensis and Inner Tracker . . . 34
  - 2.3.4. Outer Tracker . . . 36
- 2.4. Particle Identification . . . 38
  - 2.4.1. RICH Detectors . . . 38
  - 2.4.2. Calorimetry . . . 40
  - 2.4.3. Muon Chambers . . . 43
    - 2.4.3.1. MWPC . . . 44

2.4.3.2. GEM Detectors 44
2.4.3.3. Muon Identification Performance 45
2.5. Trigger System 46
2.5.1. Hardware Trigger 47
2.5.1.1. Calorimeter Trigger 47
2.5.1.2. Muon Trigger 48
2.5.2. Software Trigger 49
2.5.2.1. HLT1 49
2.5.2.2. HLT2 50
2.6. RICH HPD Efficiencies 51
2.6.1. Origins, Trends and Solution 52
2.6.2. Relation to RICH Performance 55

3. Fit Methods **57**
3.1. Negative Log Likelihood Fitting 57
3.1.1. Parameter Uncertainties 58
3.1.2. Nuisance Parameters 59
3.2. The Feldman Cousins Method and Confidence Intervals 60
3.3. Inclusion of Efficiency Corrections in the Log Likelihood 63
3.4. The ${}_s\mathcal{P}lot$ Method 65
3.4.1. Application of the ${}_s\mathcal{P}lot$ Method to an Unbinned Likelihood Fit . 66
3.5. Summary 66

4. Isolating the $B_s^0 \to \phi\phi$ decay **67**
4.1. Cut-based Isolation 68
4.1.1. Dataset 68
4.1.2. Optimisation Method and Results 68
4.2. Multi-variate Isolation 71
4.2.1. Boosted Decision Trees 73
4.2.2. Datasets 75
4.2.3. BDT Training Method 76
4.2.4. Optimisation 79
4.3. Summary 81

5. Measurement of the Polarisation Amplitudes and Triple Product Asymmetries in $B_s^0 \to \phi\phi$ Decays **83**
5.1. Dataset 83

5.2. Polarisation Amplitudes . . . 84
    5.2.1. Angular Distributions . . . 84
    5.2.2. Fit Details and Constraints . . . 87
    5.2.3. Angular Acceptance . . . 88
    5.2.4. Results . . . 89
    5.2.5. Systematic Uncertainties . . . 89
        5.2.5.1. S-wave Pollution . . . 90
        5.2.5.2. Angular Acceptance Uncertainty . . . 92
        5.2.5.3. Effect of Time Acceptance . . . 93
        5.2.5.4. Additional Uncertainties . . . 95
5.3. Triple Product Asymmetries . . . 96
    5.3.1. T-odd Observables . . . 96
    5.3.2. Fit Method . . . 97
    5.3.3. Results . . . 98
    5.3.4. Systematic Uncertainties . . . 99
        5.3.4.1. Angular Acceptance Uncertainty . . . 99
        5.3.4.2. Effect of Time Acceptance . . . 100
        5.3.4.3. Mass Model and Calibration Uncertainty . . . 101
5.4. Final Results and Summary . . . 102

6. First Measurement of the CP-Violating Phase in $B_s^0 \to \phi\phi$ Decays 105
6.1. Dataset . . . 105
6.2. Angular and Time-Dependent Decay Rates . . . 106
    6.2.1. S-wave Coupling . . . 107
6.3. Identification of B_s^0 Flavour at Production . . . 110
    6.3.1. Opposite-Side Flavour Tagging . . . 111
        6.3.1.1. Single-Particle OS Tagging Algorithms . . . 111
        6.3.1.2. Vertex Charge OS Tagging Algorithm . . . 112
    6.3.2. Same-Side Kaon Flavour Tagging . . . 113
    6.3.3. Mistag Probabilities and Combination of Different Flavour-Tagging Algorithms . . . 113
    6.3.4. Calibration of Flavour-Tagging Algorithms . . . 114
6.4. Decay Time Resolution . . . 116
6.5. Decay Time Acceptance . . . 118
6.6. Angular Acceptance . . . 119

6.7. Fit Details and Ingredients . . . 120
  6.7.1. Parameters and External Inputs . . . 121
6.8. Fit Results . . . 122
6.9. Systematic Uncertainties . . . 123
  6.9.1. Angular Acceptance . . . 124
  6.9.2. Decay Time Acceptance . . . 125
  6.9.3. S-wave . . . 126
  6.9.4. Mass Model . . . 129
  6.9.5. Peaking Background . . . 131
  6.9.6. Decay Time Resolution . . . 133
  6.9.7. Fit Bias . . . 133
6.10. Feldman Cousins Coverage Correction . . . 135
6.11. Final Results and Summary . . . 136

7. Implications and Future Prospects 139
7.1. Upgrade Components . . . 140
  7.1.1. Vertex Locator . . . 140
  7.1.2. Tracking Stations . . . 141
  7.1.3. RICH Detectors . . . 141
  7.1.4. Trigger . . . 142
7.2. Upgrade Performance . . . 144
  7.2.1. Efficiency Comparison . . . 145
  7.2.2. Comparison of Acceptances . . . 146
  7.2.3. Comparison of Resolutions . . . 147
7.3. Expected Precision of Upgrade Measurements . . . 150

8. Summary, Conclusions and Outlook 153

A. Detailed Angular Formulae Derivation 159
A.1. Form of the Amplitude . . . 159
A.2. Time Evolution . . . 163

B. T-odd Triple Products in $P \to VV$ Decays 169

C. Calculation of the B_s^0 Mixing Box Diagram 173

D. Theory of the $\overline{B}_s^0 \to \phi_L \phi_L$ Amplitude 179

E. Comparisons Between Data and Monte Carlo **185**

Bibliography **189**

“Our ideas must be as broad as Nature if they are to interpret Nature.”
— A Study in Scarlet, Sir Arthur Conan Doyle

Chapter 1.

CP Violation and the Phenomenology of B_s^0 mesons

"No man should escape our Universities without knowing how little he knows."
— J. R. Oppenheimer

1.1. Introduction to the Standard Model

The most successful description of nature to date in terms of fundamental interactions between particles is provided by a theory known as the Standard Model (SM) of particle physics. The SM as known today was finalised in the 1970s [1], and has to this date withstood every attempt to directly disprove predictions made from the theory.

The gauge group of the Standard Model (SM) is $SU(2)_L \otimes U(1)_Y \otimes SU(3)_C$, which determines the interactions between the various fundamental particles, shown in Figure 1.1. The current known fundamental particles consist of 12 fermions, along with associated anti-particles, together with 12 vector bosons describing the forces of the SM and a scalar boson responsible for fermion masses. The 12 vector bosons consist of 8 gluons mediating the strong interaction, with associated gauge group, $SU(3)_C$, the $W^{\pm}$ and Z^0 bosons mediating the weak interaction, with associated gauge group $SU(2)_L$, and the photon mediating the electromagnetic interaction, with associated gauge group, $U(1)_\gamma$.

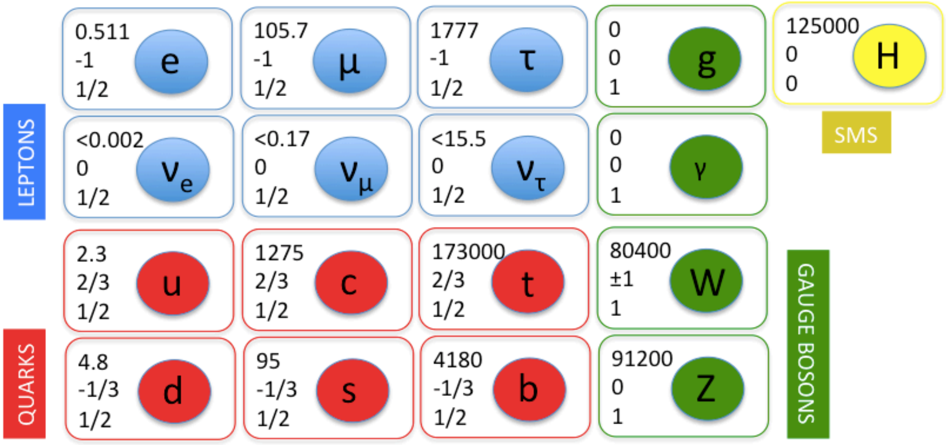


Figure 1.1.: Current understanding of the fundamental constituents of matter. In each box, the approximate mass in MeV/c^2 is given, along with the electric charge and spin of the different particle types.

1.1.1. Electroweak Interaction

The electromagnetic and weak interactions may be unified to a single force at high enough energies, known as the electroweak interaction. The Lagrangian describing the electroweak interaction consists of terms satisfying the $SU(2)_L \otimes U(1)_Y$ gauge symmetry and is given by

$$\mathcal{L}_{EW} = -\frac{1}{4}\sum_i F^i_{\mu\nu}F^{i\,\mu,\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + i\sum_j \bar{L}_j\gamma^\mu D^{EW}_\mu L_j + i\sum_k \bar{R}_k\gamma^\mu D^{U(1)}_\mu R_k, \quad (1.1)$$

where a sum over i is over the non-Abelian gauge fields of the $SU(2)_L$ group, the sum over j is over left-handed (anti) fermion doublets, $L_j(\overline{L}_j)$, and the sum over k is over right-handed (anti) fermion singlets that carry electric charge, $R_k(\overline{R}_k)$. The Dirac matrices are denoted by γ^μ. The field strength tensors, $\mathbf{F}_{\mu\nu} = i/g_1[\mathbf{D}^{(SU(2)}_\mu, \mathbf{D}^{SU(2)}_\nu]$ and $B_{\mu\nu} = i/g_2[D^{(U(1)}_\mu, D^{U(1)}_\nu]$, are defined from the covariant derivatives of the $SU(2)_L$ and $U(1)_Y$ gauge groups, respectively. These can be written as

$$\mathbf{D}^{SU(2)}_\mu = \partial_\mu \mathbf{I} - i\frac{g_1}{2}\sigma^i W^i_\mu, \quad (1.2)$$

$$D^{U(1)}_\mu = \partial_\mu - i\frac{g_2}{2}b_\mu, \quad (1.3)$$

together with a combined electroweak covariant derivative

$$D^{EW}_\mu = \partial_\mu \mathbf{I} - i\frac{g_2}{2}b_\mu \mathbf{I} - i\frac{g_1}{2}\sigma^i W^i_\mu, \quad (1.4)$$

where g_1 and g_2 are the coupling constants, of the $SU(2)_L$ and $U(1)_Y$ gauge groups, respectively. The fields $\mathbf{W}$ and b represent the fields of the gauge bosons for the $SU(2)_L$ and $U(1)_Y$ gauge groups. The matrices, σ^i, are the Pauli matrices forming a basis in the adjoint representation of $SU(2)$. The $SU(2)_L$ gauge group acts on the following left-handed fermion doublets:

$$L_e = \begin{pmatrix} e^- \\ \nu_e \end{pmatrix}_L , L_\mu = \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}_L , L_\tau = \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}_L , \tag{1.5}$$

$$L_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L , L_2 = \begin{pmatrix} c \\ s \end{pmatrix}_L , L_3 = \begin{pmatrix} t \\ b \end{pmatrix}_L . \tag{1.6}$$

The $U(1)_Y$ gauge group acts on right-handed singlets:

$$R_{u,c,t} = \Big(u, c, t\Big)_R , R_{d,s,b} = \Big(d, s, b\Big)_R \, R_{e,\mu,\tau} = \Big(e, \mu, \tau\Big)_R . \tag{1.7}$$

It should be noted that all neutrinos observed so far in nature appear with a left-handed helicity. This has a consequence that a mass term cannot be added by hand to the electroweak Lagrangian as this would violate gauge invariance.

The electroweak gauge bosons observed in nature consist of three massive particles ($W^\pm$ and Z^0) together with one massless particle (A^γ). The Brout-Englert-Higgs (BEH) mechanism [2–5], derived from Goldstone's theorem [6] accounts for this through spontaneous symmetry breaking according to $SU(2)_L \otimes U(1)_Y \rightarrow U(1)_\gamma$. This has the effect of generating fermion masses in addition to giving the weak gauge bosons mass through the coupling to a spin-0 particle known as the Higgs boson.

The BEH mechanism centres on the introduction of a scalar doublet

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} . \tag{1.8}$$

As this is a complex doublet, there are four degrees of freedom. Gauge fixing provides a way to eliminate three of these leaving one physical degree of freedom, $h(x)$, which is allowed to have a vacuum expectation value, v. This then yields

$$\Phi = \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} , \tag{1.9}$$

where $h(x) = 0$ signifies the ground state. Inserting equation 1.9 in to the electroweak Lagrangian allows us to define the physical gauge bosons as [7]

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \tag{1.10}$$

$$Z_\mu^0 = b_\mu \sin\theta_w + W_\mu^3 \cos\theta_w, \tag{1.11}$$

$$A_\mu^\gamma = b_\mu \cos\theta_w - W_\mu^3 \sin\theta_w, \tag{1.12}$$

where θ_w is the Weinberg mixing angle, which can be found from the ratio of the Gauge boson masses through the relation $M_W = M_Z \cos\theta_w$.

1.1.2. Strong Interaction

Quantum Chromodynamics (QCD) is an example of a non-Abelian gauge theory and describes the interaction that binds quarks in to mesons and baryons. This Lagrangian consists of terms satisfying $SU(3)_C$ gauge symmetry and is given by

$$\mathcal{L}_{QCD} = -\frac{1}{4}\sum_i \mathbf{G}_{\mu\nu}^i \mathbf{G}^{i\,\mu\nu} + i\sum_f \bar{q}^{(f)}\gamma^\mu D_\mu q^{(f)}, \tag{1.13}$$

where $\mathbf{G}_{\mu\nu} = i/g_s[\mathbf{D}_\mu, \mathbf{D}_\nu]$ is the gluonic field strength tensor, $\mathbf{D}_\mu = \partial_\mu \mathbf{I} - ig_s\lambda^i A_\mu^i/2$ is the covariant derivative, and λ^i are the Gell-Mann matrices that generate the SU(3) group. Note that in equation 1.13, the sum over i is of gluon fields and the sum over f is of quark flavours. Gluons are represented by A_μ, whereas colour triplet fields of a given quark flavour are represented by q^f, with g_s denoting the coupling strength of the QCD interaction. In QCD, the three colours and associated anti-colours (often called red, green and blue) are analogous to the electric charge of QED.

In quantum field theories such as QCD, the β-function provides the relationship between the coupling constant and the energy scale, μ, of an interaction. The β-function is defined as

$$\beta(g) = \frac{\partial g}{\partial \log(\mu)}. \tag{1.14}$$

In QCD, this then means that the coupling constant, $\alpha_s^2 \equiv g_s^2/(4\pi)$, varies as

$$\alpha_s(q^2) = \frac{\alpha_s(\mu^2)}{1 + \left(\frac{33-2n_f}{3}\right)\frac{\alpha_s(\mu^2)}{4\pi}\ln\left(\frac{q^2}{\mu^2}\right)}, \tag{1.15}$$

where n_f is the number of quark flavours, μ represents a reference energy scale and q the momentum transfer in an interaction. Equation 1.15 then shows that for $n_f = 6$, as the energy of an interaction increases, the coupling constant decreases. This is known as asymptotic freedom and ensures that perturbation theory becomes increasingly reliable as interactions get harder.

The QCD potential between two quarks is

$$V_{QCD}(r) = -\frac{4}{3}\frac{\alpha_s}{r} + kr, \tag{1.16}$$

where r is the radial distance and k is a positive constant. This then means that as the distance between two quarks gets larger, the potential increases and results in the confinement of quarks in to bound states.

1.2. *CP* Violation

As has been mentioned previously, the gauge group of the Standard Model is $SU(2)_L \otimes U(1)_Y \otimes SU(3)_C$, the first two groups describing the electroweak interaction, the third describing the strong interaction[1]. Due to the real nature of the gauge couplings, to discuss *CP* violation, it is necessary to introduce Yukawa interactions describing the SM dynamical sector through the Lagrangian

$$\begin{aligned} L_{Yukawa} = &-\sum_{i,j}(G_U)_{ij}(\overline{L}_{i,U}, \overline{L}_{i,D})\begin{pmatrix}\phi^0 \\ -\phi^-\end{pmatrix} R_{j,U} \\ &-\sum_{i,j}(G_D)_{ij}(\overline{L}_{i,U}, \overline{L}_{i,D})\begin{pmatrix}\phi^0 \\ -\phi^-\end{pmatrix} R_{j,D} + h.c., \end{aligned} \tag{1.17}$$

where the ϕ fields form a Higgs $SU(2)$ doublet, the $L_i(\overline{L}_i)$ and $R_i(\overline{R}_i)$ are the fermionic field doublets as introduced in Section 1.1.1. Subscripts U and D indicate up and down-

---

[1] The prospects for finding *CP* violation in the $SU(3)_C$ gauge group will be neglected in this introduction.

type fields, respectively. The $G_{U(D)}$ matrices are the couplings between up(down)-type right-handed singlets and left-handed doublets. The indices i, j run over the number of fermionic families (n). Fermion masses arise in the SM when the neutral Higgs field acquires a vacuum expectation value (v). This then means that the up-type and down-type fermionic mass matrices ($\mathcal{M}_{U(D)}$) are proportional to the Yukawa couplings (with the scale set according to v). In general, the couplings may be complex in equation 1.17. This is both sufficient and required for *CP* violation and forms the basis of the phenomenon in the SM.

The physical interpretation of the Lagrangian is given in terms of the mass eigenstates of the quarks (denoted with a superscript m). Diagonalisation is possible with the four unitary matrices, $T_{\kappa,\lambda}$, where $\kappa \in \{U, D\}$ & $\lambda \in \{L, R\}$. Neutral current interactions preserve flavour and mass eigenstates, hence flavour changing neutral current interactions are only possible in loop processes. This can be seen by looking at the transformation of the neutral current term

$$\overline{L}_U \gamma_\mu L_U = \overline{L}_U^m T_{U,L} \gamma_\mu T_{U,L}^\dagger L_U^m = \overline{L}_U^m \gamma_\mu L_U^m, \tag{1.18}$$

where unitarity of the T matrices has been used. For the case of charged currents, this is not the case, where the transformation yields

$$\overline{L}_U \gamma_\mu L_D = \overline{L}_U^m T_{U,L} \gamma_\mu T_{D,L}^\dagger L_D^m = \overline{L}_D^m \gamma_\mu \mathbf{V} L_D^m, \tag{1.19}$$

where $\mathbf{V} = T_{U,L} T_{D,L}^\dagger$ is the CKM matrix.

The number of independent physical parameters contained in $\mathbf{V}$ is constrained from group theory. First, a unitary $n \times n$ matrix (V) by definition satisfies the relation $V_{ij}^* V_{jk} = \delta_{ik}$ (summation convention assumed), which imposes n constraints for $i = k$ and $n^2 - n$ constraints otherwise. This leaves n^2 free parameters. The fields in the quark fields may be independently rotated through $L_{i,U}^m \rightarrow e^{-i\phi_i^U} L_{i,U}^m$, $L_{i,D}^m \rightarrow e^{-i\phi_i^D} L_{i,D}^m$, which can be absorbed in to $\mathbf{V}$, leading to

$$\mathbf{V} \rightarrow \begin{pmatrix} e^{-i\phi_1^U} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-i\phi_n^U} \end{pmatrix} \mathbf{V} \begin{pmatrix} e^{-i\phi_1^D} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-i\phi_n^D} \end{pmatrix}. \tag{1.20}$$

This transformation removes $2n - 1$ relative phases. Therefore the amount of independent parameters is $n^2 - 2n + 1 = (n-1)^2$. The unitary matrices are a subset of the orthogonal matrices. This then leads to $\frac{1}{2}n(n-1)$ rotation angles and $(n-1)^2 - \frac{1}{2}n(n-1)$ phases. For three quark families, this results in three angles and a single *CP*-violating phase.

There are an infinite number of ways to represent the CKM matrix. The parametrisation advocated by the Particle Data Group (PDG) is known as the Standard parametrisation, first introduced by Chau and Keung [8]. It is obtained through the product of three complex rotation matrices and leads to

$$\mathbf{V} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{13} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{13} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.21}$$

where δ is the *CP*-violating phase, $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$, $i < j \in \{1,2,3\}$. A more phenomenological parametrisation was given by Wolfenstein, where the standard parameters are transformed into Wolfenstein parameters according to $(\theta_{12}, \theta_{13}, \theta_{23}, \delta) \rightarrow (A, \rho, \eta, \lambda)$[2]. This is based on the hierarchy of the matrix elements. The parametrisation is based on an expansion in terms of $\lambda \sim |V_{us}| \simeq 0.22$. The parametrisations used are

$$s_{12} = \lambda, \tag{1.22}$$

$$s_{23} = A\lambda^2, \tag{1.23}$$

$$s_{13}e^{-i\delta} = A\lambda^3(\rho - i\eta). \tag{1.24}$$

This leads to a CKM matrix of the form (evaluated to the fourth order in λ)

$$\mathbf{V} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta + \frac{i}{2}\eta\lambda^2) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 - i\eta A^2\lambda^4 & A\lambda^2(1 + i\eta\lambda^2) \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.25}$$

Unitarity of the CKM matrix leads to

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \frac{V_{cd}V_{cb}^*}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0. \tag{1.26}$$

Each term in equation 1.26 forms a vector in the complex plane. The requirement that the sum is zero requires that the addition of the three vectors leads back to the origin, which leads to a triangle known as the unitarity triangle. Equation 1.26 can be re-written

---

[2] Parametrisations are defined to all orders in λ.

as

$$R_t e^{-i\beta} + R_u e^{-\gamma} = 1, \tag{1.27}$$

where $R_t = |V_{td}V_{tb}^*|/|V_{cd}V_{cb}^*|$ and $R_u = |V_{ud}V_{ub}^*|/|V_{cd}V_{cb}^*|$ are the two non-trivial sides of the unitarity triangle; $\beta = \arg\left(-\frac{V_{cd}|V_{cb}|^*}{V_{td}V_{tb}^*}\right)$ and $\gamma = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right)$ are two of the angles of the unitarity triangle. The third side has been normalised to 1 and the third angle (α) is given through $\alpha = \pi - \beta - \gamma$. The areas of all unitarity triangles before normalisation are given by $\mathcal{J}/2$, where $\mathcal{J}$ is the Jarlskog invariant given through the relation [7, 9]

$$\Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*) = \mathcal{J}\sum_{m,n}\varepsilon_{ikm}\varepsilon_{jln}. \tag{1.28}$$

The Jarlskog invariant is a measure of *CP* violation and has an experimental value of $(3.08 \pm 0.17) \times 10^{-5}$ [10]. As there are only four independent parameters describing the CKM matrix, the six different unitarity triangles, from the six independent unitarity relations, are highly correlated. The unitarity triangle given in equation 1.27 is usually used to describe the status of global fits to the CKM hypothesis. The apex of the unitarity triangle, located at ($\overline{\rho}$,$\overline{\eta}$), is defined though

$$\overline{\rho} \equiv \Re\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.29}$$

$$\overline{\eta} \equiv \Im\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.30}$$

such that

$$\overline{\rho} + i\overline{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \equiv 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = \frac{\sqrt{1-\lambda^2}(\rho + i\eta)}{\sqrt{1-A^2\lambda^4} + \sqrt{1-\lambda^2}A^2\lambda^4(\rho + i\eta)}. \tag{1.31}$$

The current status of global fits [11] is displayed in Figure 1.2 alongside a diagram of the parameters that define the triangle. The yellow shaded area enclosed with a red contour depicts the 95.45 % confidence region of the global fit. The global fit uses the results of both experimental and theoretical work, including perturbative and lattice methods, in order to obtain the greatest accuracy possible for the four independent parameters of the CKM matrix. The parameters that are used in the global fit are [11]:

- $|V_{ud}|$: Measured from the lifetime of $0^+ \rightarrow 0^+$ nuclear β decays.

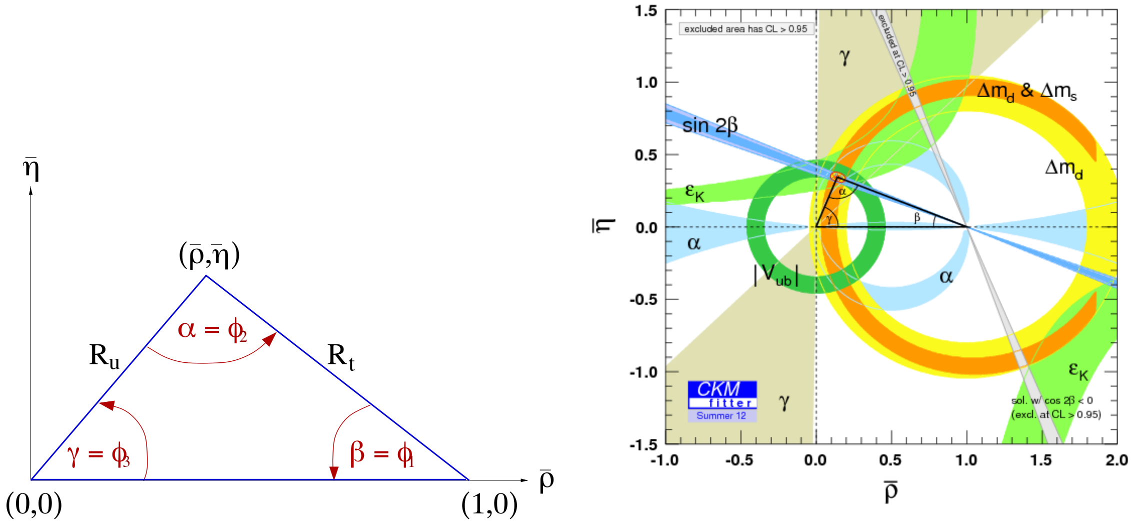


Figure 1.2.: Diagram of the parameters defining the unitarity triangle (left) alongside the current status of the global CKM fit performed by the CKMfitter group [11] (right).

- $|V_{us}|$: Measured from kaon leptonic decays $K^- \to e^-\overline{\nu}_e$, $K^- \to \mu^-\overline{\nu}_\mu$, $K^- \to \pi^0 l^- \overline{\nu}$ and the decay $\tau^- \to K^-\nu_\tau$.
- $|V_{us}|/|V_{ud}|$: Measured from the measured branching ratio ratios $\mathcal{B}(K^- \to \mu^-\nu_\mu)/\mathcal{B}(\pi^- \to \mu^-\nu_\mu)$ and $\mathcal{B}(\tau^- \to K^-\nu_\tau)/\mathcal{B}(\tau^- \to \pi^-\nu_\tau)$.
- $|V_{cd}|$: Measured from the branching ratio $\mathcal{B}(D^- \to \mu^-\overline{\nu}_\mu)$.
- $|V_{cs}|$: Measured from branching ratios of $\mathcal{B}(D_s^- \to \tau^-\overline{\nu}_\tau)$ and $\mathcal{B}(D_s^- \to \mu^-\overline{\nu}_\mu)$, where theory values of the D_s^+ fragmentation probability, $f_{D_s^+}$, are also needed.
- $|V_{ub}|$: Measured from inclusive semi-leptonic decays of B mesons and the branching ratio measurement of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$.
- $|V_{cd}|$: Measured from the inclusive semi-leptonic decays of charm mesons.
- $|V_{cb}|$: Measured mainly from knowledge of the branching ratio $\mathcal{B}(\overline{B} \to D^+\mu^-\overline{\nu}_\mu)$ in association with the form factor at zero recoil computed on the lattice.
- α: Measured from the branching ratios and CP asymmetries in decays of the form $B^0 \to \pi^+\pi^-$, $B^0 \to \rho^+\rho^-$ and $B^0 \to \rho^+\pi^-$.
- β: Measured directly from the CP asymmetry in $B^0 \to J/\psi\, K_{\mathrm{S}}^0$.

- γ: Measured from asymmetries in $B^0 \to D^* K^*$ where the final state is chosen such that both D^* and $\overline{D}^*$ mesons are allowed to contribute and thus interfere. Three popular methods are chosen to take advantage of the interference. These are the GLW method [12] in which the D meson is reconstructed as a *CP* eigenstate, the ADS method [13], where it is found in a suppressed final state or the Dalitz method, where it is found in a self-conjugate three body final state.
- $V_{tq}^* V_{tq'}$: Measured from the B_q-$\overline{B}_q$ oscillation frequencies denoted by Δm_s and Δm_d, respectively. The former is determined from $B_s^0 \to D_s^+ \pi^-$ decays and the latter from $B^0 \to D^+ \pi^-$ decays.
- $V_{tq}^* V_{tq'}, V_{cq}^* V_{cq'}$: Measured from *CP* violation in neutral kaon mixing (ε_K).

1.3. Phenomenology of B_s^0 mixing

The Schrödinger equation describes the time evolution, including mixing and decay of B_s^0 mesons (denoted $B(t)$), and is given by

$$i\hbar \frac{\partial}{\partial t} |B(t)\rangle = \left(\mathbf{M} - \frac{i}{2} \mathbf{\Gamma} \right) |B(t)\rangle, \tag{1.32}$$

where $\mathbf{M}$ and $\mathbf{\Gamma}$ denote the mass and decay matrices, respectively. A natural solution to equation 1.32 is given by

$$|B_j(t)\rangle = \exp\left(-i \left(M_j - \frac{i}{2} \Gamma_j \right) \right) |B_j(0)\rangle, \tag{1.33}$$

where j refers to either of the two mass eigenstates. The relationship between the flavour eigenstates and the mass eigenstates ($|B_H\rangle$ and $|B_L\rangle$) can in general be written as

$$|B_H\rangle = p|B_s^0\rangle + q|\overline{B}_s^0\rangle, \tag{1.34}$$

$$|B_L\rangle = p|B_s^0\rangle - q|\overline{B}_s^0\rangle, \tag{1.35}$$

where p and q are complex numbers and the absence of a time argument implies the state at $t = 0$. It is simple to invert equations 1.34 and 1.35, and by substituting the solution of equation 1.33 into the Schrödinger equation, to arrive at the general equations for the

Term	$\vert B_s^0(t)\rangle$	$\vert\overline{B}_s^0(t)\rangle$
a	$\vert A(f)\vert^2[\frac{1}{2}(1+\vert\frac{q}{p}\overline{\rho}(f)\vert^2)+\Re(\frac{q}{p}\overline{\rho}(f))]$	$\vert\overline{A}(f)\vert^2[\frac{1}{2}(1+\vert\frac{p}{q}\rho(f)\vert^2)+\Re(\frac{p}{q}\rho(f))]$
b	$\vert A(f)\vert^2[\frac{1}{2}(1+\vert\frac{q}{p}\overline{\rho}(f)\vert^2)-\Re(\frac{q}{p}\overline{\rho}(f))]$	$\vert\overline{A}(f)\vert^2[\frac{1}{2}(1+\vert\frac{p}{q}\rho(f)\vert^2)-\Re(\frac{p}{q}\rho(f))]$
c	$\vert A(f)\vert^2(1-\vert\frac{q}{p}\overline{\rho}(f)\vert^2)$	$\vert\overline{A}(f)\vert^2(1-\vert\frac{p}{q}\rho(f)\vert^2)$
d	$-2\vert A(f)\vert^2\Im(\frac{q}{p}\overline{\rho}(f))$	$-2\vert\overline{A}(f)\vert^2\Im(\frac{p}{q}\rho(f))$

Table 1.1.: Terms describing the time dependent decay rate in equation 1.43.

time evolution of the flavour eigenstates $|B_s^0(t)\rangle$ and *CP* conjugate $|\overline{B_s^0}(t)\rangle$

$$|B_s^0(t)\rangle = g_+(t)|B_s^0\rangle + \frac{q}{p}g_-(t)|\overline{B}_s^0\rangle, \tag{1.36}$$

$$|\overline{B}_s^0(t)\rangle = \frac{p}{q}g_-(t)|B_s^0\rangle + g_+(t)|\overline{B}_s^0\rangle, \tag{1.37}$$

where

$$g_\pm(t) = \frac{1}{2}\left(e^{-(iM_L+\Gamma_\mathrm{L}/2)t} \pm e^{-(iM_H+\Gamma_\mathrm{H}/2)t}\right) \tag{1.38}$$

$$= \frac{1}{2}e^{-iM_L t}e^{-\frac{1}{2}\Gamma_\mathrm{L}t}[1 \pm e^{-i\Delta m_s t}e^{-\frac{1}{2}\Gamma_s t}]. \tag{1.39}$$

The symbols $M_{H(L)}$ and $\Gamma_{H(L)}$ refer to the masses and decay rates of the heavy (light) mass eigenstates, respectively, and Δm_s is the B_s^0-$\overline{B}_s^0$ oscillation frequency. The average decay rate, Γ_s, and decay rate difference, $\Delta\Gamma_s$, are defined through

$$\Gamma_s = \frac{1}{2}(\Gamma_\mathrm{L} + \Gamma_\mathrm{H}), \tag{1.40}$$

$$\Delta\Gamma_s = \Gamma_\mathrm{L} - \Gamma_\mathrm{H}. \tag{1.41}$$

The decay amplitudes of the states B_s^0 and $\overline{B}_s^0$ to a final state f, denoted $A(f)$ and $\overline{A}(f)$, respectively, lead to the expression

$$\Gamma(\overset{(-)}{B_s^0}(t) \to f) = |\langle f|H_{\Delta B=1}|\overset{(-)}{B_s^0}(t)\rangle|^2 \propto \frac{1}{2}e^{-\Gamma_\mathrm{L}t}\cdot \overset{(-)}{G}_f(t), \tag{1.42}$$

where $H_{\Delta B=1}$ is the Hamiltonian of the decay and

$$\overset{(-)}{G}_f = \overset{(-)}{a} + \overset{(-)}{b}e^{\Delta\Gamma_s t} + \overset{(-)}{c}e^{\Delta\Gamma_s t/2} + \overset{(-)}{d}e^{\Delta\Gamma_s t/2}\sin\Delta m_s t. \tag{1.43}$$

The $(\overset{(-)}{a}, \overset{(-)}{b}, \overset{(-)}{c}, \overset{(-)}{d})$ terms are given in Table 1.1, where $\overline{\rho}(f) \equiv \overline{A}(f)/A(f)$ and $\rho(f) \equiv$

$A(f)/\overline{A}(f)$. It can be seen from equation 1.42, that $G_f(t) \neq \overline{G}_f(t)$ implies CP is violated. In the absence of mixing, CP may only be violated if the decay amplitudes differ, i.e. $|A(f)| \neq |\overline{A}(f)|$, known as direct CP violation. For the case of non-zero mixing and decays to a common final state, it is useful to introduce a complex parameter, λ_f, defined through

$$\lambda_f \equiv \eta_{CP} \frac{q}{p} \frac{\overline{A}_f}{A_f}, \tag{1.44}$$

where $\eta_{CP} = +1(-1)$ for the case that the final state is CP-even (CP-odd). On deriving the form of the CP-violating asymmetry

$$\frac{\Gamma(B_s^0 \to f) - \Gamma(\overline{B}_s^0 \to f)}{\Gamma(B_s^0 \to f) + \Gamma(\overline{B}_s^0 \to f)}, \tag{1.45}$$

using equation 1.42 under the assumption that $|\lambda_f| = 1$, the time-dependent decay rates become

$$\begin{aligned}\Gamma(B_s^0 \to f) \propto\ & 1 + e^{\Delta\Gamma_s t} + (1 - e^{\Delta\Gamma_s t}) \cos(\phi_{\Delta B=1} + \phi_{\Delta B=2}) \\ & - 2e^{\frac{1}{2}\Delta\Gamma_s t} \sin(\Delta m_s t) \sin(\phi_{\Delta B=1} + \phi_{\Delta B=2}), \qquad (1.46)\\ \Gamma(\overline{B}_s^0 \to f) \propto\ & 1 + e^{\Delta\Gamma_s t} + (1 - e^{\Delta\Gamma_s t}) \cos(\phi_{\Delta B=1} + \phi_{\Delta B=2}) \\ & + 2e^{\frac{1}{2}\Delta\Gamma_s t} \sin(\Delta m_s t) \sin(\phi_{\Delta B=1} + \phi_{\Delta B=2}), \qquad (1.47)\end{aligned}$$

where $\overline{\rho} = e^{i\phi_{\Delta B=1}}$ and $q/p = e^{i\phi_{\Delta B=2}}$ and $\phi_{\Delta B=1}$ and $\phi_{\Delta B=2}$ denote CP-violating phases in the decay and mixing, respectively. This leads to an asymmetry of the form

$$\frac{-2e^{\frac{1}{2}\Delta\Gamma_s t} \sin(\Delta m_s t) \sin(\phi_{\Delta B=1} + \phi_{\Delta B=2})}{1 + e^{\Delta\Gamma_s t} + (1 - e^{\Delta\Gamma_s t}) \cos(\phi_{\Delta B=1} + \phi_{\Delta B=2})}. \tag{1.48}$$

It can be seen from equation 1.48 that in order to have observable indirect CP violation, two conditions must be met. These are that the mixing oscillations are present to generate a non-zero Δm_s and that the two CP-violating phases in the mixing and decay do not cancel, i.e. $\phi_{\Delta B=1} + \phi_{\Delta B=2} \neq 0$.

1.4. $b \longrightarrow s$ Transitions

While the phenomenology of mixing-induced *CP* violation as explained in Section 1.3 is relatively straight-forward and elegant, the predictions of *CP* violating parameters in specific physics scenarios are much more complex. This will be discussed for two flavour changing neutral current (FCNC) $b \to s$ transitions that form the basis of measurements in the $B_s^0 \to \phi\phi$ decay.

1.4.1. B_s^0-$\overline{B}_s^0$ Mixing

1.4.1.1. Standard Model B_s^0-$\overline{B}_s^0$ Mixing

It is useful to revisit equation 1.32 with the substitution $\mathbf{\Sigma} = \mathbf{M} - \mathbf{i}\mathbf{\Gamma}/\mathbf{2}$. As a consequence that both $\mathbf{M}$ and $\mathbf{\Gamma}$ are Hermitian, the off-diagonal elements, M_{12} and Γ_{12} can be found through looking at the dispersive and absorptive parts of $\mathbf{\Sigma}$, respectively. Through rearrangement of the characteristic equation, we know that

$$(\sigma_H - \sigma_L)^2 = 4\Sigma_{12}\Sigma_{21} \tag{1.49}$$

$$\Rightarrow \left(\Delta m_s + i\frac{\Delta\Gamma_s}{2}\right)^2 = 4\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)\left(M_{12}^* - i\frac{\Gamma_{12}}{2}^*\right), \tag{1.50}$$

where $\sigma_{L(H)} = M_{L(H)} - i\frac{\Gamma_{L(H)}}{2}$ are the eigenvalues derived in Section 1.3. Taking the real and imaginary parts of equation 1.50 yields the simultaneous equations

$$\begin{aligned}(\Delta m_s)^2 - \frac{(\Delta\Gamma_s)^2}{4} &= 4(|M_{12}|^2 - |\Gamma_{12}|^2) + 2\Im(\Gamma_{12}M_{12}^* + M_{12}\Gamma_{12}^*) \\ &= 4(|M_{12}|^2 - |\Gamma_{12}|^2) \qquad (1.51)\\ \Delta m_s \Delta\Gamma_s &= -4\Re\left(\frac{\Gamma_{12}M_{12}^*}{2} + \frac{\Gamma_{12}^*M_{12}}{2}\right) \\ &= 4|M_{12}||\Gamma_{12}|\cos\phi_s, \qquad (1.52)\end{aligned}$$

where in the last line, the definition $\phi_s \equiv \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right)$ has been used. Combining the two equations yields

$$(\Delta m_s)^2 - \frac{(\Delta\Gamma_s)^2}{4} = 4|M_{12}|^2 - \frac{(\Delta m_s)^2(\Delta\Gamma_s)^2}{16|M_{12}|^2\cos^2\phi_s}. \tag{1.53}$$

On noting that it is experimentally well known that $(\Delta m_s)^2 >> (\Delta\Gamma_s)^2$, the mixing parameters Δm_s and $\Delta\Gamma_s$ take the form

$$\Delta m_s \approx 2|M_{12}|, \tag{1.54}$$

$$\Delta\Gamma_s \approx 2|\Gamma_{12}|\cos\phi_s. \tag{1.55}$$

From equations 1.34 and 1.35, it can be seen that the matrix needed to diagonalise $\mathbf{\Sigma}$ is given by

$$\mathbf{Q} = \begin{pmatrix} p & p \\ q & -q \end{pmatrix} \tag{1.56}$$

such that $[\mathbf{Q}^{-1}\mathbf{\Sigma}\mathbf{Q}]_{21} = [\mathbf{Q}^{-1}\mathbf{\Sigma}\mathbf{Q}]_{12} = 0$. Explicitly, this means

$$\frac{q}{p} = -\frac{\Delta m_s + i\Delta\Gamma_s/2}{2M_{12} - i\Gamma_{12}} = -\frac{2M_{12}^* - i\Gamma_{12}^*}{\Delta m_s + i\Delta\Gamma_s/2}. \tag{1.57}$$

This can then be manipulated to give

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^*}{M_{12}}\left(\frac{1 + i|\Gamma_{12}/(2M_{12})|e^{i\phi_s}}{1 + i|\Gamma_{12}/(2M_{12})|e^{-i\phi_s}}\right). \tag{1.58}$$

It is known experimentally that $|\Gamma_{12}/(2M_{12})|$ is small and can therefore be used as an expansion parameter. Performing this expansion gives the result

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^*}{M_{12}}(1 - |\Gamma_{12}/M_{12}|\sin\phi_s) + \mathcal{O}(|\Gamma_{12}/(2M_{12})|^2). \tag{1.59}$$

The magnitude $|q/p|^2$ is given by

$$|q/p|^2 = 1 - |\Gamma_{12}/M_{12}|\sin\phi_s \equiv 1 - a_s, \tag{1.60}$$

where the factor a_s is used to parameterise *CP* violation in B_s^0 mixing.

The *CP*-violating phase in B_s^0-$\overline{B}_s^0$ mixing is found through looking at the phase of $q/p = -M_{12}^*/|M_{12}|[1 + \mathcal{O}(a_s)]$. In the Standard Model, the leading order Feynman diagram that contributes to B_s^0-$\overline{B}_s^0$ mixing is the box diagram shown in Figure 1.3. This means that *CP* violation is found through the phase of $-(V_{tb}^* V_{ts})/(V_{tb} V_{ts}^*)$ as the SM box diagram is dominated by the top quarks. At the time of writing the SM predictions

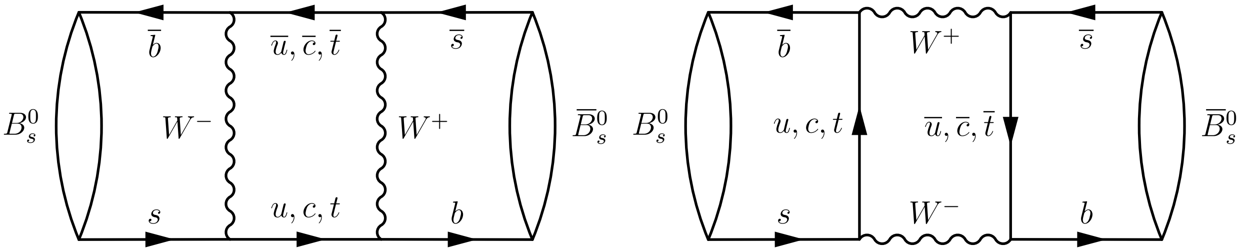


Figure 1.3.: Feynman diagrams contributing to B_s^0-$\overline{B}_s^0$ mixing.

of the *CP*-violating phase in B_s^0 mixing, ϕ_s^{SM}, and the *CP* asymmetry, a_s^{SM}, have been calculated by Lenz and Nierste (2011) [14] to be

$$\phi_s^{SM} = 0.0038 \pm 0.0010 \text{ rad}, \tag{1.61}$$

$$a_s^{SM} = (1.9 \pm 0.3) \cdot 10^{-5}. \tag{1.62}$$

The theoretical calculation of the SM box diagram is explained in more detail in Appendix C.

1.4.1.2. Prospects for New Physics in B_s^0-$\overline{B}_s^0$ Mixing

The phenomenology of B_s^0-$\overline{B}_s^0$ mixing allows for New Physics (NP) extensions to be parametrised in a largely model-independent way. Through the inclusion of the NP complex parameter, Δ_s, the description of B_s^0-$\overline{B}_s^0$ mixing can be extended to

$$M_{12} \rightarrow M_{12}^{SM} \cdot \Delta_s, \tag{1.63}$$

where $\Delta_s \equiv |\Delta_s| e^{i\phi_s^\Delta}$. This means that the *CP*-violating phase in B_s^0-$\overline{B}_s^0$ mixing undergoes the simple extension of $\phi_s \rightarrow \phi_s^{SM} + \phi_s^\Delta$. In the SM, $|\Delta_s| = 1$ and $\phi_s^\Delta = 0$.

Experimentally, the SM is tested through measurements in $B_s^0 \rightarrow J/\psi K^+K^-$ and $B_s^0 \rightarrow J/\psi \pi^+\pi^-$ decays. As a tree level process, the *CP*-violating phase measured in $b \rightarrow sc\bar{c}$ transitions, $-2\beta_s$, is given through $\arg(V_{tb}V_{ts}^*/V_{cb}V_{cs}^*)$, where penguin contributions are neglected. This is a measurement of *CP* violation in mixing as the imaginary component of $V_{cb}V_{cs}^*$ is suppressed to $\mathcal{O}(\lambda^4)$ in the Wolfenstein parametrisation. The combined recent LHCb measurement of the *CP*-violating phase in $B_s^0 \rightarrow J/\psi K^+K^-$ and $B_s^0 \rightarrow J/\psi \pi^+\pi^-$ decays is found to be $\phi_s = 0.01 \pm 0.07(\text{stat}) \pm 0.01(\text{syst})$ rad [15]. The 68 %, 95 % and 99 % confidence levels are shown in Figure 1.4 along with the SM prediction for the fit to $B_s^0 \rightarrow J/\psi K^+K^-$ events. It can be seen that the measured result

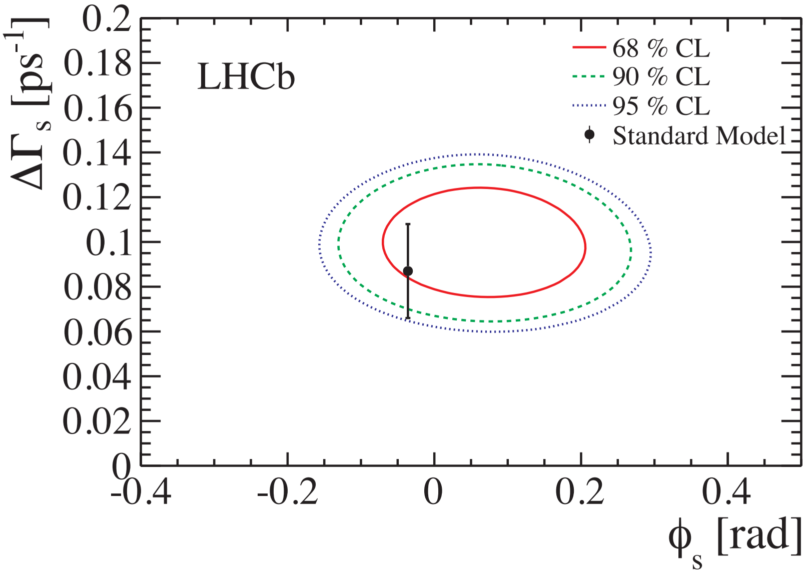


Figure 1.4.: LHCb Measurement of the *CP*-violating phase, ϕ_s vs. $\Delta\Gamma_s$ in $B_s^0 \to J/\psi\, K^+K^-$ decays [15].

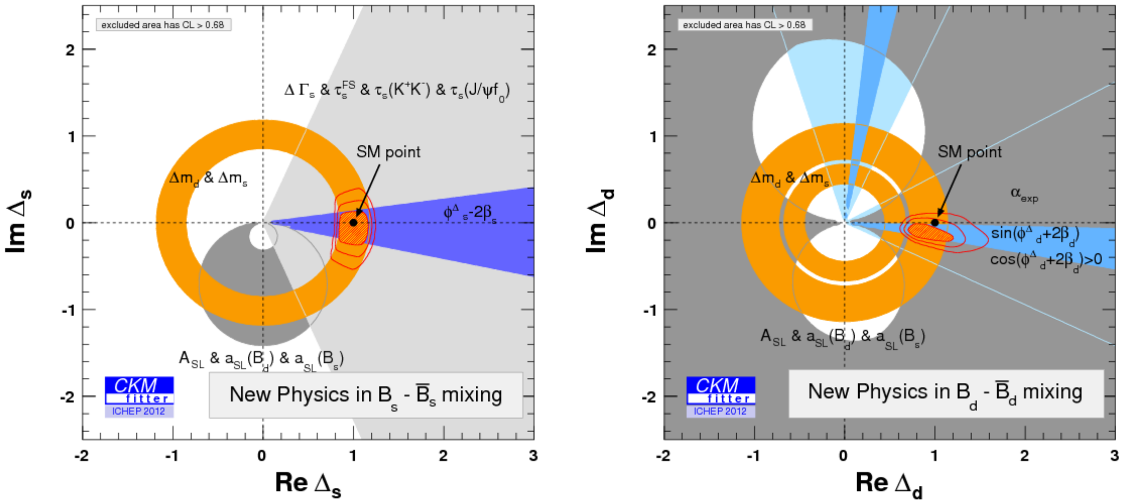


Figure 1.5.: Allowed regions, bounded by the 68 % CL, of Δ_s (left) and Δ_d (right) from the global CKMfitter group [16]. The shaded areas show allowed regions from individual constraints to the global fit. The red hatched area shows the 68% CL for the global fit with all constraints included. Additional red contours show the 2σ and 3σ confidence regions.

is in good agreement with the theoretical prediction and therefore puts strong constraints on NP.

As can be seen from this Chapter, there are a large number of observables that can be measured in the decays of B_s^0 mesons, along with equivalent observables in the decays of B^0 mesons. This therefore means that there are large numbers of experimental inputs that can only be properly put in to context with the help of a global fit to all available data. This fit has most recently been performed by the CKMFitter Group in collaboration with Lenz and Nierste [17]. In addition to the *CP*-violating phase and B_s^0 physical eigenstate lifetimes measured in $b \to sc\bar{c}$ transitions, the global fit also uses other experimental and theoretical inputs. The main constraining inputs to the Δ_s parameter include the B^0 and B_s^0 oscillation frequencies, the lifetimes of the physical eigenstates in the B_s^0 system, the *CP* asymmetries in semileptonic B^0 and B_s^0 decays, denoted by a_{SL}^d and a_{SL}^s, respectively, in addition to the combined semileptonic *CP* asymmetry, A_{SL}. Figure 1.5 shows the constraining power of the result of the global fit, where Δ_s and the corresponding parameter in the B^0 system, Δ_d, have been allowed to vary independently. As can be seen from the left plot of Figure 1.5, the global fit is entirely consistent with the SM value, which is by definition located at the coordinate $(\Re(\Delta_s) = 1, \Im(\Delta_s) = 0)$. However, a conflict appears in the constraints from measurements of A_{SL} (dark grey circle) and $-2\beta_s$ (blue cone).

Many extensions to the SM have been proposed in which *CP* violation is described by the same CKM mechanism as is present in the SM. The interested reader is directed to the work of Bigi (2009) [18] for examples consisting of the popular theory of Supersymmetry and theories with extra spatial dimensions. The class of theories that use the so-called minimal flavour violation (MFV) hypothesis [19] have become popular largely due to the success of the SM in describing experimental measurements, for example the measurement of the *CP*-violating phase in $B_s^0 \to J/\psi K^+K^-$ decays.

The global fit shows the advantages of the model independent parameterisation shown in equation 1.63, where the global fit can with minimal assumptions on the Δ_s parameter, encompass theories that impose MFV.

1.4.2. The $B_s^0 \to \phi\phi$ Decay

As an example of a flavour changing neutral current interaction, the leading order contributions to the $B_s^0 \to \phi\phi$ decay in the Standard Model are from penguin diagrams, shown in Figure 1.6.

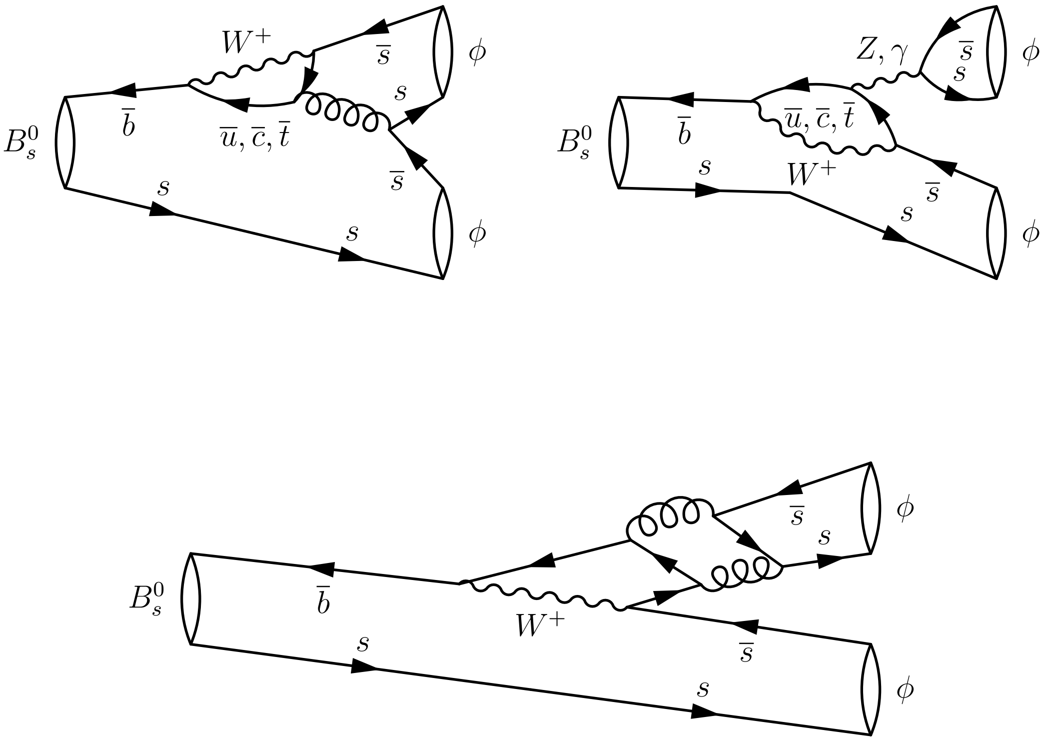


Figure 1.6.: Feynman diagrams contributing to the $B_s^0 \to \phi\phi$ decay, consisting of a gluonic penguin (top-left), electroweak penguin (top-right), and a 2-loop gluonic penguin (bottom).

The $B_s^0 \to \phi\phi$ decay is an example of a pseudo-scalar to vector vector ($P \to VV$) process. This means that there is a mixture of polarisation amplitudes in the final state. These are the *CP*-even longitudinal (A_0), the *CP*-odd transverse ($A_\perp$) and the *CP*-even transverse ($A_\parallel$) polarisations. In order to measure *CP* violation in this decay, the polarisations, along with their associated interferences must be disentangled by measuring decay angle distributions. A popular choice for the $B_s^0 \to \phi\phi$ decay is the helicity basis (explained in detail in Figure 1.7). The need for the measurement of the decay angles complicates the determination of the *CP*-violating phase in the interference between mixing and decay and therefore results in a more complex expression than that given in equation 1.48. The derivation of the form is explained in greater detail in Appendix A.

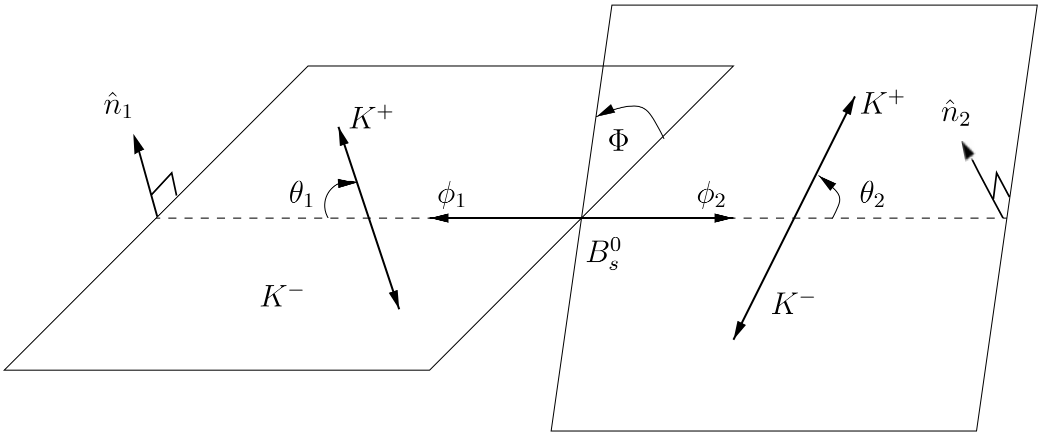


Figure 1.7.: Decay angles for the $B_s^0 \to \phi\phi$ decay, where the K^+ momentum in the $\phi_{1,2}$ rest frame, and the parent $\phi_{1,2}$ momentum in the rest frame of the B_s^0 meson, $\theta_{1,2}$ is the angle between the K^+ track momentum in the $\phi_{1,2}$ meson rest frame and the parent $\phi_{1,2}$ momentum in the B_s^0 rest frame, Φ is the angle between the two ϕ meson decay planes and $n^{1,2}$ is the unit vector normal to the decay plane of the $\phi_{1,2}$ meson.

For each polarisation state, h, a generic penguin amplitude may be written as

$$A_h = V_{tb}^* V_{ts} P_{t,h} + V_{cb}^* V_{cs} P_{c,h} + V_{ub}^* V_{us} P_{u,h}, \tag{1.64}$$

where $P_{j,h}$ is a penguin amplitude for a quark of flavour j in the loop transition. The CKM unitarity relation $V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0$ can be used to write equation 1.64 as

$$\begin{aligned} A_h &= V_{tb}^* V_{ts} P_{tc,h} + V_{ub}^* V_{us} P_{uc,h} \\ &= |V_{tb}V_{ts}| e^{-i\beta_s} P_{tc,h} + |V_{ub}V_{us}| e^{i\gamma} P_{uc,h}, \end{aligned} \tag{1.65}$$

where $P_{jk,h} \equiv P_{j,h} - P_{k,h}$, and the CKM phases, β_s and γ, have been explicitly written. If the penguin amplitudes are redefined to include the CKM matrix elements and with strong phases factored out, then equation 1.65 can be written as

$$A_h = e^{-i\beta_s}[e^{i\delta_{tc,h}} P_{tc,h} + e^{i\delta_{uc,h}} e^{i(\gamma+\beta_s)} P_{uc,h}], \tag{1.66}$$

where $\delta_{ij,h}$ can be understood as the CP-conserving strong phase corresponding to the $P_{ij,h}$ amplitude. Equation 1.66 can be re-written as [20]

$$A_h = e^{-i\beta_s} e^{i\delta_{tc,h}} P_{tc,h}[1 + e^{i(\gamma+\beta_s)} e^{i\Delta_h} R_h], \tag{1.67}$$

where

$$R_h \equiv \frac{P_{uc,h}}{P_{tc,h}}, \tag{1.68}$$

$$\Delta_h \equiv \delta_{uc,h} - \delta_{tc,h}. \tag{1.69}$$

Historically, amplitudes are parameterised using the functional form $A_h = |A_h|e^{i\phi}e^{i\delta_h}$, where ϕ is the weak phase and δ is the strong phase. It can be seen that the amplitude in equation 1.67 takes this form in two situations:

1. R_h is helicity-independent and $\Delta_h = 0$.
2. $R_h = 0\,\forall h \in \{0, \perp, \|\}$, which would correspond to an amplitude with no *CP*-violating phase present.

1.4.2.1. The $B_s^0 \to \phi\phi$ Decay in the Standard Model

Various approaches are used to calculate the form of the penguin amplitudes, P, shown in equation 1.64. These include perturbative QCD (pQCD) [21] and QCD factorisation (QCDF) [22]. Throughout this dissertation, the methods of QCDF will be discussed in detail due to the success [23] and the comparative elegance of the method.

The theoretical prediction for *CP* violation in the $B_s^0 \to \phi\phi$ decay in the SM is known to be hard to calculate, due to the non-factorisable nature of the transverse polarisation states. This results in approximations being introduced to allow for a quantitative analysis. It has been shown by Bartsch *et al.* (2008) [22], that $B_s^0 \to \phi\phi$ decays to the longitudinal polarisation can be calculated, where weak annihilation effects are suppressed by factors of Λ_{QCD}/m_b.

For the case of the SM prediction of *CP* violation in $B_s^0 \to \phi\phi$, it is useful to evaluate the expression in equation 1.45 as

$$\frac{\Gamma(B_s^0 \to \phi_L\phi_L) - \Gamma(\overline{B}_s^0 \to \phi_L\phi_L)}{\Gamma(B_s^0 \to \phi_L\phi_L) + \Gamma(\overline{B}_s^0 \to \phi_L\phi_L)} = S_\phi \sin(\Delta m_s t) - C_\phi \cos(\Delta m_s t). \tag{1.70}$$

The coefficients S_ϕ and C_ϕ take the form

$$S_\phi = 2\lambda^2\eta\Re\left(\frac{a^c - a^u}{a^c}\right), \tag{1.71}$$

$$C_\phi = 2\lambda^2\eta\Im\left(\frac{a^c - a^u}{a^c}\right), \tag{1.72}$$

where λ and η are Wolfenstein parameters. The symbols a^p represent coefficients of the $B_s^0 \to \phi_L\phi_L$ penguin amplitudes in the framework of QCDF, where p is the quark propagator in the penguin loop. This is described in detail in Appendix D. The term S_ϕ is comparable to the coefficient of $\sin(\Delta m_s t)$ in equation 1.48. The difference $a^c - a^u$ is calculable in the framework of QCDF and has the advantage of the cancellation of annihilation contributions. The evaluation of Bartsch *et al.* leads to

$$|a^c - a^u|/\,\mathrm{GeV}^3 = 0.057^{+0.023}_{-0.029}, \tag{1.73}$$

where errors arise from form factor uncertainties, scale uncertainties and the uncertainty on the mass of the charm quark. The calculation of a^c requires experimental input. Bartsch *et al.* [22] have shown that the expression for the branching ratio may be written as

$$\mathcal{B}(B_s^0 \to \phi_L\phi_L) = \frac{\tau_{B_s^0}\mathrm{G}_F^2|V_{cb}V_{cs}^*|^2}{64\pi m_{B_s^0}}|a^c|^2, \tag{1.74}$$

which can be re-arranged to give

$$|a^c| = 0.177\,\mathrm{GeV}^3\left(\frac{\mathcal{B}(B_s^0 \to \phi_L\phi_L)}{15\cdot 10^{-6}}\right)^{1/2}\left(\frac{1.53\ \mathrm{ps}}{\tau_{B_s^0}}\right)^{1/2}. \tag{1.75}$$

At the time of the publication of Bartsch *et al.* (2008) [22], the CDF collaboration had announced the first evidence for the $B_s^0 \to \phi\phi$ decay [24] and provided a branching ratio measurement of $\mathcal{B}(B_s^0 \to \phi\phi) = (14^{+6}_{-5}(\mathrm{stat}) \pm 6(\mathrm{syst})) \times 10^{-6}$, based on eight $B_s^0 \to \phi\phi$ events. This did not allow for an amplitude analysis to find the longitudinal polarisation fraction.

Assuming that direct *CP* violation is negligible, S_ϕ essentially equates to the sine of the *CP*-violating phase in the interference between mixing and decay in $b \to s s\bar{s}$

transitions. An upper limit on S_ϕ is found through the inequality

$$0 \leq S_\phi \leq 2\lambda^2\eta\frac{|a^c - a^u|}{|a^c|} \leq \lambda^2\eta\left(\frac{\mathcal{B}(B_s^0 \rightarrow \phi_L\phi_L)}{15 \cdot 10^{-6}}\right)^{-1/2} \sim 0.02. \tag{1.76}$$

This therefore shows that *CP* violation in the interference between B_s^0 mixing and the $B_s^0 \rightarrow \phi\phi$ decay is expected to be small in the SM.

1.4.2.2. Prospects for New Physics in Penguin Decays

Due to the suppressed branching fraction of penguin decays, larger data samples must be collected in order to achieve the same precision as that obtained in tree level decays, such as $B_s^0 \rightarrow J/\psi\,\phi$. It is mainly for this reason that penguin decays remain largely unexplored experimentally.

The $B^0 \rightarrow \phi K_\mathrm{S}^0$ decay represents another example of a $b \rightarrow s\bar{s}s$ transition and contains similar physics as is present in the decay of $B_s^0 \rightarrow \phi\phi$ (neglecting the effects of neutral meson mixing). In 2005, the BaBar collaboration presented a measurement of the time-dependent asymmetry in $B^0 \rightarrow \phi K_\mathrm{S}^0$ [25]. The study resulted in the measurement of a time-dependent *CP* asymmetry of $\sin(2\beta_s^{eff}) = 0.50 \pm 0.25^{+0.07}_{-0.04}$. This followed an earlier measurement by the Belle collaboration with considerably larger experimental uncertainties [26]. The BaBar result of the asymmetry as a function of decay time is shown in Figure 1.8. From this result, it can be concluded that there is a possibility of a sizeable time-dependent asymmetry in this transition and hence a disagreement with the SM prediction shown in equation 1.76, though errors are quite large.

The CDF collaboration has performed a time-integrated study of the $B_s^0 \rightarrow \phi\phi$ decay, in which the polarisation fractions were measured in addition to the measurement of T-violating triple product asymmetries, A_U and A_V [27]. With the safe assumption of *CPT* conservation[3], T-violation implies *CP* violation. The two T-odd observables, U and V are defined through [28]

$$U = \sin(2\Phi), \tag{1.77}$$

$$V = \pm\sin(\Phi), \tag{1.78}$$

where the + sign is taken if the *CP*-even quantity $\cos\theta_1\cos\theta_2 \geq 0$, else the − sign is used. The helicity angles, θ_1, θ_2 and Φ are defined in Figure 1.7. The U and V observables

---

[3] *CPT* conservation is known to hold in all quantum field theories, of which the SM is an example.

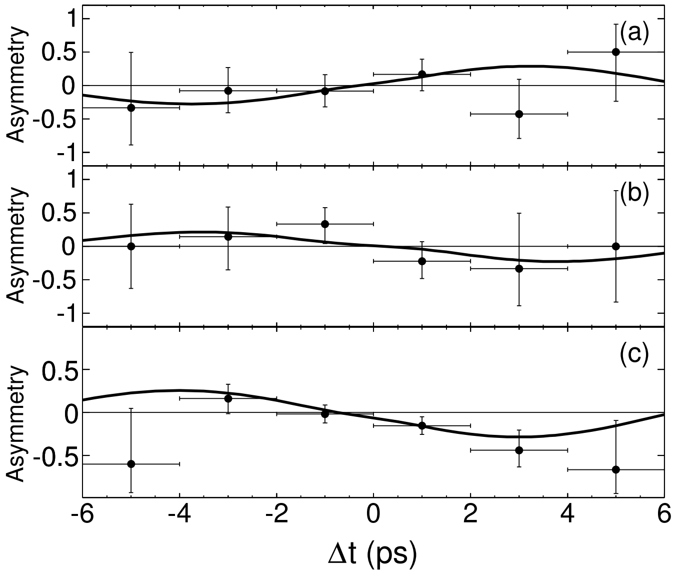


Figure 1.8.: The time-dependent *CP* asymmetry in (a) $B^0 \to \phi K_\mathrm{S}^0$ decays, (b) $B^0 \to \phi K_\mathrm{L}^0$ decays and (c) $B^0 \to K^+K^-K_\mathrm{S}^0$ decays (non-resonant) as measured by the BaBar collaboration in 2005 [25].

correspond to the T-odd triple products [28]

$$\sin\Phi = (\hat{n}_1 \times \hat{n}_2) \cdot \hat{p}_1, \tag{1.79}$$

$$\sin(2\Phi)/2 = (\hat{n}_1 \cdot \hat{n}_2)(\hat{n}_1 \times \hat{n}_2) \cdot \hat{p}_1, \tag{1.80}$$

where $\hat{n}_i$ $(i = 1, 2)$ is a unit vector perpendicular to the ϕ_i decay plane and $\hat{p}_1$ is a unit vector in the direction of the ϕ_1 momentum in the B_s^0 rest frame shown in Figure 1.7. The expressions in equations 1.79 and 1.80 show explicitly how the U and V observables are T-odd. The exact relation to the *CP*-violating observables in the time-dependent method is given in Appendix B. The power of triple product asymmetries is the ability to probe *CP* violation in the interference between mixing and decay with a time-integrated method, where the initial flavour of the B_s^0 meson does not need to be known. The triple product asymmetries were measured by the CDF collaboration to be

$$A_U = -0.007 \pm 0.064 \pm 0.018, \tag{1.81}$$

$$A_V = -0.120 \pm 0.064 \pm 0.016. \tag{1.82}$$

The measurements of the triple product asymmetries from CDF, as for the case of the B-factory measurements of the time-dependent *CP* asymmetry in $B^0 \to \phi K_\mathrm{S}^0$, shows a central value that deviates from zero, but with large uncertainties.

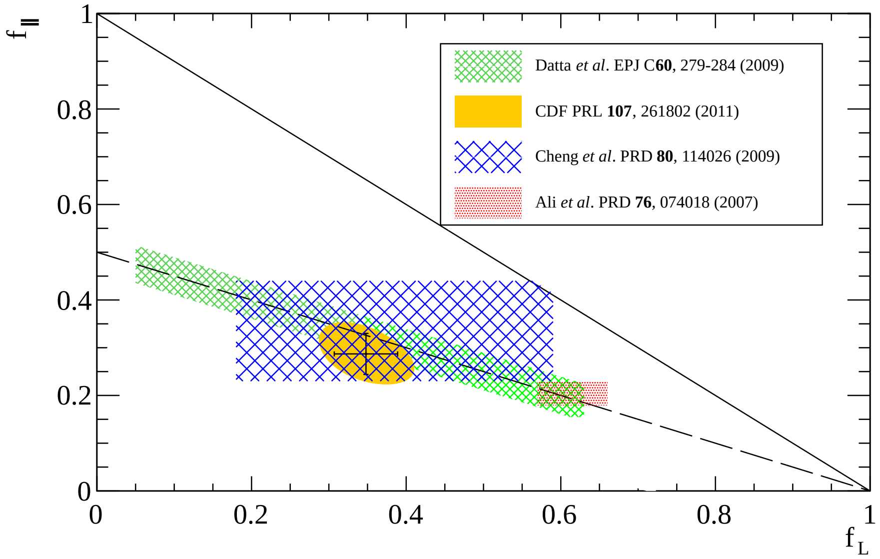


Figure 1.9.: Comparison of the CDF measurement of the polarisation fractions (solid yellow) [27], with the perturbative QCD prediction of Ali *et al.* (medium hatched red) [21], the QCDF prediction of Datta *et al.* (small hatched green) [29], and the QCDF prediction of Cheng *et al.* (large hatched blue) [30].

Measurements of the polarisation fractions would provide further tests of the SM if the SM predictions were known accurately. However, we can see from Figure 1.9 that this is not the case. The predictions from QCDF are shown to contain large errors particularly for the case of f_L, where the polarisation fractions are defined through

$$f_L = \frac{|A_0|^2}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}, \tag{1.83}$$

$$f_\parallel = \frac{|A_\parallel|^2}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}. \tag{1.84}$$

The CDF measurement of the polarisation fractions clearly favours the prediction offered by QCDF over pQCD and also finds uses as an input to theoretical predictions.

In light of possible hints of NP effects in $b \to s\bar{s}s$ transitions and large NP effects in B mixing largely ruled out in the results of global fits, models describing extensions of the SM must accommodate this difference. The recent discovery of a boson consistent with the SM Higgs boson [31, 32] has led to greater interest in an extended Higgs sector[4]. This would generate additional gauge bosons, equivalent to the W and Z bosons of the SM. Such a new boson, X, would most likely have a mass greater than the masses of the

---

[4] Indeed, the SM does not forbid additional Higgs doublets.

SM W and Z bosons, otherwise we could reasonably expect to have already confirmed the existence of such a particle. The heavy mass would allow for X to be described by an effective theory, leading to the Hamiltonian [20]

$$H_{NP} = \frac{4\mathrm{G}_F}{\sqrt{2}} \sum_{A,B=L,R} f_s^{AB} \bar{b}\gamma_A s \bar{s}\gamma_B s, \tag{1.85}$$

where $\gamma_{L(R)} = \frac{1-(+)\gamma_5}{2}$ is the left(right)-handed projection using the Dirac γ_5 matrix and f_s^{AB} allows for the coupling constants to vary for the different components of the sum. In the work of Bhattacharya *et al.* [20], it was assumed that the NP contribution to the $b \to s\bar{s}s$ transition are sub-leading with respect to the SM amplitude. This leads to the introduction of the parameter $\epsilon \equiv |NP|/|SM|$. Ignoring the small SM *CP*-violating phase and assuming NP strong phases take values 0 or π, equation 1.64 can be extended phenomenologically to account for new physics contributions through

$$A_h^{SM} = \overline{A}_h^{SM} = |A_h^{SM}|e^{i\delta_h}, \tag{1.86}$$

$$A_h^{NP} = \overline{A}_h^{NP*} = |A_h^{NP}|e^{i\phi_h}, \tag{1.87}$$

where δ and ϕ represent the usual strong and weak phases and a SM weak phase is neglected. Bhattacharya *et al.* have shown that the different scenarios resulting from the different operators (f_s^{LL}, f_s^{RR} and f_s^{LR}) reveal different patterns. For the cases of f_s^{LL} and f_s^{RR}, the longitudinal helicity amplitude was shown to contain suppressed new physics amplitudes, which was not found for the transverse amplitudes. The opposite scenario was found for operators of the form f_s^{LR}, where the longitudinal polarisation suffered no suppression in new physics amplitudes, but strong suppression of $\mathcal{O}(\Lambda_{QCD}/m_b)$ was seen for transverse polarisations. The results for the three cases are given in Table 1.2, where the heavy quark approximation has been used to provide the relation $A_\parallel^{SM} = -A_\perp^{SM}$. It

Amplitude	f_s^{LL} Ext.	f_s^{RR} Ext.	f_s^{LR} Ext.
A_0	A_0^{SM}	A_0^{SM}	$A_0^{SM} + A_0^{NP(LR)} - A_0^{NP(RL)}$
$A_\parallel$	$-A_\perp^{SM} - A_\perp^{NP(LL)}$	$-A_\perp^{SM} - A_\perp^{NP(RR)}$	$-A_\perp^{SM}$
$A_\perp$	$A_\perp^{SM} + A_\perp^{NP(LL)}$	$A_\perp^{SM} + A_\perp^{NP(RR)}$	$A_\perp^{SM}$

Table 1.2.: Extensions to amplitudes in the presence of different NP scenarios [20]. Note in the row of $A_\parallel$, terms of $\mathcal{O}(\lambda_{QCD}/m_b)$ have been neglected.

can be seen from Table 1.2, that different NP scenarios provide a different pattern to the observables measured in the $B_s^0 \to \phi\phi$ decay. This therefore shows that these observables are not only sensitive to NP, but also that discriminating power between different NP scenarios can be achieved.

1.5. Summary

In this Chapter, *CP* violation in the SM has been introduced. This includes not only the elegance of the CKM mechanism for *CP* violation in the SM, but also the vast effort, most notably from experimental results and lattice QCD, to determine the values and provide one of the most important tests of the SM.

The phenomenology of neutral meson mixing has been derived and the wealth of observables available to provide tests of the SM have been introduced. The latest tests of the SM in the context of global fits to measurements in the B^0 and B_s^0 sector have shown that results are largely consistent with SM expectations, while the discrepancy between the constraints provided by the *CP*-violating phase in B_s^0 mixing and the constraints provided by semileptonic asymmetries remains currently unexplained.

Penguin decays, specifically in $b \to s\bar{s}s$ transitions have been shown to exhibit small quantities of *CP* violation in the SM. Though still relatively unexplored experimentally, intriguing hints of larger *CP*-violating effects remain possible from the two results of studies on $b \to s\bar{s}s$ transitions thus far. The observables measured in $b \to s\bar{s}s$ transitions have been shown to be not only sensitive to NP, but also to have power in discriminating the type of NP, if present.

Chapter 2.

The LHCb Experiment

"Tu ne prévois les événements que lorsqu'ils sont déjà arrivés."
— Eugéne Ionesco

2.1. The Large Hadron Collider

Located at the European Organisation for Nuclear Research (CERN), the Large Hadron Collider (LHC) hosts four major particle physics experiments. The LHC, 27 km in circumference, is designed to collide protons at 14 TeV with instantaneous luminosities of up to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [33]. The four experiments are the ATLAS (A Toroidal LHC Apparatus) and the Compact Muon Solenoid (CMS) general purpose detectors, designed to search for the Higg's boson and physics beyond the Standard Model (SM) through the decays of undiscovered fundamental particles; the LHC Beauty (LHCb) experiment, designed to search for new physics mainly through CP violation in rare decays of B mesons and ALICE (A Large Ion Collider Experiment), designed mainly to investigate the quark gluon plasma through ions collided in the LHC ring.

To reach the high energies for protons in the LHC, the protons must be accelerated before injection. This process begins in the Linear Accelerator (LINAC2), where electrons are removed from hydrogen gas and the resulting protons accelerated to an energy of 50 MeV. The protons are then accelerated in a booster ring to 1.4 GeV. The Proton Synchrotron is responsible for bringing the protons up to an energy of 25 GeV (99.9 % of the speed of light). In the Super Proton Synchrotron, energies of 450 GeV are reached

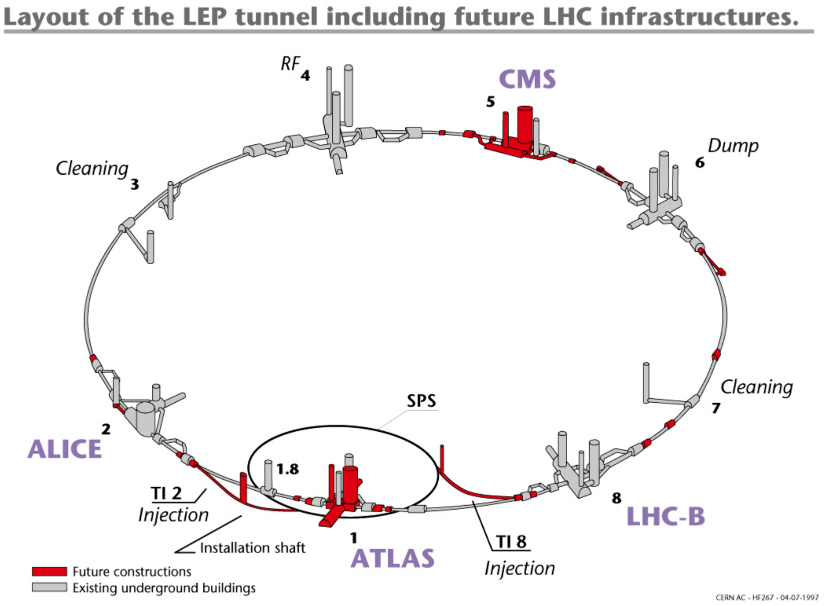


Figure 2.1.: Location of the four major experiments on the LHC ring in addition to the injection and dump points of the proton beams [34].

and protons are then injected in to the LHC at points 2 and 8 in the ring (shown in Figure 2.1). The LHC is designed to accelerate each proton beam to an energy of 7.0 TeV, although during 2011 and 2012 operation, energies of 3.5 TeV and 4.0 TeV were obtained, respectively.

2.2. The LHCb Detector

For proton-proton collisions in the TeV range, both b and $\bar{b}$ quarks are produced in the same forward or backward region at the LHC. This is due to the kinematics of the gluons and quarks inside the proton. For this reason, it is only necessary to build a detector in the forward region in order to detect b-$\bar{b}$ quark pairs produced in proton collisions. The LHCb detector (shown in Figure 2.2) is a forward arm spectrometer, designed to measure long-lived charged particles originating from the decays of B mesons with high precision, thus providing accurate momentum resolution necessary for measurements of decay angles, which in turn gives improved accuracy in measurements of CP violation. The rest of this Chapter is devoted to the description of the detector that allows LHCb to perform B physics in the challenging environment provided by a hadron collider. These focal points can be divided into the description of the tracking of charged particles, the identification of the nature of the particles through particle identification (PID) and

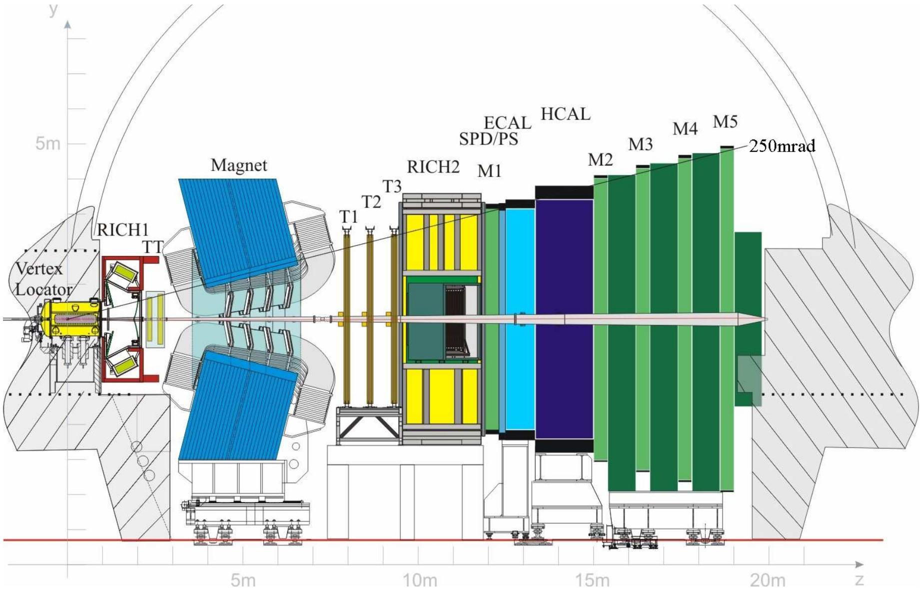


Figure 2.2.: The LHCb detector.

the trigger technology used to reduce the background to a level that can be written to storage media for offline analysis. This strategy allows the LHCb experiment to make precise measurements of the properties of B and D meson decays.

2.3. Tracking System

Tracking at LHCb is performed by the VErtex LOcator (VELO), designed to accurately determine the decay vertex of B and D mesons; the Silicon Tracker, comprised of the Tracking Turicensis (TT) and Inner Tracker (IT), which both use silicon microstrip sensors to detect the presence of charged tracks with a single hit spatial resolution of 50 µm and the Outer Tracker (OT), a drift-time detector designed to cover a larger acceptance area with drift coordinate resolution of 200 µm.

2.3.1. Magnet

The LHCb dipole magnet is used to aid in the determination of the momentum of charged particles through the force a charged particle experiences perpendicular to the motion in the presence of a magnetic field. This force follows the well-known left hand rule of John Fleming. The radius of curvature for the charged track is directly proportional to

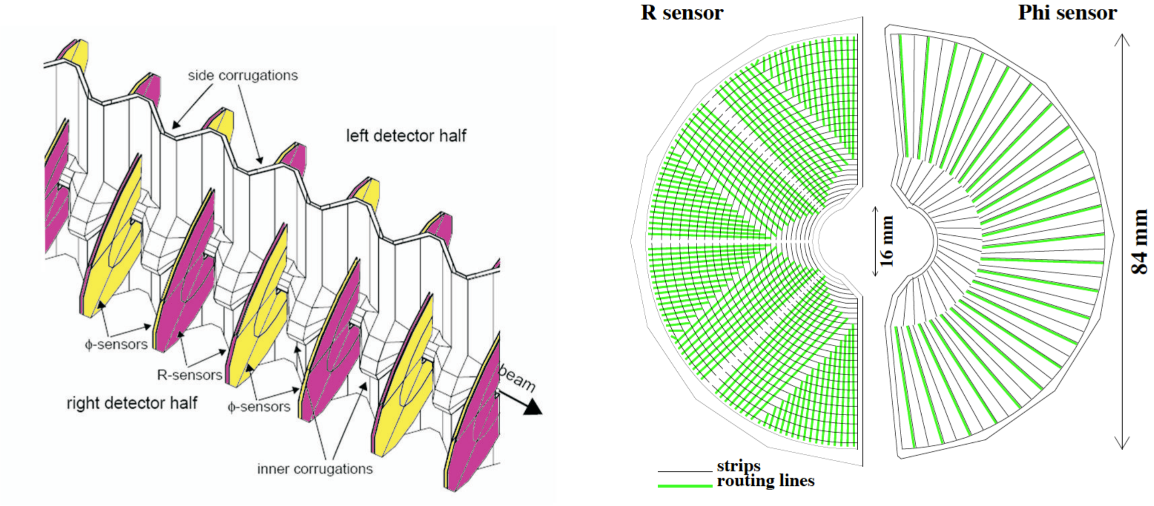


Figure 2.3.: Diagram of the VELO sensor locations with respect to the LHC beam (left) and schematic design of the sensors in each module (right).

the momentum of the particle. For the LHCb detector, the magnet consists of 54 tons of saddle-shaped coils wrapped around a 1500 ton window-frame yoke. The magnet is designed to provide an integrated magnetic field of 3.62 Tm, allowing charged particle tracks of momenta up to 200 GeV/c to be determined with 0.5 % uncertainty.

2.3.2. Vertex Location

The VELO consists of a series of modules positioned along the LHC beam axis (shown in Figure 2.3). Each module consists of an r-type and a ϕ-type silicon sensor fixed to a common support, designed to ascertain the position of a given track in cylindrical polar coordinates for the radius and azimuthal angle, respectively. Each sensor strip consists of oxegenated n^+-on-n silicon with an n-type implant on an n-type bulk, except for two sensors which are n^+-on-p for the purposes of the LHCb upgrade, attached to a backplane p^+ implant (shown in Figure 2.4). A nominal bias voltage of 150 V is applied to the sensor strips to ensure a depletion region is created. When a charged particle is incident on the depleted region, electron-hole pairs are produced, which create a current pulse with an amplitude dependent on the number of electron-hole pairs produced. These sensors allow for a spatial resolution of $\sim 4\,\mu\text{m}$ for tracks at an angle of 100 mrad from the beam axis in the smallest strip pitch region.

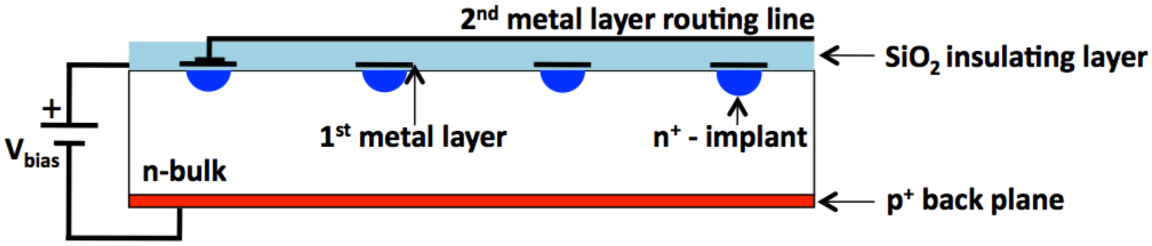


Figure 2.4.: Schematic diagram of an r-type VELO sensor.

The proton interaction point at LHCb is located inside the vertex locator, which operates as close as 8 mm from the beam axis during data-taking conditions. When the LHC is setting-up collisions at the four interaction points on the ring, it is not safe for devices to be in such close proximity to the beam axis as the proton beams are wider and more unstable than for data-taking. The VELO is therefore retracted to a safe distance of at least 40 mm. Due to the proximity to the LHC beam, the VELO was designed with radiation hardness in mind. During one nominal year of LHCb data-taking, i.e. $2\,\text{fb}^{-1}$ data collected, the silicon in the inner region of the VELO modules is exposed to an equivalent of 1 MeV neutrons at a flux of $1.3 \times 10^{14}\,\text{n}_{\text{eq}}/\text{cm}^2$. Radiation damage introduces impurities that affect the doping of the semiconductor material. This causes the depletion voltage to change over time, as described by the Hamburg model [35]. The dominant mechanism is expected to be the inactivation of phosphorous dopants in addition to the introduction of acceptors. This causes the depletion voltage to decrease for an n^+-on-n semiconductor and then increase when the type changes to p-on-n. This is shown in Figure 2.5 for the sensors in the VELO [36], where the effective depletion voltage[1] (EDV) is used due to the inability to determine the depletion voltage after installation. The current pulse when a charged particle is incident on a sensor is measured with an analogue to digital converter (ADC). The EDV is defined as voltage at which the most probable value of the ADC distribution achieves 80 % of the maximum. It should be noted that the EDV does not correspond to the depletion voltage close to type-inversion. In Figure 2.5, it can be seen that far from type inversion, agreement with the Hamburg model is seen.

The performance of the VELO has been evaluated on 2011 data [37] in terms of the primary vertex (PV) location and impact parameter (IP), where the IP is defined as the closest radial distance of a track with respect to the primary vertex. The PV resolution is shown as a function of track multiplicity in Figure 2.6. This is measured by randomly

---

[1] This is equivalent to the depletion voltage in most circumstances.

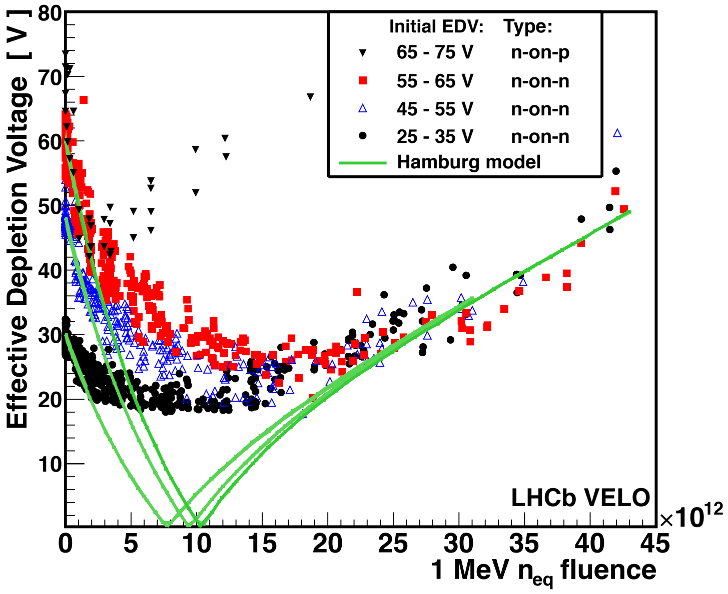


Figure 2.5.: EDV as a function of fluence for VELO sensors [36]. The data is compared to predictions from the Hamburg model [35], where good agreement is seen for fluences far from type inversion.

splitting the track sample in two and reconstructing two independent PVs. The resolution is then taken as the width of a Gaussian fit to the distribution of the distances between the two PVs divided by $\sqrt{2}$. The resolution for a single PV with a track multiplicity of 25 is measured to be $(13.1, 12.5, 69.2)$ µm for the (x, y, z) components [37]. The IP resolution is shown as a function of $1/p_{\mathrm{T}}$ in Figure 2.6 for 2011 data and simulation, and is measured from the widths of the IP distributions projected on to the x and y axes. Measurements on 2011 data have yielded an IP resolution of $(13.2 + 24.7/p_{\mathrm{T}})$ µm [37]. Simulated events are known to underestimate the resolution due to the description of the vertex locator.

2.3.3. Tracker Turicensis and Inner Tracker

The TT and the IT have been designed as a joint project known as the Silicon Tracker (ST), hence the two subdetectors share much of the technology and design features. The TT is a planar tracking station covering an area of 1.5 m × 1.3 m, thus covering the whole acceptance of the LHCb detector upstream of the LHCb dipole magnet. The IT covers a 1.2 m × 0.4 m cross-shaped region in the innermost area of the three tracking stations downstream of the LHCb dipole magnet.

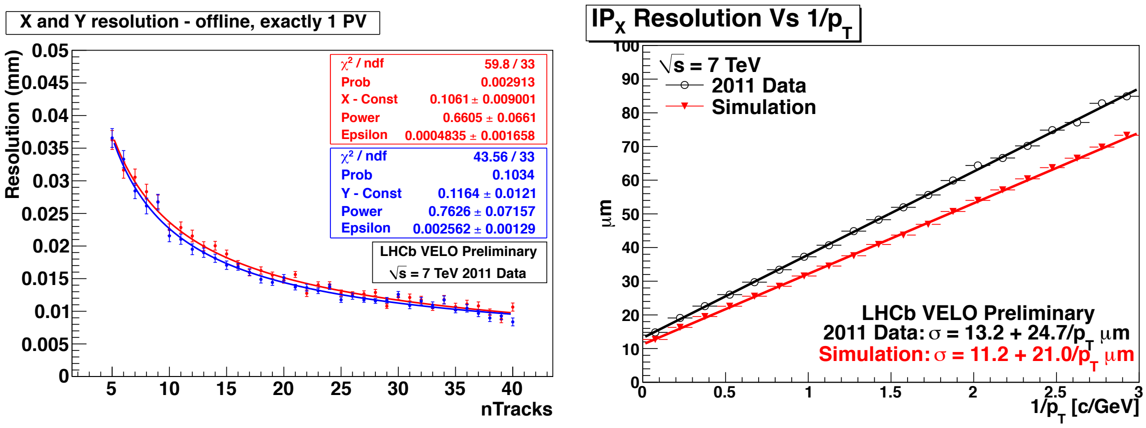


Figure 2.6.: Performance of the VELO in terms of PV resolution (left) and IP resolution (right) [37]. The PV resolutions are shown for the x and y directions in red and blue, respectively, for offline-reconstructed events consisting of one PV. IP resolutions are shown as a function of $1/p_{\mathrm{T}}$ for 2011 data and simulation in black and red, respectively.

Detectors forming the ST are composed of two half-stations housing silicon sensors in four layers, arranged in the formation x, u, v, x, where x indicates silicon strips positioned vertically, u indicates silicon strips rotated by an angle of $-5°$ and v indicates silicon strips rotated at an angle of $5°$. The TT detectors are 500 µm thick single-sided p^+-on-n silicon sensors divided into 512 readout strips. The IT detectors are p^+-on-n silicon sensors divided into 384 readout strips. Thickness is chosen to achieve the best compromise between the signal-to-noise ratio and material budget. Bias voltages for the silicon sensors may be applied to sectors individually to account for the different ageing rates based on proximity to the LHC beamline. Each layer of the TT is made from half-modules consisting of seven silicon sensors. Each layer of the IT is made from modules containing one or two silicon sensors depending on location. The thickness of the IT detectors is 320 µm for one-sensor modules and 410 µm for two-sensor modules. The layout of the detection layers of the TT and IT is shown in Figure 2.7. The need to split silicon sensors into sectors originates from the density of tracks, as the sensors closest to the beamline will have a higher occupancy, which will place a larger strain on the readout system than sensors far from the beamline. For the case of half-modules of the TT close to the beam pipe, the silicon sensors are organised into a formation of four sensors in the L sector, two in the M sector and one in the K sector, with each sector connected to a readout front-end. For the case of half-modules of the TT away from the beam pipe, silicon sensors are organised into an L sector consisting of four sensors and an M sector

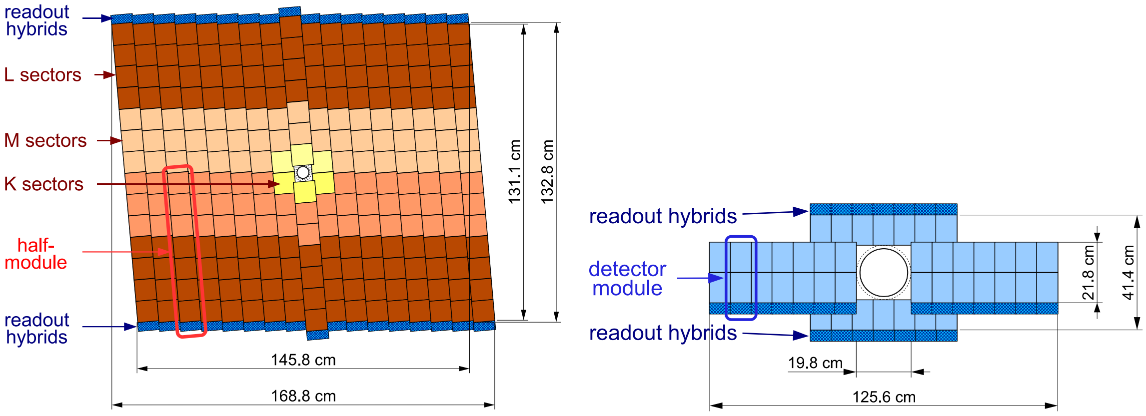


Figure 2.7.: Diagram of the sensor locations in the third TT layer (left) and first IT layer (right), where different shades of orange represent the different silicon sensor types in the TT layer.

containing the three remaining sensors with the L sector connected to a front-end and the M sector connected to the remaining two.

Adjacent modules within a detection layer are staggered by 1 cm in z and overlap by a few millimeters in x to avoid acceptance gaps and allow for the alignment of adjacent layers. For the case of the u and v layers, each module is individually rotated by the stereo angle. The use of detection layers in a (x, u, v, x) formation allows for multiple tracks to be identified simultaneously with no ambiguities. The measured hit resolution from the TT and IT is measured in data to be 59 µm and 50 µm, respectively, compared with the design resolution of 50 µm [37].

2.3.4. Outer Tracker

The design of the Outer Tracker is based on arrays of gaseous straw tubes, where measuring the drift time allows trajectories of charged particles to be determined. Each straw tube is comprised of a cathode on the outer edge and an anode wire at the centre of the tube, with the volume filled with gas. A charged particle traversing will ionize the gas releasing drift electrons, which will in turn be accelerated to the anode wires in a time typically less than 10 ns. The resulting signal on the anode is then amplified, with the drift time of the electrons giving a measure of the radial distance of the charged particle from the anode. The OT subdetector consists of three detectors, each positioned outside a corresponding inner tracker on the three stations downstream of the LHCb dipole magnet (T1, T2 and T3 in Figure 2.2). Each of the stations consists of four

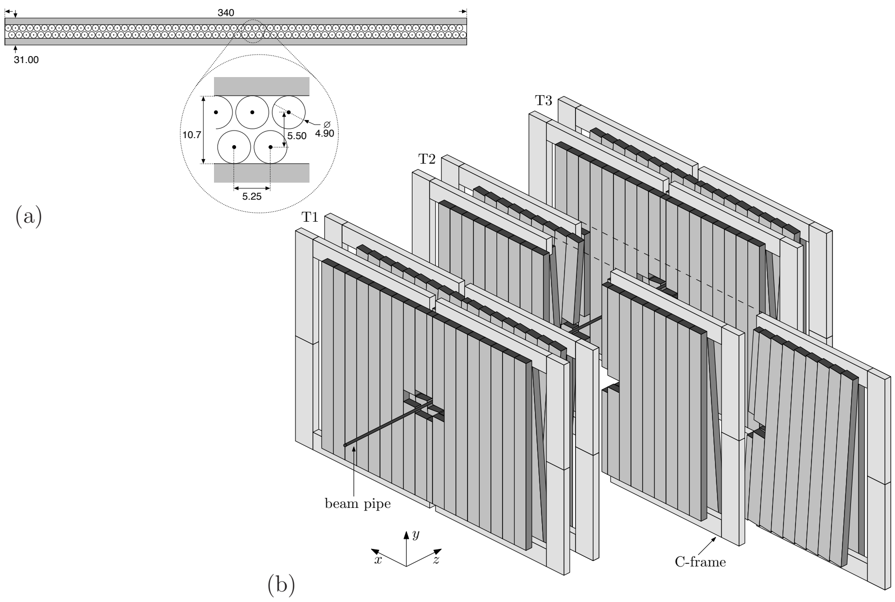
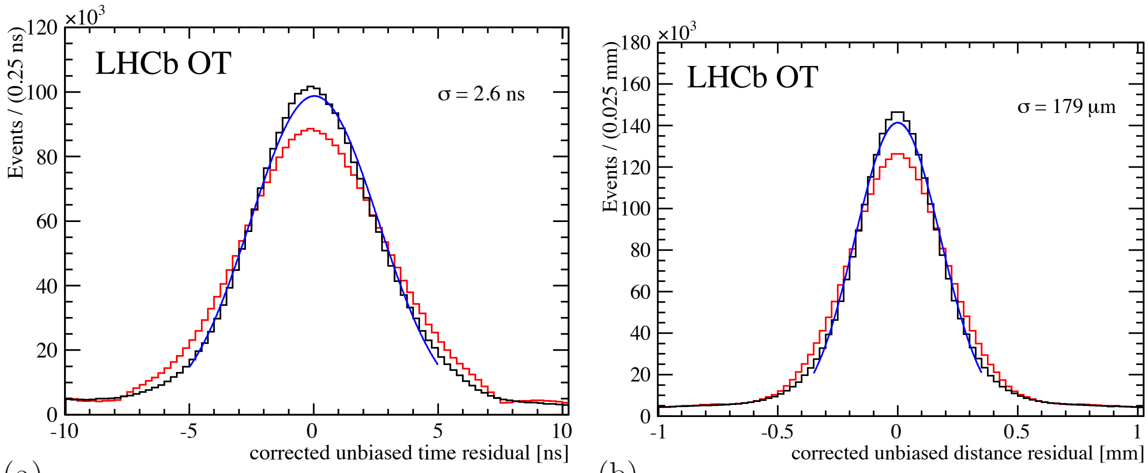


Figure 2.8.: (a) Cross section of an OT module and (b) diagram of the OT arrangement in the tracking stations [38].

layers arranged in a x-u-v-x formation as for the silicon trackers, where the u and v configurations represent stereo angles of $+5°$ and $-5°$, respectively. A diagram of the station arrangement is shown in Figure 2.8. The outer boundary of the OT detectors is designed to coincide with an acceptance of 300 mrad in the horizontal plane and 250 mrad in the vertical plane. The inner boundary is determined by the requirement that the occupancies should not exceed 10 % at a nominal luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

Each module contains two monolayers of drift tubes of diameter 4.9 mm. A cross-section of a layer of the module is shown in Figure 2.8. Two types of modules are present in the OT. These are the F-type modules containing two hundred and fifty six straws and achieving an active length of 4850 mm. The second type of module is the shorter S-type, containing half the number of straw tubes and covering approximately half of the active length of the F-type. Each layer of the OT is built from seven F-type modules and four S-type modules. A mixture of 70 % argon and 30 % carbon dioxide is used as a counting gas in order to achieve drift-time resolution of ~ 2.6 ns and a good drift resolution of $\sim 179\,\mu\mathrm{m}$ [38], where the resolutions have been evaluated from data. The data samples used required the charged particles to have momentum greater than 10 GeV/c and a track fit χ^2 per degree of freedom less than 2 without the hit in question being used in the track fit. The drift time and position residuals are shown in Figure 2.9.



(a) (b)

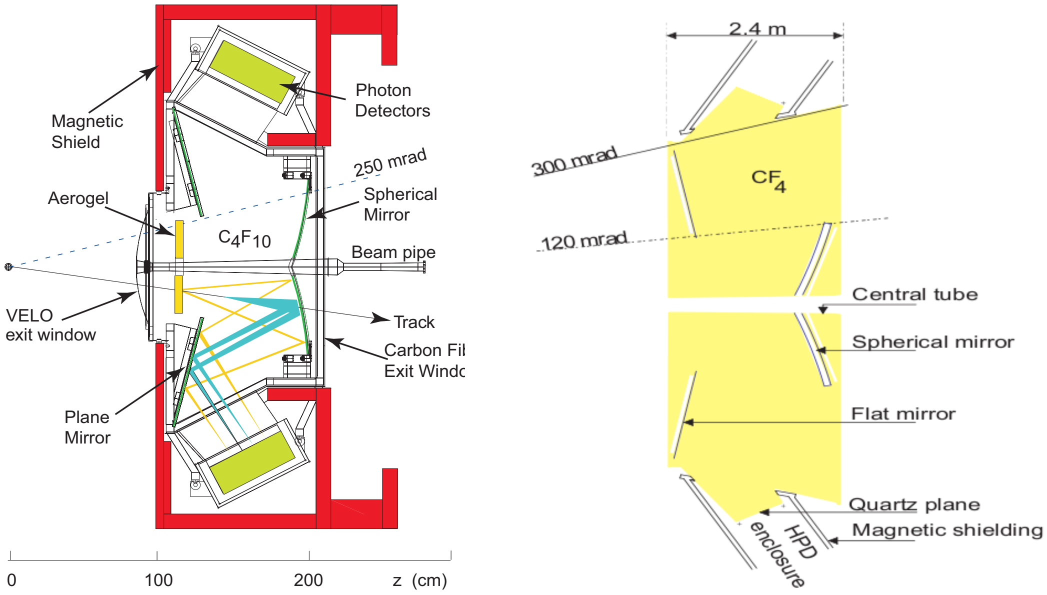
Figure 2.9.: (a) Drift time residual distribution and (b) hit distance residual distribution. The core of the distributions is fitted with a Gaussian function, which is used to measure the resolutions [38].

2.4. Particle Identification

At LHCb, particle identification is provided by Ring-Imaging CHerenkov detectors (RICH) [39], which allow for species of charged particles to be distinguished in the momentum range 2-100 GeV/c; an electromagnetic calorimeter (ECAL), which determines the energy of electrons and photons; a hadronic calorimeter (HCAL), which determines the energy of hadrons, and muon chambers, which determine the energy of muons in the forward region.

2.4.1. RICH Detectors

Cherenkov radiation is a well known phenomenon that arises in relativistic electromagnetism whereby no charged particle may travel faster than the speed of light in a given medium continually, and no particle may travel faster than light in a vacuum. A consequence of special relativity is that action is not instantaneous. When a charged particle enters a medium faster than light may traverse the medium, Cherenkov radiation is produced in a similar way to the sonic boom produced when an object travels faster than the speed of sound. This radiation is emitted by the particle at an angle, θ, from the direction of motion that contains a dependence on the velocity of the particle. This



(a) (b)

Figure 2.10.: Diagram of the layouts of the RICH1 (a) and RICH2 (b) detectors.

relation is

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

where n is the index of refraction of the material and $\beta = v/c$, with v representing the velocity of the particle and c the speed of light in a vacuum.

The layout of the two RICH detectors is shown in Figure 2.10. When Cherenkov radiation is emitted from a radiator, aerogel or C_4F_{10} gas in RICH 1 and CF_4 gas in RICH 2, the photons are reflected off a spherical mirror, then a flat mirror and focused onto the detection planes. The detection planes in RICH 1 are divided into two halves located above and below the beamline, while the detection planes in RICH 2 are divided into two halves at either side of the beamline.

The Cherenkov photons are detected with Hybrid Photon Detectors (HPDs). A schematic diagram of the HPD design is given in Figure 2.11. When a photon is incident on the window of an HPD, a photo-electron is produced at the photo-cathode. This is then accelerated towards a reverse-biased silicon sensor by a voltage of 18 kV. When the photo-electron impacts on the silicon sensor, electron-hole pairs are created in the silicon for every 3.6 eV on average. Two planes consisting of 98 HPDs each in RICH 1

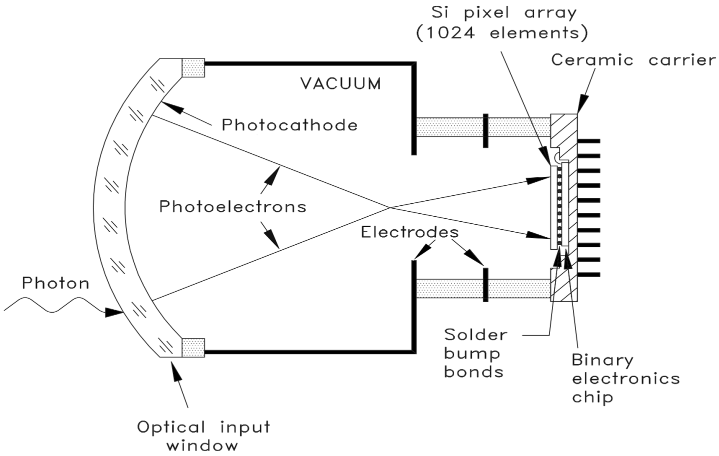


Figure 2.11.: Schematic diagram of the Hybrid Photon Detector design used in RICH detectors.

and 144 HPDs each in RICH 2 are used to identify rings of Cherenkov photons in order to provide a measurement of the velocity of the charged particle, which in combination with momentum information from the trackers, provides a mass measurement.

Angular resolutions of $1.618 \pm 0.002\,\mathrm{mrad}$ and $0.68 \pm 0.02\,\mathrm{mrad}$ are achieved for the C_4F_{10} and CF_4 radiators, respectively. The angular resolution of the aerogel radiator is approximately 5.6 mrad. The reason for the 2-100 GeV momentum range can be seen from Figure 2.12 as it is only for these ranges, that charged particles produce rings of Cherenkov photons that can be measured with an accuracy that distinguishes between different species of particles.

2.4.2. Calorimetry

LHCb adopts the traditional approach to calorimetry, in which the Hadronic CALorimeter (HCAL) is placed behind an ELectromagnetic CALorimeter (ECAL). The principle of the sampling calorimeters used in LHCb is based on the use of a scintillating medium, which releases photons with yields given through Birks' law [40]

$$\frac{\mathrm{d}L}{\mathrm{d}x} = L_0 \frac{\mathrm{d}E/\mathrm{d}x}{1 + K_B \mathrm{d}E/\mathrm{d}x}, \tag{2.2}$$

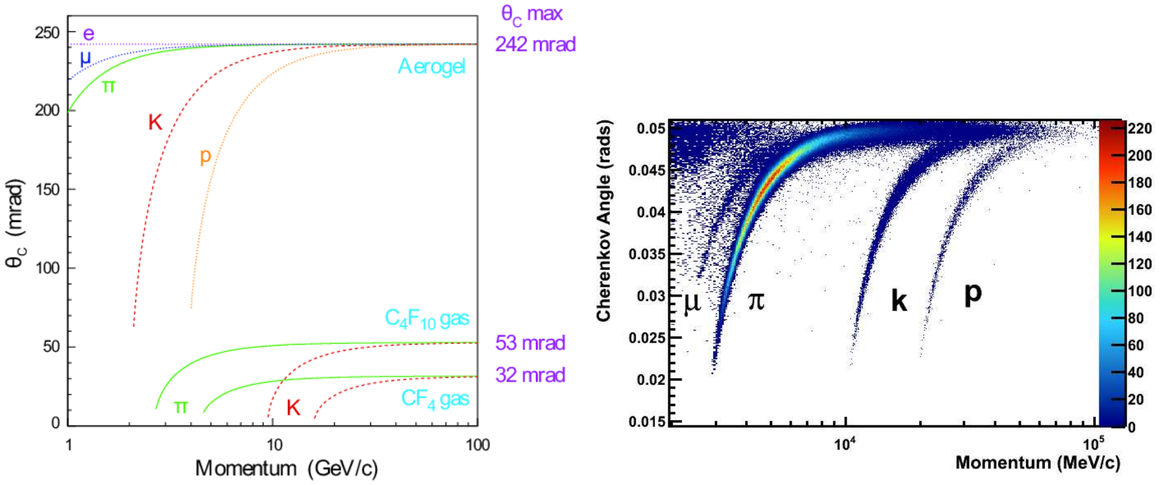


Figure 2.12.: Theoretical distribution of the expected Cherenkov angle versus momentum for different species of particles for the three radiators (left) and corresponding data distribution for the C_4F_{10} radiator (right) [39].

where L_0 is the scintillation efficiency, K_B is a constant depending on the scintillation medium, $\mathrm{d}L/\mathrm{d}x$ represents the light yield in a given path length and $\mathrm{d}E/\mathrm{d}x$ is the energy loss of the particle for a given path length.

The ECAL has been designed to be able to trigger photons and electrons, above a transverse energy (E_T) threshold, and ECAL measurements are used extensively in the reconstruction of neutral particles. The ECAL cells are comprised of alternating layers of 2 mm thick lead and 4 mm thick scintillating tiles wrapped in 120 μm thick, white reflecting TYVEK paper. The layers have a longitudinal length of 42 cm, corresponding to $25X_0$, where X_0 is the radiation length. Scintillator tiles are made from polystyrene. Optical fibres are used to collect light from the scintillating layers and traverse the entire cell to phototubes at the end of the cell. The cells of the ECAL may be considered in three regions: The inner region containing 3072 cells, the middle region containing 3584 cells and the outer region containing 5376 cells with dimensions 4×4, 6×6 and $12 \times 12\,\mathrm{cm}^2$, respectively. The cell structure of the calorimeters is shown in Figure 2.13.

A Preshower Detector (PS) is present that is longitudinally separated from the rest of the ECAL with a Scintillating Pad Detector (SPD) placed in front of the PS. The PS and the SPD are introduced to reject the large backgrounds of charged and neutral pions, respectively. The SPD/PS detector consists of 15 mm of a lead converter (two radiation lengths thick) that is sandwiched between two layers of scintillating pads with 12032 detection channels. The sensitive area of the detector is 6.2 m high and 7.6 m wide. Each layer of the SPD/PS is divided in two to allow for access and maintenance. There

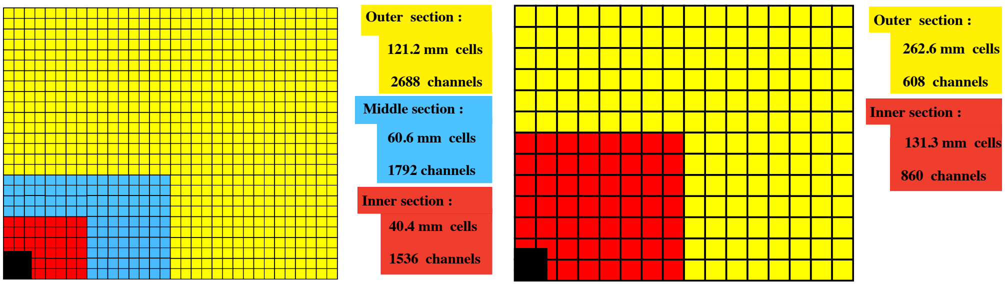


Figure 2.13.: Diagram of the cell structures in the Electronic Calorimeter (left) and Hadronic Calorimeter (right).

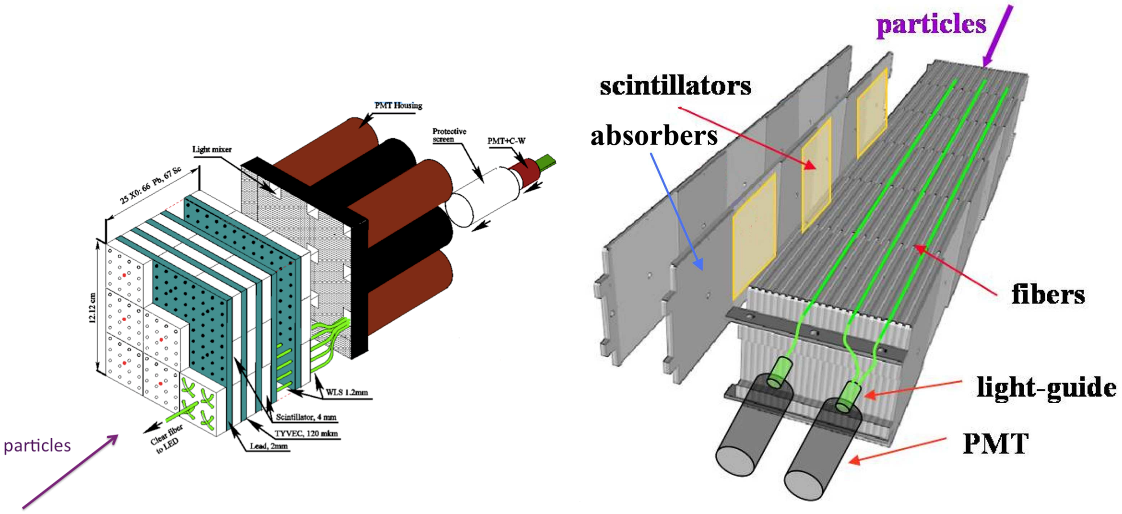


Figure 2.14.: Diagram of the internal elements of an ECAL module (left) and HCAL module (right) [42, 43].

is a one-to-one projective correspondence between the SPD, PS and the ECAL cells. A diagram showing the structure of an ECAL module is shown in Figure 2.14. The ECAL provides an energy resolution of $(0.8/\sqrt{E} \oplus 0.9)\,\%$, where E is the energy in GeV [41].

The hadronic calorimeter is designed to trigger on hadrons with large E_{T} and is separated into cells in a similar fashion as for the ECAL. The hadronic calorimeter cells are divided into two regions instead of the three for the ECAL. Cells in the inner region of the HCAL are $13.1 \times 13.1\,\mathrm{cm}^2$, whereas cells in the outer region of the HCAL are $26.3 \times 26.3\,\mathrm{cm}^2$. A diagram showing the structure of a HCAL cell is shown in Figure 2.14, where two layers are separated from the stack for clarity. The cell consists of alternating layers of iron and scintillator layers with the thickness of the scintillators corresponding to 5.6 interaction lengths. The iron layers are 1 cm thick and have a longitudinal length

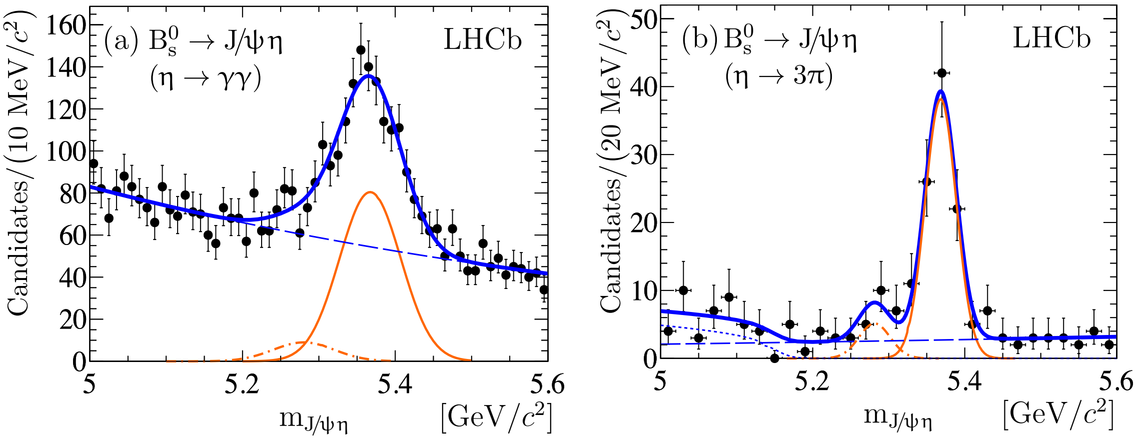


Figure 2.15.: Invariant mass distributions for (a) $B_s^0 \to J/\psi\eta(\to \gamma\gamma)$ and (b) $B_s^0 \to J/\psi\eta(\to \pi^+\pi^-\pi^0)$ decays, obtained from 1.0 fb^{-1} of 2011 LHCb data [44]. Solid orange lines show the signal B_s^0 contribution, dashed orange lines show the B^0 contribution, dashed blue lines show the combinatoric background, and dotted blue lines show the partially reconstructed background.

equal to the hadronic interaction length in steel. As in the case of the ECAL, scintillation photons are collected by optical fibres traversing the length of the cell. Photons are detected at the end of the cell by photo-multiplier tubes. The HCAL provides an energy resolution of $(69/\sqrt{E} \oplus 9)\,\%$, where E is the energy in GeV [41].

The performance of the ECAL is shown in Figure 2.15 for $B_s^0 \to J/\psi\eta(\to \gamma\gamma)$ and $B_s^0 \to J/\psi\eta(\to \pi^+\pi^-\pi^0)$ decays, obtained from 1.0 fb^{-1} of 2011 LHCb data [44], where information has been used from the ECAL, to fully reconstruct B_s^0 candidates. Resolutions of the B_s^0 signals are found to be $40.1 \pm 3.6\,\mathrm{MeV}/c^2$ and $20.3 \pm 2.3\,\mathrm{MeV}/c^2$ for the case of $B_s^0 \to J/\psi\eta(\to \gamma\gamma)$ and $B_s^0 \to J/\psi\eta(\to \pi^+\pi^-\pi^0)$ decays, respectively.

2.4.3. Muon Chambers

Muon chambers at LHCb are located furthest from the interaction point, with the exception of the M1 muon chamber in Figure 2.2, which is positioned in front of the calorimeters. This is due to the fact that the vast majority of muons traverse the entire length of the detector. The M2-M5 chambers are interleaved with 80 cm thick iron absorbers to filter muons from hadrons that may have passed through the calorimeters (the iron absorbers are shown as dark green in Figure 2.2). The technology used at LHCb is that of multi-wire proportional chambers (MWPC) and triple gas electron multiplier

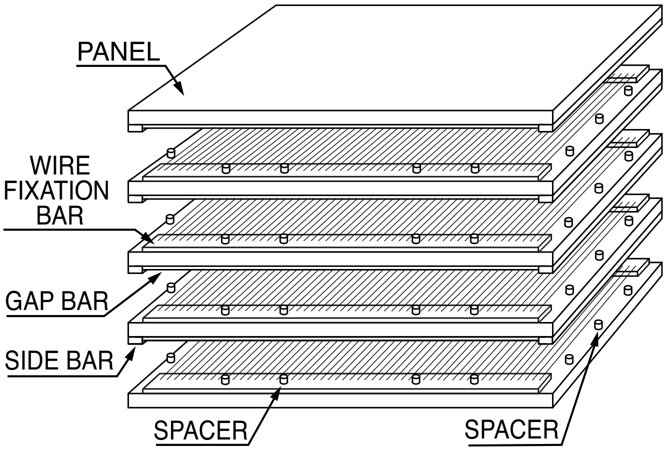


Figure 2.16.: Diagram of a MWPC module.

(GEM) detectors in the region closest to the beamline in the M1 chamber, that provide tracking information and muon identification information.

2.4.3.1. MWPC

The principle of Multi-Wire Proportional Chambers is based on the ionisation of gas in the chamber by a passing muon. The ionisation electrons are then collected and amplified on the wires in the chamber. The charge collected by the wires is used to detect the presence of a muon. An expanded view of a muon chamber is shown in Figure 2.16. The multi-wire proportional chambers in LHCb operate with ionisation of a $Ar/CO_2/CF_4$ gas mixture, with corresponding ratios 8:11:1. This gas ratio is chosen to satisfy the requirements that each station must be greater than 99 % efficient within a time window of 25 ns in order that the trigger efficiency is greater than 95 % [45]. The chambers consist of alternating layers of 9 mm thick insulating cores and conducting planes in which 30 µm gold-plated tungsten wires are positioned 2 mm apart in order to achieve time resolution of 5 ns. Four wire layers are present in each of M2-M5, two in the case of M1.

2.4.3.2. GEM Detectors

The principle of operation of gas electron multiplier detectors relies on two electrodes separated over a short distance with many small holes inserted. With such a design, the electric field may be produced in a manner that is much more resistant to radiation damage

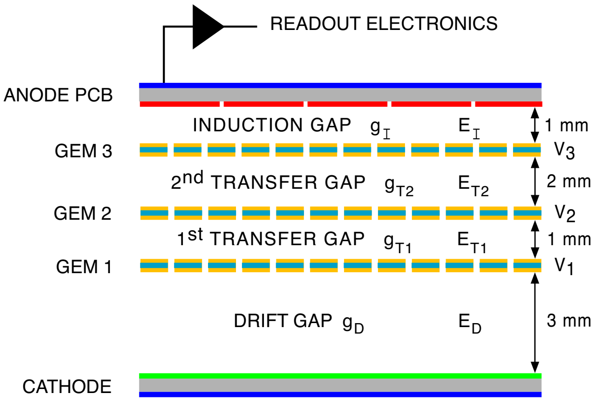


Figure 2.17.: Principle of operation of a triple-GEM module.

than in a MWPC. This advantage allows for the high efficiency and time requirements to be satisfied in the high flux environment of the innermost region of M1. The triple-GEM detection system present in LHCb uses three layers of foils as shown in Figure 2.17. The optimum values of the electric fields between foils and the potential differences placed across them have been determined from the trade-off between discharge probability and time resolution. These typically take the values of $E_D = 3.5\,\mathrm{kV/cm}$, $E_T = 3.5\,\mathrm{kV/cm}$, $E_I = 5\,\mathrm{kV/cm}$, $V_1 = 440\,\mathrm{V}$, $V_2 = 430\,\mathrm{V}$ and $V_3 = 410\,\mathrm{V}$, where E_D and E_I are the drift gap and induction gap electric fields, respectively, and V_1, V_2 and V_3 are the voltages applied on the three GEMs, as shown in Figure 2.17. The triple-GEM detector achieves a time resolution of around 3 ns, using holes with a diameter of 50 µm and a pitch of 140 µm along with an RMS gain variation of $\sim 10\,\%$. This performance is achieved with an optimised detector geometry based on the considerations that mechanical constraints mean a minimum separation of 1 mm between foils is required; the drift gap (g_D) is made large enough to allow all charged tracks to be detected; the first transfer gap (g_{T1}) is small enough that primary electrons produced from the ionisation of the gas produce as few further ionisations as possible; the second transfer gap (g_{T2}) is large enough to allow the ionisation electrons to spread over multiple holes and hence reduce the discharge probability.

2.4.3.3. Muon Identification Performance

A decision on whether a candidate was a muon or not can be taken based on the hits in the muon detectors corresponding to extrapolations from the tracking stations. For a candidate with momentum, p, in the range $3\,\mathrm{GeV}/c < p < 6\,\mathrm{GeV}/c$, hits must be found

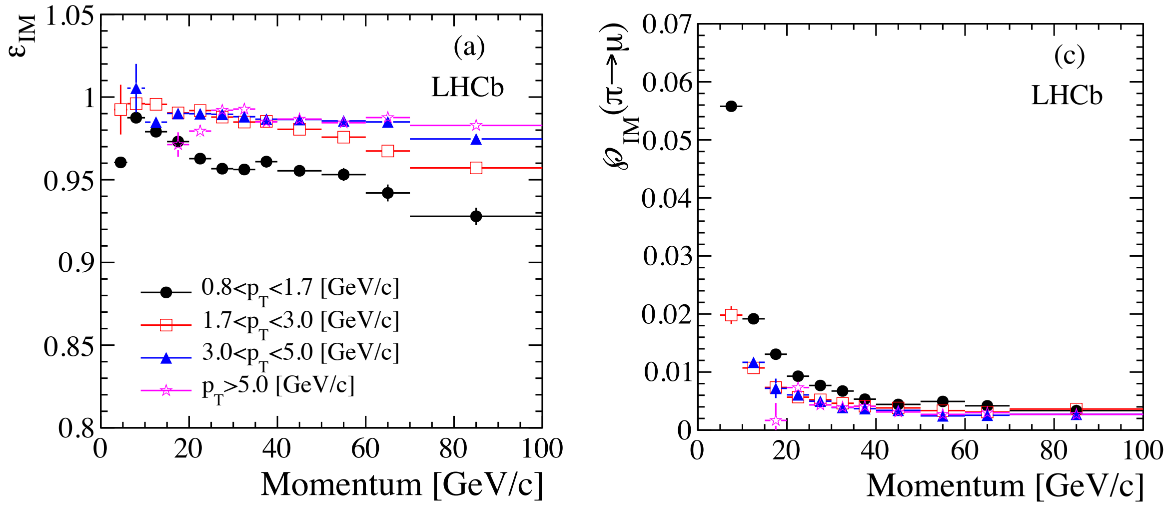


Figure 2.18.: Performance of the muon identification decision in terms of efficiency (left) and pion mis-identification probability (right) [46] from 2011 LHCb data. Efficiencies have been evaluated from 2.4 million $J/\psi \rightarrow \mu^+\mu^-$ candidates and pion mis-identification probabilities have been evaluated from 11.7 million $D^{*+} \rightarrow D^0(\rightarrow \pi^+K^-)\pi^+$ candidates.

in M2 and M3, for a candidate with momentum $6\,\mathrm{GeV}/c < p < 10\,\mathrm{GeV}/c$, hits must be found in M2, M3 and either M4 or M5, and for a candidate with momentum greater than $10\,\mathrm{GeV}/c$, hits must be found in all muon stations to be classed as a muon. The performance of this muon identification decision is shown in Figure 2.18.

2.5. Trigger System

The design luminosity of LHCb is $2\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. At this luminosity, visible pp interactions are expected to produce 100 kHz of b-$\bar{b}$ quark pairs. For the purposes of studying *CP* violation, the majority of branching fractions for interesting decays are expected to be $< 10^{-3}$. This therefore requires a large amount of background rejection. At LHCb, this background rejection is performed by passing events through a hardware trigger (L0) to reduce the LHC beam crossing rate from 40 MHz to 1.1 MHz, which is a rate that allows information from all subdetectors to be used. Events passing the L0 trigger are then required to pass through a first layer of software triggers (HLT1), reducing the rate further to 11 kHz. A final layer of software triggers performs the task of reducing the output rate to 2 kHz.

2.5.1. Hardware Trigger

The L0 decision unit, responsible for making the decision at the hardware stage for a given event receives information from two main sources in order to distinguish interesting B and D decays from the background. These are the calorimeter trigger, which uses data from the ECAL, HCAL, PS and SPD, and the muon trigger, which uses data from the 5 muon stations.

2.5.1.1. Calorimeter Trigger

The calorimeter component of the L0 trigger uses the ECAL and HCAL information in terms of 2×2 cells, with 32 cells allocated to each front-end. The decays of B mesons tend to produce final state particles with large transverse momentum (p_{T}) and energy (E_{T}), where the p_{T} is defined as the momentum projected onto the x, y plane in the LHCb co-ordinate system defined in Figure 2.2. The E_{T} is the energy measured, multiplied by the sine of the angle between the z axis and a straight line from the cell to the middle of the interaction region. Particles are selected based on the E_{T} variable. Three types of candidates are reconstructed by the calorimeter. These are hadron candidates, photon candidates and electron candidates. Hadron candidates are defined from the highest E_{T} HCAL cluster, and the sum of the E_{T} in the ECAL and HCAL if an ECAL cluster lies directly in front of the corresponding HCAL cluster. The hadron candidate is selected if the measured E_{T} is greater than $3.5\,\mathrm{GeV}/c$. Photon candidates are defined from the highest ECAL cluster, with up to 2 cells being hit in the PS directly in front of the corresponding ECAL cell and no hits in the corresponding SPD cells. Up to 4 cells in the PS are allowed for the inner regions of the ECAL. The E_{T} of the photon candidate is measured from the ECAL only. Electron candidates are defined in the same way as photon candidates, but require at least one hit in the corresponding SPD cells. The photon and electron candidates are selected if the measured E_{T} is greater than $2.5\,\mathrm{GeV}/c$. The efficiency of the L0 hadron trigger as a function of p_{T} is shown for 4 hadronic decay modes in Figure 2.19, where efficiencies have been determined from data [47]. In the rest of this dissertation, trigger performances said to be evaluated from the data are measured using the TISTOS method [47]. Candidates are selected as TOS (Triggered On Signal), if the particles measured in the final state of a decay would have passed the trigger requirements with no other tracks present. Candidates are selected as TIS (Triggers Independent of Signal), if the trigger requirements would have been met independent of the presence of the signal decay. The efficiency of the trigger in question is then defined

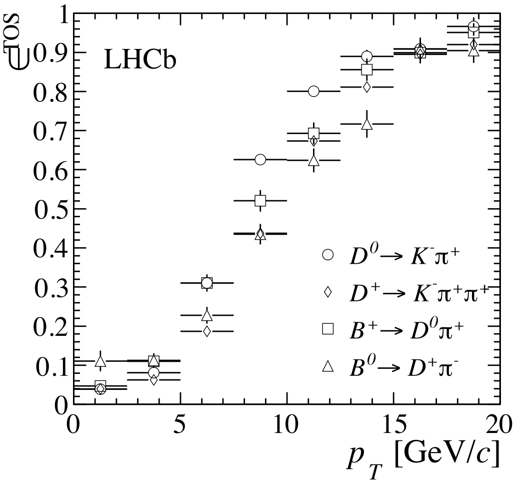


Figure 2.19.: Efficiency of the L0 hadron trigger as a function of p_{T} for $D^0 \to \mathrm{K}^+\pi^-$, $D^+ \to \mathrm{K}^-\pi^+\pi^-$, $B^+ \to D^0\pi^+$, and $B^0 \to D^+\pi^-$ candidates [47].

as

$$\varepsilon^{\mathrm{TOS}} = \frac{N^{\mathrm{TIS\&TOS}}}{N^{\mathrm{TIS}}}, \tag{2.3}$$

where N^{TIS} is the number of events triggered as TIS, and $N^{\mathrm{TIS\&TOS}}$ is the number of events triggered as both TOS and TIS. Equation 2.3 then allows for the trigger efficiency to be determined from data events.

2.5.1.2. Muon Trigger

Muons are present in the final states of many B and D decays, that are interesting not only for studies of CP violation and rare decays, but also for use in the identification of the initial flavour of neutral mesons. Muon tracks are identified using hits in M3 as a starting point. From a straight line connecting the interaction point and the hit location in M3, positions in M2, M4 and M5 are extrapolated. A search is then performed in fields of interest (FoI) centred on the extrapolated points. The size of the FoI depends on the minimum-bias retention, the background and the location in the muon subdetector. A muon is identified if for a given seed in M3, corresponding hits are found in all other muon detectors. The position in M1 is made through the straight line extrapolation from the points in M2 and M3[2]. The locations of muon tracks in the first two muon stations

---

[2] Such extrapolations are possible due to the projective design of the muon stations, i.e. there is a one-to-one mapping between pads in M2, M3, M4 and M5 and also a one-to-one mapping between pads in M1 and pairs of pads in M2 and M3.

allows for a p_{T} resolution of 25 % with respect to the momentum of the muon. Muon candidates are selected if the measured tracks have a p_{T} that is greater than 1.48 GeV/c.

2.5.2. Software Trigger

The High Level Trigger (HLT) is designed as a C++ software trigger that is executed on a dedicated Event Filter Farm (EFF). The HLT is executed in two stages referred to as HLT1 and HLT2. The purpose of HLT1 is to perform L0 confirmation, in which the tracking stations are used to confirm the existence of charged tracks that correspond to the calorimeter deposits or confirm that no such tracks exist in the case of photons and neutral pions. The purpose of HLT2 is to use inclusive and exclusive trigger algorithms, in which B or D decays are partially or fully reconstructed.

2.5.2.1. HLT1

The HLT1 trigger stage is based on a single track trigger, which is designed to look for a single, high transverse momentum track of good quality and displaced from the primary vertex (PV), due to the relatively long lifetime of B and D mesons. PVs are reconstructed by requiring vertices with at least 5 VELO tracks originating from them. Vertices are considered to be PVs if measured to be within 300 μm of the mean PV position in the x, y plane. Tracks are found by HLT1 through the application of selections in three stages. The first stage selects tracks in the VELO most likely to come from a B decay. This is performed through cuts on the distance between the track and the PV perpendicular to the z-axis, known as the impact parameter (IP) with respect to the PV. Selections based on the number of hits deemed to have originated from the VELO track and the difference between this number of hits and the expected number given the distance travelled in the VELO[3] are also used. The second stage uses an algorithm [48] to match the VELO tracks to the tracking stations. This algorithm selects tracks with momentum above 8 GeV/c and p_{T} greater than 800 MeV/c in order to reduce the size of the search window in the tracking stations. In the final stage the tracks are fitted with a Kalman-based method. This allows for a more accurate cut in the track χ^2 and provides a covariance matrix allowing a cut on the IP χ^2 to be performed.

---

[3] This preselection is necessary as time constraints prevent all VELO tracks from being matched to candidates in the tracking stations.

A final note is worth mentioning for final states involving muons or photons, where the HLT1 strategy previously outlined can be extended. In the case of muons, i.e. events passing the muon triggers in the L0 stage, the cuts used in the baseline trigger may be loosened, in order to improve efficiency. In the case of photons, meaning events passing the photon trigger in the L0 stage, offline selections use a tight cut on the energy at 2.4 GeV. This can be applied at the HLT1 stage again allowing for baseline selections to be loosened.

2.5.2.2. HLT2

The HLT2 level uses a combination of cut-based and multi-variate methods in order to select signal candidates. Due to the large number of parallel selection criteria used in the HLT2 stage to allow for the wide physics range of LHCb to be achieved, only the most relevant triggers for the $B_s^0 \to \phi\phi$ decay, with $\phi \to K^+K^-$, will be discussed in this Section.

The majority of $B_s^0 \to \phi\phi$ candidates are provided through a HLT2 trigger searching for B decays involving a single ϕ meson. The reason for this strategy is due to the relatively low online track reconstruction efficiency. Therefore ϕ candidates are selected using loose K^+K^- invariant mass, vertex χ^2 and p_{T} cuts, in addition to a cut on the difference in the likelihoods of the kaon and pion hypotheses from RICH information.

While the inclusive ϕ trigger is important for selecting $B_s^0 \to \phi\phi$ decays, inclusive selections exploiting the topological structure of B decays are also useful. These can be applied in a cut-based method or by using a multi-variate classifier to enhance selection efficiencies. The principle of the topological triggers revolves around a corrected mass quantity, which is used to trigger B decays using a subset of the final state particles. The corrected mass is defined through

$$m_{corrected} = \sqrt{m^2 + |{p_{\mathrm{T}}}'_{missing}|^2} + |{p_{\mathrm{T}}}'_{missing}|, \tag{2.4}$$

where ${p_{\mathrm{T}}}'_{missing}$ is the missing momentum transverse to the direction of flight of the B meson candidate, and m is the invariant mass of the subset of the particles in the final state of the B decay candidate. The topological triggers require that $m_{corrected}$ lies in the range $4 < m_{corrected} < 7\,\mathrm{GeV}/c$. The robustness of the inclusive method to missing daughter particles through the $m_{corrected}$ variable means tight requirements can be imposed on the tracks. These consist of 1.5 GeV/c on the hardest track p_{T} and at least

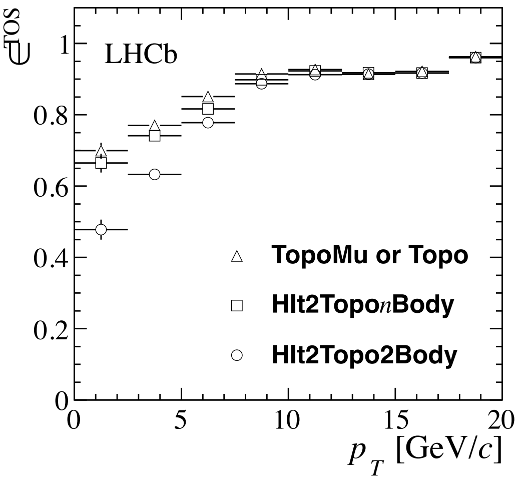


Figure 2.20.: Efficiency of the topological HLT2 triggers trigger as a function of p_{T} measured using $B^+ \to J/\psi \mathrm{K}^+$ candidates [47], where TopoMu or Topo means that at least one of the n-body topological or muon topological triggers selected the events, for $n = 2, 3$,

one track having χ^2 of the track fit less than 3. Also, in order to remove backgrounds from D decays, the $(n-1)$-body objects of the n-body topological trigger are required to have an invariant mass greater than $2.5\,\mathrm{GeV}/c^2$ or an IP χ^2 with respect to the primary vertex greater than 16. Due to the relatively small mass difference between pions and kaons compared to the B mass, all tracks are chosen to have the kaon mass.

The performance of the topological HLT2 triggers is shown in Figure 2.20 using $B^+ \to J/\psi \mathrm{K}^+$ candidates that have passed the HLT1 requirements [47], where TopoMu or Topo means that at least one of the n-body topological or muon topological triggers selected the events, Hlt2Topo$n$body means at least one of the n-body topological triggers selected the events, for $n \in \{2, 3\}$, and Hlt2Topo2Body means only the 2-body topological trigger selected the events. As for the case of the HLT1 trigger level, selection requirements for muon candidates are relaxed.

2.6. RICH HPD Efficiencies

During the course of 2011 data-taking, an effect on RICH HPD images started occurring, in which a central band became visible. Figure 2.21 shows the detected hits on a silicon pixel array in a single HPD over the course of an hour of data-taking. A central band with an increased quantity of hits can clearly be seen. This required further investigation in

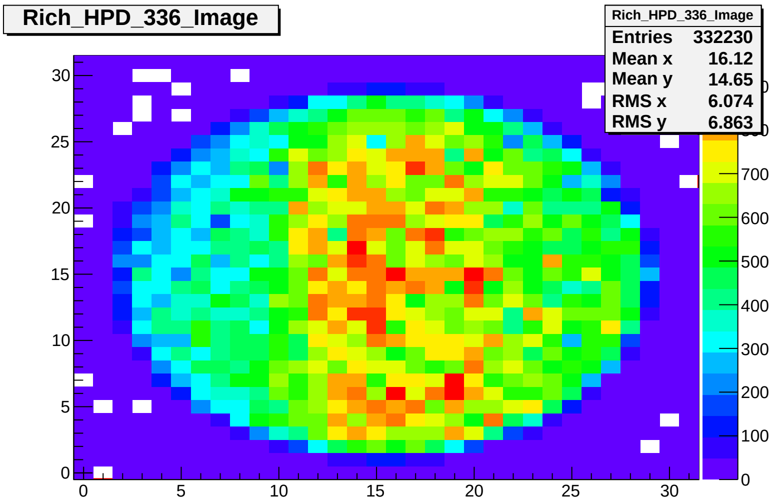


Figure 2.21.: Example of the banding effect seen in RICH HPDs.

to the possible causes of the effect. The following subsections detail the work performed on the causes and trends associated with the central band of increased detected hit-rate and the explanation of how the problem was fixed.

2.6.1. Origins, Trends and Solution

In order to understand the effect and possible trends, a way of quantifying the banding was defined. There are 484 HPDs present in the RICH detectors, therefore it is practical to analyse many images for each run with automation, where a run is a period of LHCb data-taking that can be one hour long at most. A program was written in order to achieve this automation. The program was designed to operate in the following way:

1. The centre of the image was found through the projection of the 2D image on to the x and y axes. A Gaussian fit was then performed to each projection. The resulting means (μ_x, μ_y) and widths (σ_x, σ_y) were then used to define a rectangular search region for the bands, where the cartesian coordinates (x, y) are inside the search region if $x < |\mu_x - 1.5\sigma_x|$ and $y < |\mu_y - 1.5\sigma_y|$.

2. In the search region, pixels are scanned row-by-row and images have been rotated to ensure central bands appear vertically. If the scanned pixel has a value more than 1.5× the value of a pixel scanned the time before last, a tally is incremented.

3. If the tally is greater than nine, then banding is said to be present in the HPD.

The methodology previously described was applied to a total of eighteen runs in 2011 and 2012 data, equating to a total of 8,712 images analysed. These runs were chosen to have varying conditions to identify possible trends. For example, they had varying trigger rates and varying instantaneous luminosities, but were chosen also to have lasted as long as possible, to have a larger number of events contributing. As can be seen from the black crossed points in Figure 2.22, a clear trend with the number of events read out from the L0 hardware trigger (termed L0 rate) can be seen. A trend with L0 rate points to the cause of the problem being related to the voltage distribution across the silicon pixel chip. In order to further isolate the causes, the occupancy difference was checked between high and low L0 rate runs to see if the banding effect could be correlated with a drop in photons recorded by the HPDs and therefore associated with a drop in efficiency. Figure 2.23 shows a fractional difference in occupancy for high and low L0 rate runs, where a drop in occupancy is seen with a high correlation with banding detection.

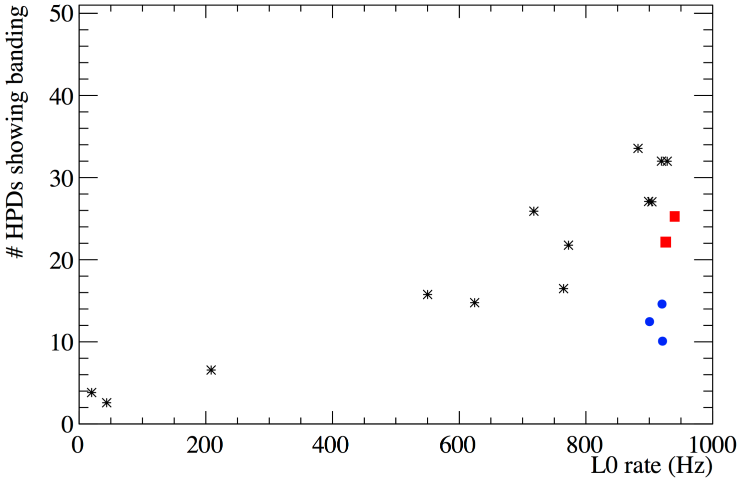


Figure 2.22.: Number of HPDs showing banding versus L0 rate. Black crosses show data points before the RICH settings change, red squares show data points after the setting changes were applied to RICH 1 and blue circles show data points after setting changes were applied to both RICH detectors.

A range of voltages must be supplied to the circuitry in the silicon pixel chip. This is achieved through the use of digital-to-analogue converters (DAC), which are configurable and must be externally biased by stable and precise reference voltages. The correct reference voltages were found and optimised during the testing of the hardware before operation in 2005 [49]. In the course of the studies conducted in 2005, it was noted that the measured output voltage of one of these reference voltages, referred to as DRHi, showed a stronger temperature dependence than the others. This is interesting as

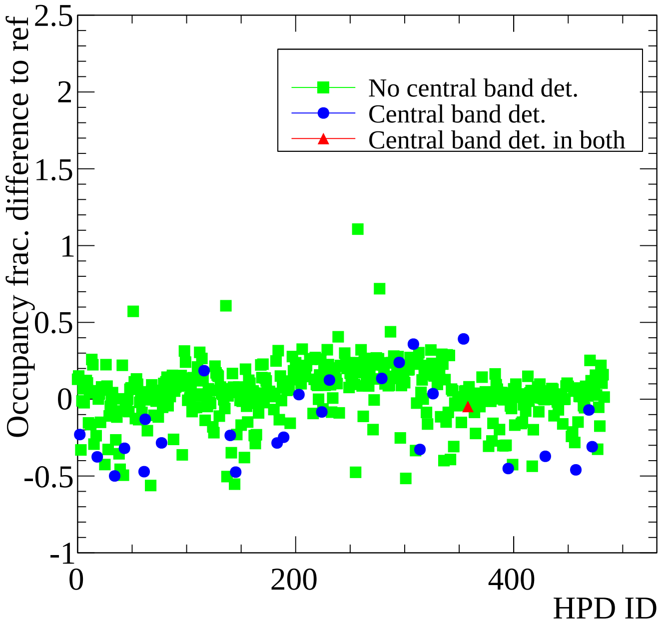


Figure 2.23.: Fractional change in occupancy between a low and high L0 rate runs for each HPD in the RICH detectors (identified by HPD ID). Green squares indicate no banding was detected, blue circles that banding was detected in the high L0 rate run but not in the low L0 rate run and red triangles that banding was detected in both runs.

the increased banding and decreased occupancies explained earlier show a dependence on the L0 rate. It is known that the temperature on the silicon pixel chip increases when the L0 rate increases as this is accompanied by an increase in photon detections. Further evidence that this was related to operating temperature became apparent in the monitoring of the day-to-day operation of the RICH detector. A decrease in the number of detected hits was seen when the detector was reset while proton-proton collisions were occurring. The temperature of the silicon pixel chip would therefore steadily increase during proton-proton collisions and a reset would therefore have been performed at a different temperature to the initial configuration.

The combination of the evidence of a temperature effect and a link to the DRHi voltage led to this voltage being optimised to see if a higher occupancy could be achieved during high L0 rate conditions. The optimisation resulted in the modification of the DRHi input voltage from 1.80V to 1.78V. This change was first applied for all HPDs in RICH 1 and then all HPDs in RICH 2 over the course of June to July 2012. The effect of this can be seen in Figure 2.22, where the two red squares show the results from runs in which the settings modification was only present in RICH 1 and the three blue dots from runs in which the settings modification was present in both RICH detectors. The level of banding detections decreased with both setting changes.

2.6.2. Relation to RICH Performance

At first glance, the central band shown in Figure 2.21 appears to be due to a decrease in efficiency at the edges of the silicon pixel array. However this may not necessarily be the case as this effect could equally be due to an excess of detected hits located in the higher occupancy region that do not originate from Cherenkov photons. It is also not apparent if any drop in efficiency corresponds directly to a drop in the efficiency of detecting Cherenkov photons from signal tracks. Therefore, it requires more evidence to link the central band to a drop in RICH performance. Particle identification (PID) performance is depicted using a 2D plot of efficiency versus mis-identification probability and is measured with data using $D^{*+} \to D^0(\to \mathrm{K}^+\pi^-)\pi^+$ decays. The reason for this is that if the identification efficiency increases, but the rate that the identification is incorrect also increases, then PID performance does not necessarily increase. Figure 2.24 shows the comparison in kaon identification of both a high L0 rate and a low L0 rate run before the setting changes described in the previous section and also the average performance in 2012 data (with field polarity down) before and after the setting changes. It can be seen that the low L0 rate performance did indeed have a better PID performance than that of the high L0 rate data before setting changes and that the PID performance improved with the changes in the DRHi setting.

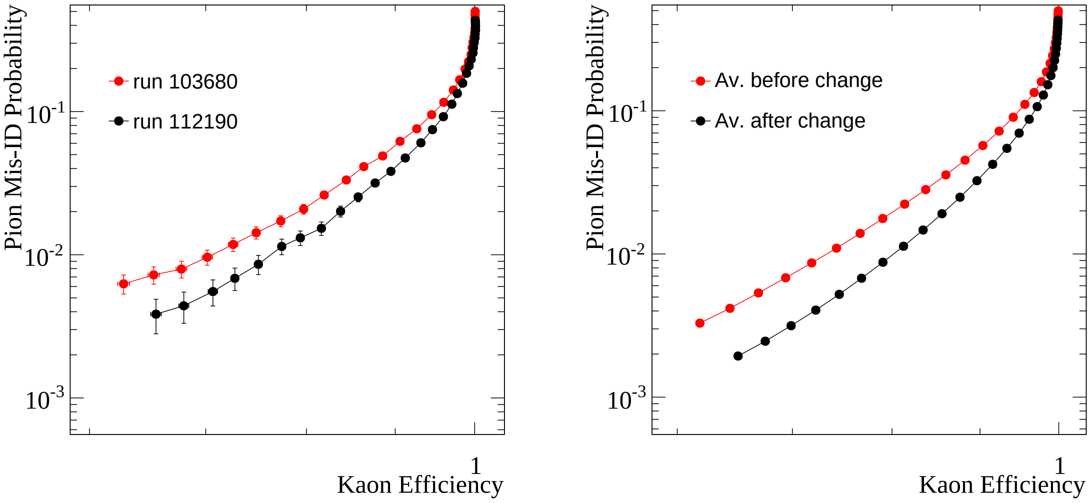


Figure 2.24.: PID performance for a single run consisting of an hour long data-taking period with high L0 rate (run 103680) and low L0 rate (run 112190) before setting changes (left), and average PID performances before and after setting changes (right).

Chapter 3.

Fit Methods

"I do not fear computers, I fear the lack of them."
— Isaac Asimov

Measurements of *CP* violation require that observables in the data be fitted to Probability Density Functions (PDFs). This is done through the minimisation of a quantity known as the Negative Log Likelihood (NLL). The NLL method is used extensively in this dissertation, not only for the extraction of *CP* and *T*-violating parameters, but also in the isolation of $B_s^0 \to \phi\phi$ candidates from the background.

3.1. Negative Log Likelihood Fitting

The likelihood function, $\mathcal{L}$, is defined as

$$\mathcal{L}(\vec{x}_1, ..., \vec{x}_N; \vec{a}) = \prod_{i=1}^{N} P(\vec{x}_i, \vec{a}), \tag{3.1}$$

where $\vec{x}_i$ represents the set of observables for the i^{th} event, with N in total, P is the PDF expected to describe the functional form of the observables seen in the data, and $\vec{a}$ is the set of parameters to be fitted. The values of the $\vec{a}$ parameters that are favoured by a given dataset are therefore those for which $\mathcal{L}$ achieves the maximum value. Alternatively, the values of $\vec{a}$ for which $-\ln\mathcal{L}$ achieves the minimum value is in more common use among the high energy physics community. The product of potentially very small numbers in the likelihood is transformed in to the sum of numbers in the NLL, and is therefore easier to deal with in practice. The values of $\vec{a}$ for which the NLL is minimised satisfy the simultaneous equations

$$-\frac{\partial}{\partial a_j}\ln\mathcal{L} = -\frac{\partial}{\partial a_j}\sum_{i=1}^{N}\ln P(\vec{x}_i,\vec{a}) = 0, \tag{3.2}$$

where a_j represents the j^{th} fit parameter.

3.1.1. Parameter Uncertainties

Having obtained the central values from fitting to the data, it is important to know the accuracy to which the fitted parameters are known. In the limit that the number of events in the dataset is large, the likelihood takes the form of a multi-dimensional Gaussian function [50]. This can be shown simply for the case of a likelihood of one parameter. A Taylor expansion around the fitted minimum provides the relation

$$-\ln\mathcal{L}(a) = -\ln\mathcal{L}(a_0) - (a-a_0)\frac{\mathrm{d}}{\mathrm{d}a}\ln\mathcal{L}(a)\bigg|_{a=a_0} - \frac{(a-a_0)^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}a^2}\ln\mathcal{L}(a)\bigg|_{a=a_0} + \mathcal{O}((a-a_0)^3), \tag{3.3}$$

where a_0 is the fitted central value. In the large N limit, $(a-a_0)$ is a small number, therefore only terms up to order $(a-a_0)^2$ need to be kept. The first derivative is 0 by definition at $a=a_0$. Exponentiating equation 3.3 therefore yields

$$\mathcal{L}(a) = -\mathcal{L}(a_0)\cdot\exp\left(-\frac{(a-a_0)^2}{2\sigma^2}\right), \tag{3.4}$$

where $\sigma^2 = -\left(\frac{\mathrm{d}^2}{\mathrm{d}a^2}\ln\mathcal{L}(a)\Big|_{a=a_0}\right)^{-1}$. The likelihood may then be seen to have a Gaussian shape[1] with standard deviation σ, hence variance σ^2. For the case of multiple fitted parameters, the variance generalises to a covariance matrix given by

$$\mathrm{cov}(a_i,a_j) = -\left(\frac{\partial^2\ln\mathcal{L}}{\partial a_i\partial a_j}\bigg|_{\vec{a}=\vec{a}_0}\right)^{-1}, \tag{3.5}$$

where the values on the diagonal provide the variances for the individual parameters. The covariance matrix, in addition to providing the errors of parameters, also allows for

---

[1] This means that the NLL will appear to be parabolic in the large N limit.

the definition of the correlation matrix, $\rho(a_i, a_j)$, through

$$\rho(a_i, a_j) = \frac{\mathrm{cov}(a_i, a_j)}{\sigma_{a_i}\sigma_{a_j}}, \tag{3.6}$$

where σ_{a_i} refers to the standard deviation of parameter a_i. The elements of the correlation matrix are known as correlation coefficients and describe the extent to which parameters depend on each other. The coefficients themselves range between -1 and 1, where a value of 0 implies no correlation.

3.1.2. Nuisance Parameters

It is often the case that a PDF has parameters that are not directly of interest, but are required to be known. Examples include experimental resolutions in the case of fits to the invariant mass or the oscillation frequency of the B_s^0 meson in the case of fitting for time-dependent *CP* asymmetries. In some cases, it is possible to fit for such nuisance parameters directly, however it is often advantageous to use external information to aid the fit. In some cases, where there is negligible uncertainty, the value of the parameter may be fixed to that obtained from the external source. Often there is a substantial uncertainty. The NLL method allows for this uncertainty to be accounted for with the use of Gaussian constraints. This then increases the statistical uncertainty of the unconstrained parameters to reflect the uncertainty on the constrained parameter. Gaussian constraints are implemented in the likelihood through

$$\mathcal{L}(\vec{x}_1, ..., \vec{x}_N; \vec{a}) \to \mathcal{L}(\vec{x}_1, ..., \vec{x}_N; \vec{a}) \prod_{k=1}^{M} \exp\left(-\frac{(a_k - \bar{a}_k^c)^2}{2\sigma^2_{a_k^c}}\right), \tag{3.7}$$

where a_k is the k^{th} constrained parameter of a total M to be constrained, $\bar{a}_k^c$ is the central value of the external input, and $\sigma_{a_k^c}$ is the associated error. In the context of an NLL, the extension is represented as

$$-\ln\mathcal{L}(\vec{x}_1, ..., \vec{x}_N; \vec{a}) \to -\ln\mathcal{L}(\vec{x}_1, ..., \vec{x}_N; \vec{a}) - \sum_{k=1}^{M} \left(-\frac{(a_k - \bar{a}_k^c)^2}{2\sigma^2_{a_k^c}}\right). \tag{3.8}$$

The introduction of Gaussian constraints then allows for some systematic uncertainties to be elegantly included in the statistical uncertainty.

3.2. The Feldman Cousins Method and Confidence Intervals

In the case that small datasets are fitted, where small implies that the Gaussian limit of the likelihood has not yet been established, it becomes more important that the definition of probability being used is understood. There are two definitions in common usage among the high energy physics community which affect the interpretation of a given uncertainty. These are termed Bayesian and frequentist. The uncontroversial Bayes theorem states that

$$p(a|b)p(b) = p(b|a)p(a), \tag{3.9}$$

where $p(a|b)$ denotes the probability of observing a given b and $p(a)$ denotes the probability of a. The application of Bayes theorem to the testing of an experimental result is an example of subjective probability as a prior PDF is needed to describe the distribution of parameters. The interpretation of probability found through frequentist statistics describes the frequency of the observation in a repeatable experiment. For results presented in this dissertation, the frequentist approach is followed.

In the previous section, the assumption of the Gaussian limit allowed for a parameter uncertainty, σ to be provided. This then gave the range of values $[a-\sigma, a+\sigma]$, for which the parameter a would be observed in a repeatable experiment $68.27\,\%$ of the time. Such an interpretation is available from the Neyman construction of a confidence interval. Using the PDF, $P(x, a)$, where x represents an observable, a pre-specified probability, denoted by $(1-\alpha)$, can be related to bounds on x for a given value of a through

$$1 - \alpha = \int_{x_1}^{x_2} P(x, a)\mathrm{d}x. \tag{3.10}$$

Values of x_2 and x_1 can be found for each value of a that satisfies equation 3.10 for the chosen value of $(1-\alpha)$ before any measurement has even been performed. This is shown in Figure 3.1, where the region between the two curves $x_1(a)$ and $x_2(a)$ is known as the confidence belt. After a measurement has been made, the measured value of x_0 allows the confidence interval $[a^-, a^+]$ to be found. This is the principle of the Neyman construction of a confidence interval. The Neyman construction allows for the choice of the test statistic. In the previous description of the Neyman construction, this consisted of the observable itself. Another popular choice is the likelihood ratio or equivalently log

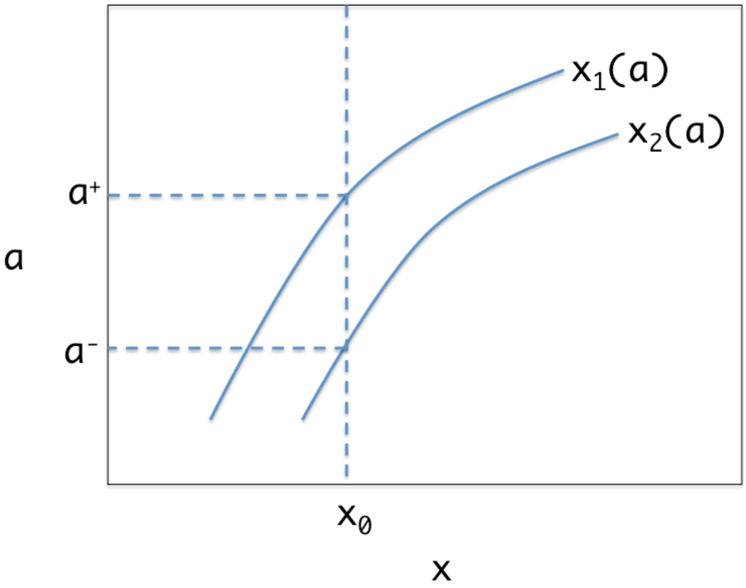


Figure 3.1.: Illustration of the construction of Neyman confidence intervals for a one parameter (a), one observable (x) PDF, with the observable as the test statistic.

likelihood ratio, λ_R, defined through

$$\lambda_R \equiv \frac{\ln \mathcal{L}}{\ln \mathcal{L}_0}, \tag{3.11}$$

where $\mathcal{L}_0$ indicates the maximum value of the likelihood. The lower and upper bounds of the 68.27 % confidence interval correspond to $2\Delta \ln \mathcal{L} \equiv 2(\ln \mathcal{L} - \ln \mathcal{L}_0)$ of value 1 in the case of the estimation of a single parameter in the Gaussian limit. This can be seen by directly observing the likelihood ratio

$$\frac{\ln \mathcal{L}}{\ln \mathcal{L}_0} = -\ln \mathcal{L}_0 + \left(\frac{(a-a_0)^2}{2\sigma^2}\right) + \ln \mathcal{L}_0, \tag{3.12}$$

with the substitution $\sigma = a - a_0$. The values of the likelihood and associated confidence levels are shown in Table 3.1, for the joint estimation of up to three parameters [10]. It is due to the well known values in the Gaussian limit given in Table 3.1, that scans of the likelihood ratio, referred to simply as likelihood scans, provide a powerful tool not only in the evaluation of statistical uncertainties, but also in the determination of the validity of the Gaussian limit.

When the assumption of the Gaussian limit does not apply, the establishment of a confidence interval that covers the required probability correctly becomes more difficult.

C.L. (%)	$n = 1$	$n = 2$	$n = 3$
68.27	1.00	2.30	3.53
90.00	2.71	4.61	6.25
95.45	4.00	6.18	8.03
99.73	9.00	11.83	14.16

Table 3.1.: Values of the $2\Delta \ln \mathcal{L}$ and associated confidence levels for the joint estimation of n parameters [10].

It is for this reason that Cousins and Feldman [51] introduced a method utilising the freedom provided by the Neyman construction to solve the problem. This freedom is that of ordering according to a test statistic. For the purposes of Feldman-Cousins confidence levels in this dissertation, the test statistic used is λ_R. The Feldman-Cousins method is then evaluated as follows for a confidence level of probability α:

1. The dataset is fitted to find the parameters for which the NLL achieves the minimum value.

2. For each value of the parameter of interest, a_i, the value of the likelihood is calculated. Note that the values of all other parameters are allowed to vary as in (1).

3. A large number of simulated datasets are then generated for $a = a_i$. Nuisance parameters when generating simulated datasets are set to be the same as that found in the fit to the data in (1).

4. Each simulated dataset is fitted with the value of a fixed to a_i and also fitted with a allowed vary, the result being denoted as $\tilde{a}_i$. The test statistic, $\lambda^k_{R,i} = \ln \mathcal{L}^k(a_i)/\ln \mathcal{L}^k(\tilde{a}_i)$ is then evaluated for each simulated dataset, k.

5. The fraction of events, α_i for which $\lambda_R < \lambda_R^{\text{data}}$ is calculated, where $\lambda_R^{\text{data}} = \ln \mathcal{L}(a_i)/\ln \mathcal{L}(a_0)$ and a_0 is the overall best fit value of the parameter a in the data.

6. If the value of $\alpha_i > \alpha$ then a_i lies inside the Feldman-Cousins confidence region.

7. This procedure is followed until all of the boundaries of the required confidence region have been established.

3.3. Inclusion of Efficiency Corrections in the Log Likelihood

One of the most powerful features of log likelihood fitting is the ability to include efficiencies that can be factorised from fitted parameters directly as numerical weights, without the need of a histogram to describe the efficiency. This has been used extensively in this dissertation to account for the efficiency as a function of the helicity angles, defined in Section 1.4.2. The so-called method of normalisation weights to describe efficiencies as a function of helicity angles has been well established in measurements of *CP* violation in $B_s^0 \to J/\psi\,\phi$ decays [52, 53].

Recall from equation 3.2 that maximising the log likelihood amounts to solving the equation

$$\frac{\partial \ln \mathcal{L}}{\partial a_j} = \frac{\partial}{\partial a_j} \sum_i \ln \frac{s(\vec{x_i}|\vec{a})}{\int s(\vec{x}|\vec{a}) \mathrm{d}\vec{x}} = 0, \tag{3.13}$$

where s indicates an unnormalised signal PDF, $\vec{a}$ is the set of parameters to be fitted, $\vec{x}$ is the set of observables and a subscript i refers to a sum over events. An efficiency $\varepsilon(\vec{y})$ over a subset of observables $\vec{y} \subset \vec{x}$ can be included in equation 3.13 through

$$\frac{\partial \ln \mathcal{L}}{\partial a_j} = \frac{\partial}{\partial a_j} \sum_i \ln \frac{s(\vec{x_i}|\vec{a})\varepsilon(\vec{y_i})}{\int\int s(\vec{y},\vec{z}|\vec{a})\varepsilon(\vec{y}) \mathrm{d}\vec{y}\mathrm{d}\vec{z}} = 0, \tag{3.14}$$

where $\vec{z}$ denotes the rest of the observables that are not included in $\vec{y}$. In the case where an efficiency does not depend on any of the fitted parameters as is the case of efficiency as a function of helicity angles in the $B_s^0 \to \phi\phi$ decay, the logarithmic identity $\ln(AB) = \ln(A) + \ln(B)$ may be used such that the efficiency no longer needs to be included in the numerator. Therefore equation 3.14 reduces to

$$\frac{\partial \ln \mathcal{L}}{\partial a_j} = \frac{\partial}{\partial a_j} \sum_i \ln \frac{s(\vec{x_i}|\vec{a})}{\int\int s(\vec{y},\vec{z}|\vec{a})\varepsilon(\vec{y}) \mathrm{d}\vec{y}\mathrm{d}\vec{z}} = 0. \tag{3.15}$$

Normalisation weights, ξ_j, may be defined through the equation

$$\xi_j \equiv \int f_j(\vec{y})\varepsilon(\vec{y}) \mathrm{d}\vec{y}, \tag{3.16}$$

where it has been assumed that the PDF factorises according to $s(\vec{y},\vec{z}) = \sum_i K_i(\vec{z}) f_i(\vec{y})$. It should be noted that this PDF form is the same as that of the angular distribution for $B_s^0 \to \phi\phi$ decays. Other parameters that can effect whether or not an event is accepted or rejected such as the transverse momentum or impact parameter may be denoted as $\vec{w}$. Note that this requires a re-definition of $\vec{z}$ to be the components of $\vec{x}$ that are not in $\vec{y}$ or $\vec{w}$. The efficiency can then be written as

$$\varepsilon(\vec{y}) = \frac{\int \varepsilon(\vec{y}|\vec{w}) S(\vec{z},\vec{y},\vec{w}|\vec{a}) \mathrm{d}\vec{w}}{S(\vec{z}|\vec{a})}. \tag{3.17}$$

The equation for the weights may then be written as

$$\xi_j = \int \frac{\int f_j(\vec{y}) \varepsilon(\vec{y},\vec{w}) S(\vec{z},\vec{y},\vec{w}|\vec{a}) \mathrm{d}\vec{w}}{S(\vec{z},\vec{y}|\vec{a})} \mathrm{d}\vec{y}, \tag{3.18}$$

where $S(\vec{z},\vec{y},\vec{w}|\vec{a})$ now denotes the normalised PDF. Noting that $S(\vec{z},\vec{y},\vec{w}|\vec{a}) \mathrm{d}\vec{w} \mathrm{d}\vec{y}$ is the probability to generate an event at a given time with a given set of $\vec{a}$ with observables between $[\vec{y},\vec{y}+\mathrm{d}\vec{y}]$ and $[\vec{w},\vec{w}+\mathrm{d}\vec{w}]$, equation 3.18 may then be written as

$$\xi_j \approx \frac{1}{N_{\text{gen}}} \sum_{i \in \{\text{generated}\}} \frac{f_j(\vec{y_i}) \varepsilon(\vec{y_i},\vec{w_i})}{S(\vec{z_i},\vec{y_i},\vec{w_i}|\vec{a})}, \tag{3.19}$$

$$\xi_j = \frac{1}{N_{\text{gen}}} \sum_{i \in \{\text{accepted}\}} \frac{f_j(\vec{y_i})}{S(\vec{z_i},\vec{y_i},\vec{w_i}|\vec{a})}, \tag{3.20}$$

where different values of ξ_j can be found for different bins of $\vec{z}$ if required. However, for the purposes of the analyses in this dissertation, normalisation weights are assumed to be constant with $\vec{z}$. This then means that equation 3.15 reduces to

$$\frac{\partial}{\partial a_k} \ln \mathcal{L}(\vec{a}) = \frac{\partial}{\partial a_k} \sum_i \frac{K_n(\vec{z_i}) f_n(\vec{y_i})}{\int K_m(\vec{z},\vec{a}) \xi_m \mathrm{d}\vec{z}} = 0, \tag{3.21}$$

where the normalisation weights are found using equation 3.20 and repetition of an index indicates summation.

3.4. The ${}_s\mathcal{P}lot$ Method

In the following chapters, heavy use is made of the ${}_s\mathcal{P}lot$ technique [54] in order to optimise the selection of the $B_s^0 \to \phi\phi$ decay and also in the measurement of the *CP*-violating phase in the decay time-dependent analysis of $B_s^0 \to \phi\phi$. The ${}_s\mathcal{P}lot$ technique allows for distributions of variables, for example helicity angles or the B_s^0 decay time, to be observed independently for a given fit component, using a discriminating variable. For all use cases in this dissertation, the discriminating variable used is the four-kaon invariant mass. This is only possible for the case where the variables of interest and discriminating variable are uncorrelated. In the ${}_s\mathcal{P}lot$ technique, so-called s-weights for the n^{th} component of a fit to the discriminating variable (with N_s species present in the sample), defined through

$$
{}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}, \tag{3.22}
$$

where y_e is the value of the discriminating variable for the e^{th} event, f_j refers to the PDF of the j^{th} component of the fit to the discriminating variable, y, and N_k represents the yield of the fit component k. The matrix V_{nj} is the $N_s \times N_s$ covariance matrix of the species yields defined through

$$
V_{nj}^{-1} = \frac{\partial^2(-\mathcal{L})}{\partial N_n \partial N_j} = \sum_{e=1}^{N} \frac{f_n(y_e) f_j(y_e)}{(\sum_{k=1}^{N_s} N_k f_k(y_e))^2}. \tag{3.23}
$$

The distribution of a control variable can then be obtained through plotting with the associated s-weight, which will on average reproduce the true distribution of the control variable. If a discriminating variable is used that is correlated with a control variable, a bias will be introduced on the s-weights that will be difficult to understand and hence correct for, especially if the PDFs used to fit the discriminating variable do not completely describe the dataset.

It is important when using s-weights that uncertainties are taken in to account properly. For a given bin in a histogram of a control variable, the uncertainty on a given bin is simply the square root of the sum of the s-weights in that bin.

3.4.1. Application of the ${}_s\mathcal{P}lot$ Method to an Unbinned Likelihood Fit

A powerful feature of s-weights, used to a large extent in the time-dependent analysis of $B_s^0 \to \phi\phi$, is the ability to use s-weights to disentangle a required component of a dataset from other components in an unbinned NLL fit to the control variables. To do this, the s-weight for a given event is applied as a multiplicative factor when building the likelihood, i.e.

$$\mathcal{L}(\vec{x}_1, ..., \vec{x}_N; \vec{a}) = \prod_{i=1}^{N} P(\vec{x}_i, \vec{a}) \rightarrow \prod_{i=1}^{N} P(\vec{x}_i, \vec{a}) \cdot {}_s\mathcal{P}_n(y_i), \tag{3.24}$$

where y_i is the value of the observable y for the i^{th} event, that is not contained nor correlated with any of the observables $\{\vec{x}_1, ..., \vec{x}_N\}$. The naive use of the s-weights in equation 3.24 will lead to undercoverage as the errors on the s-weights themselves have not been taken in to account. This can be corrected by scaling the s-weights before the application in equation 3.24 with an α factor calculated as [54]

$$\alpha = \frac{\sum_i {}_s\mathcal{P}_n(y_i)}{\sum_i {}_s\mathcal{P}_n(y_i)^2}. \tag{3.25}$$

3.5. Summary

The principles underlying log likelihood fitting have been introduced including the methods of obtaining the central values of parameters and associated uncertainties in the Gaussian limit. Confidence levels have been explained along with the concept of coverage, which is especially relevant when datasets are not large enough to establish the Gaussian limit.

The so-called normalisation weights have been described, which allow efficiencies dependent on observables that may be factorised from fitted parameters to be corrected for elegantly in the fit.

The ${}_s\mathcal{P}lot$ method has been introduced, which finds uses not only in the development of data-driven selections but also directly in the likelihood, to isolate the $B_s^0 \to \phi\phi$ signal distributions of the helicity angles and B_s^0 decay time when fitting for CP violation.

Chapter 4.

Isolating the $B_s^0 \to \phi\phi$ decay

"Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organisation of the entire tapestry."
— Richard P. Feynman

Proton-proton collisions provide a challenging environment to observe $B_s^0 \to \phi\phi$ decays with each $\phi \to \mathrm{K}^+\mathrm{K}^-$. This is due to the high number of tracks from the primary vertex (PV) that provide large backgrounds to the $B_s^0 \to \phi\phi$ signal. The $\mathrm{K}^+\mathrm{K}^-\mathrm{K}^+\mathrm{K}^-$ final state is used as the $\phi \to \mathrm{K}^+\mathrm{K}^-$ branching fraction is $\mathcal{B}(\phi \to \mathrm{K}^+\mathrm{K}^-) = (48 \pm 0.5)\,\%$ [10]. The particle identification (PID) offered by the LHCb RICH detectors in the form of differences in the log likelihoods between the kaon and pion mass hypotheses ($\mathrm{DLL}_{K\pi}$) is therefore essential to separate the $B_s^0 \to \phi\phi$ signal from the background. The application of the ${}_s\mathcal{P}lot$ technique, described in detail in Section 3.4, allows for a data-driven method for obtaining the optimal selection requirements. This is used to obtain optimum values for cut-based selections, described in detail in Section 4.1, and to obtain optimum requirements on the multi-variate classifier, described in detail in Section 4.2.

The two different methods of isolating the $B_s^0 \to \phi\phi$ decay are introduced for two reasons. The main reason is that the cut-based selection was used to create the dataset used for decay time integrated measurements, described in Chapter 5, while the multi-variate classifier was used for the decay time dependent analysis of the $B_s^0 \to \phi\phi$ decay, described in Chapter 6.

4.1. Cut-based Isolation

4.1.1. Dataset

The optimisation of the requirements of the cut-based selection was performed on 220 pb^{-1} of LHCb data collected during the course of 2011 data-taking. The optimisation method described in Section 4.1.2 allows for a purely data-driven optimisation, hence no simulated datasets were used. The dataset was produced with the initial version of the reconstruction available at the time the data was collected. This so-called prompt reconstruction[1] makes use of the alignment and calibrations available at the time of data-taking. The selection at trigger level required that events passed the hadron, electron and muon triggers at the L0 stage. In addition, events were required to pass the HLT1 triggers selecting all tracks from L0 and the corresponding muon HLT1 trigger. At the HLT2 stage, events were required to pass any of the topological BDT triggers or the inclusive ϕ trigger. The details of the triggers used may be found in Section 2.5.

4.1.2. Optimisation Method and Results

The metric used to judge the quality of a given set of requirements, known as a Figure of Merit (FoM), in the cut-based selection is $S/\sqrt{S+B}$, where S and B refers to the number of signal and background candidates passing a given set of requirements, respectively. This FoM was recursively optimised with the use of the Cut Recursive OPtimiser program (CROP) [55]. In order to first use CROP, a set of loose requirements, known as a preselection, was applied to the dataset in order to create a $B_s^0 \to \phi\phi$ signal peak in the $\mathrm{K}^+\mathrm{K}^-\mathrm{K}^+\mathrm{K}^-$ mass spectrum. A peak is required such that a fit may be performed to assign s-weights as described in Section 3.4, allowing further optimisation to take place. The preselection criteria, shown in Table 4.1, consisted of:

- Impact parameter (IP) χ^2 of the B_s^0 meson and the kaon tracks with respect to the primary vertex,
- transverse momentum, p_T, of the kaon tracks and the product of the p_T from each ϕ meson,
- difference in log likelihoods between the kaon and pion mass hypotheses ($\mathrm{DLL}_{K\pi}$),

---

[1] Reconstruction version 10, selected with DaVinci version 28r3p1.

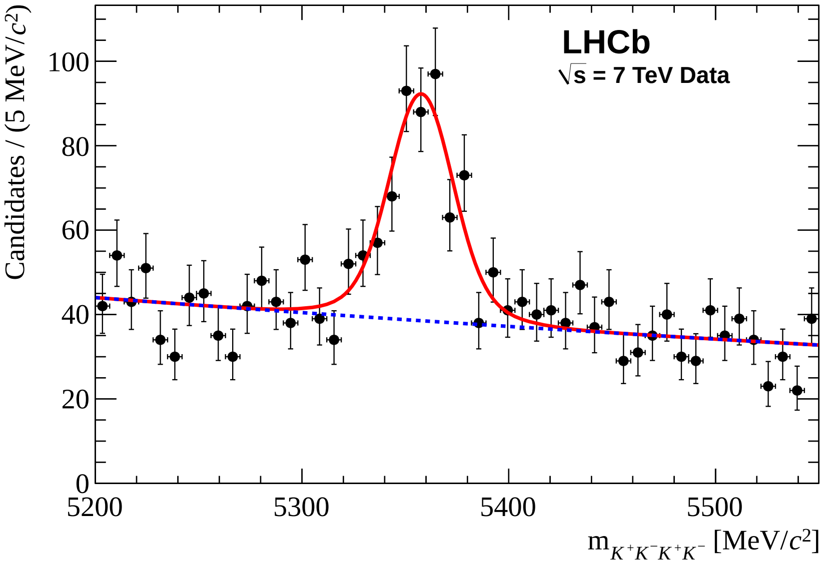


Figure 4.1.: $K^+K^-K^+K^-$ invariant mass distribution for $B_s^0 \to \phi\phi$ candidates after preselection events found in the 220 pb^{-1} dataset. The result of a fit to a Gaussian signal and exponential background component is superimposed.

- χ^2 per Number of Degrees of Freedom (NDF) of the ϕ and B_s^0 vertex fits,
- invariant mass of the $K^+K^-K^+K^-$ final state,
- invariant mass of the K^+K^- pairs originating from ϕ mesons,
- χ^2 of the B_s^0 Flight Distance (FD) with respect to the PV.

After the preselection was applied, a suitable signal was seen for optimisation. This is shown in Figure 4.1 together with a fit to a Gaussian signal and exponential background component. The associated signal and background s-weights for each event, calculated using the method described in Section 3.4, were then used to unfold the data into the signal and background components. The s-weighted distributions for each variable in the optimisation resulting from these signal and background components are shown in Figure 4.2, where the separation power for each of the variables can clearly be seen.

The optimisation of the FoM is performed by scanning through the possible combinations of values of the variables used for optimisation. At each given scan point, the requirements are applied and the FoM is calculated. The final values chosen by the optimisation are those for which the FoM is maximised. The optimisation was performed in a signal region defined to be $5200 < m_{K^+K^-K^+K^-} < 5550\,\text{MeV}/c^2$. The final requirements, found via the optimisation procedure, are shown in Table 4.1. The

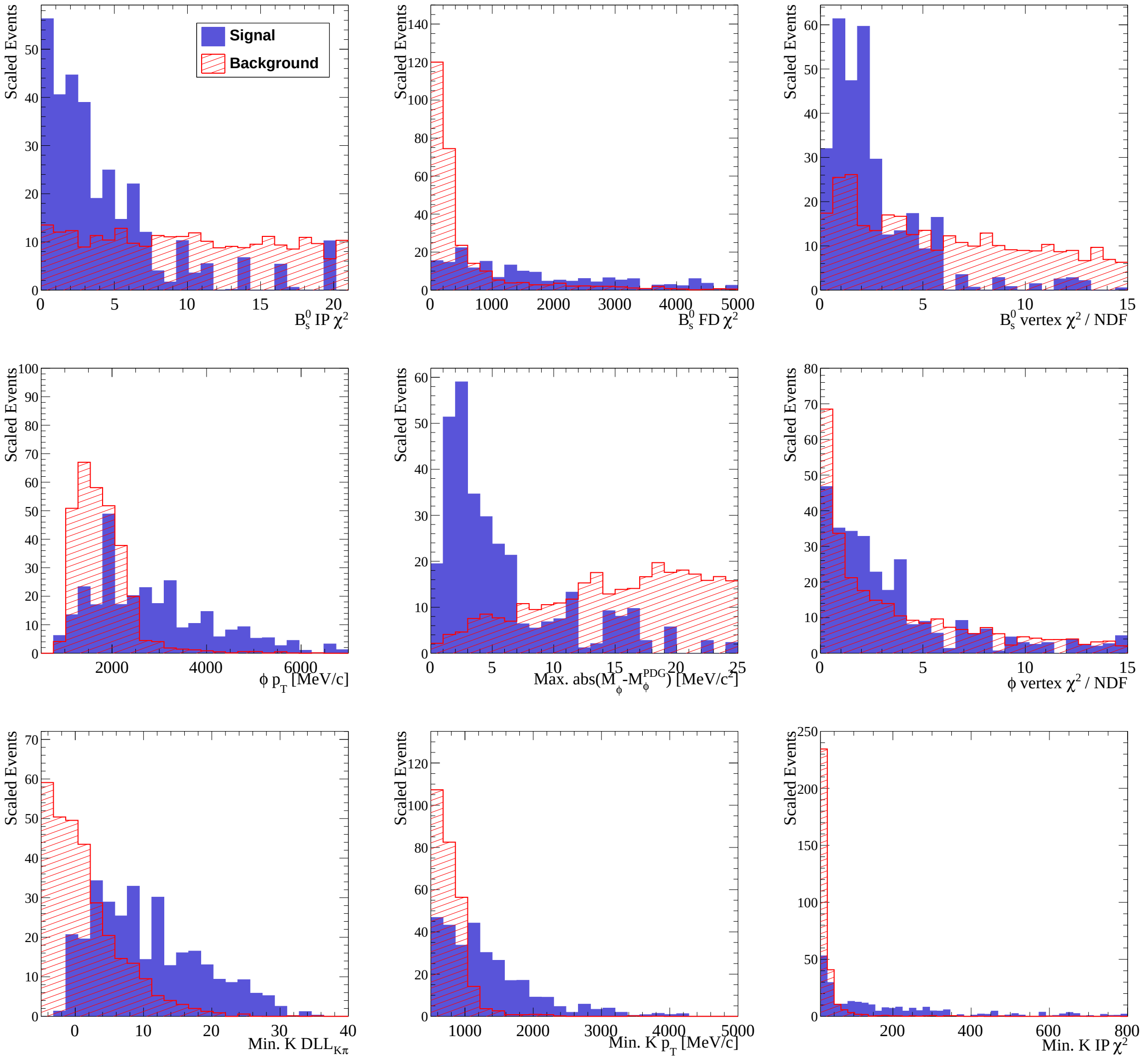


Figure 4.2.: Signal and background s-weighted distributions for the variables used in the cut-based selection optimisation, where the signal and background histograms are scaled to have the same area.

requirements with optimised values found to be close to the preselection values were allowed to remain at the preselection value.

With a cut-based optimisation, it is important not to use variables that are heavily correlated. The introduction of such variables can lead to decreased performance in the optimisation as the order in which the selections are optimised becomes important. Before optimisation, the input variables were checked for the degree of correlation. The results of this check are shown in Figure 4.3. It is clear from Figure 4.3 that very little correlation can be seen between variables with the exception of the χ^2 of the fit to the

Requirement	Preselection value	Optimised value
$\lvert M_{\mathrm{K^+K^-K^+K^-}} - M_{B_s^0}^{PDG}\rvert$ [MeV/c^2]	< 300	
K IP χ^2	> 15	> 21
K p_{T} [MeV/c]	> 500	
Min. DLL$_{K\pi}$	> -5	> 0
ϕ vertex χ^2 per NDF	< 25	
ϕ^1 p_{T} $\times$ ϕ^2 p_{T} [(GeV/c)^2]	> 2	
$\lvert M_\phi - M_\phi^{PDG}\rvert$ [MeV/c^2]	< 25	< 12
B_s^0 vertex χ^2 per NDF	< 15	< 7.5
B_s^0 FD χ^2	> 100	> 270
B_s^0 IP χ^2	< 25	< 15

Table 4.1.: Preselection and optimised requirements used in the cut-based selection. Events passing the preselection were used as input to the s-weight based cut optimisation, where NDF, IP and FD refer to the number of degrees of freedom, impact parameter and flight distance, respectively, and DLL$_{K\pi}$ is the difference in the global likelihood between the kaon and pion mass hypotheses from RICH subdetector information.

vertex of the B_s^0 meson and the χ^2 of the fit to the vertex of the ϕ meson. Selection efficiencies were calculated from simulated events. The efficiencies of selections applied individually from simulation are shown in Table 4.2. Note that the efficiencies quoted from simulation are relative to simulated events that have been reconstructed by the LHCb detector, passing the nominal trigger requirements with a $\mathrm{K^+K^-K^+K^-}$ invariant mass within 500 MeV/c^2 of the PDG B_s^0 mass. The nominal trigger requirements, used throughout this dissertation unless stated otherwise consist of TOS requirements on the L0 hadron trigger[2], TOS requirements on the HLT1 trigger using all tracks from the L0 stage, and TOS requirements on the topological and Inclusive ϕ HLT2 triggers. The signal-to-background ratio is calculated from 1 fb^{-1} of 2011 data to be 11.7, in the four-kaon invariant mass range $5200 < m_{\mathrm{K^+K^-K^+K^-}} < 5550$ MeV/c^2.

4.2. Multi-variate Isolation

The aim of a multi-variate classifier is to use the information for a given event in the most efficient way possible. As a simple example, consider a cut-based optimisation consisting

---

[2] Recall that TOS requires that the trigger was passed with the signal candidate and associated tracks.

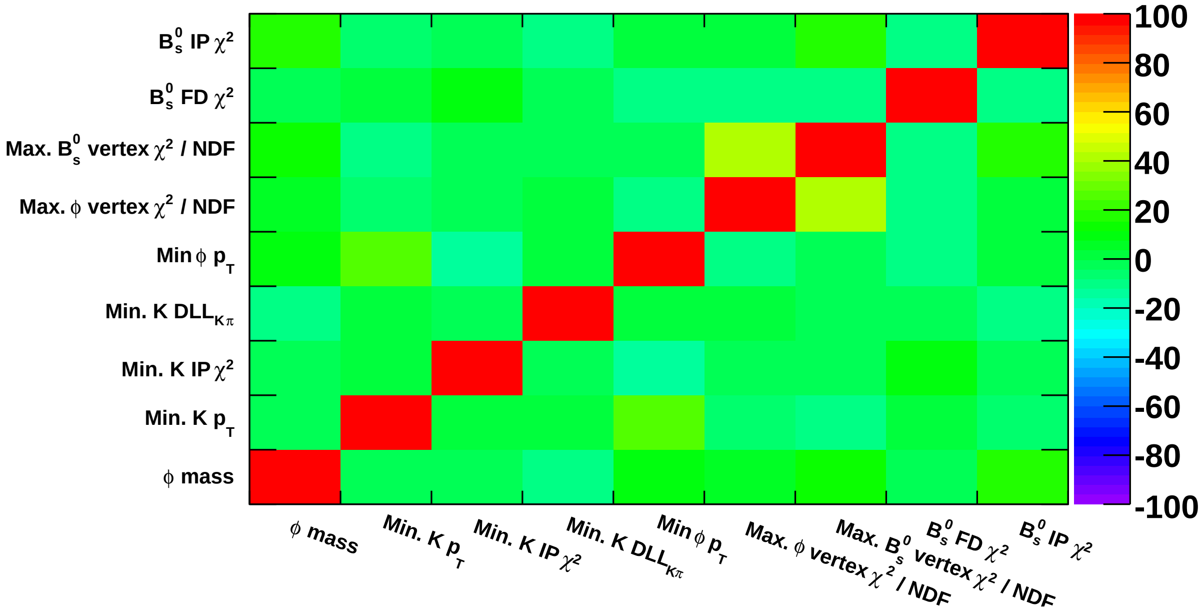


Figure 4.3.: Degree of correlation between input variables used in the cut-based requirement optimisation, where correlation is provided in percent.

Variable	Efficiency (%)
K IP χ^2	89.18 ± 0.12
K p_{T}	96.81 ± 0.07
Min. DLL$_{K\pi}$	92.28 ± 0.10
ϕ vertex χ^2 per NDF	96.88 ± 0.07
ϕ^1 $p_{\mathrm{T}} \times \phi^2$ p_{T}	100
$\lvert M_\phi - M_\phi^{PDG} \rvert$	80.14 ± 0.16
B_s^0 vertex χ^2 per NDF	98.77 ± 0.04
B_s^0 FD χ^2	94.44 ± 0.09
B_s^0 IP χ^2	96.88 ± 0.07
Total	60.36 ± 0.19

Table 4.2.: Exclusive selection efficiencies for individual requirements from the cut-based optimisation calculated from simulated $B_s^0 \to \phi\phi$ events.

of the B_s^0 IP χ^2 and the χ^2 of the B_s^0 vertex fit. In the case of a signal event that has a vertex fit χ^2 that is slightly worse than the optimised requirement but an IP χ^2 that is much smaller than the optimised requirement, then the event would not pass based on the one failed variable. However, as will be explained in the next Section, an optimised multi-variate classifier is in theory able to use the information provided by all variables to provide better separation of the signal from the background.

Multi-variate isolation techniques include Boosted Decision Trees (BDTs) [56] and Artificial Neural Networks (ANN) [57] that may be trained on signal and background samples, using variables that provide discrimination between the two. The purpose of the multi-variate techniques is then to combine all knowledge of the signal and background contributions gained from the training sample in to a single variable, known as a multi-variate classifier.

4.2.1. Boosted Decision Trees

The underlying principle of Boosted Decision Trees (BDTs) is that of the decision tree itself. A decision tree works by initially taking a variable and finding the value of the variable that provides the best separation between the signal and background. In this way, the training samples are divided into two branches created from the original sample, termed the Root Node. This procedure is repeated for each variable until a certain purity is reached, or when further division will result in too few events being present in the resulting samples. Note that samples in a given branch are termed nodes and the node at the end of a branch is termed a leaf. A weight for the leaf is calculated as the signal purity and the leaf is given the label as signal or background depending on the dominant contribution to the leaf. This principle of operation is illustrated in Figure 4.4.

Decision trees have been in use since the mid 1980s and are known to be powerful but unstable as small changes in the training sample can cause significant changes in the decision tree. Reliability has been ensured more recently through the use of boosting. The principle of boosting relies on the re-weighting of signal events that have been wrongly classified as background and vice versa. Practically, this means first building a decision tree and finding cases of signal events landing on a background leaf or background events landing on a signal leaf. The weights of such events are then increased (boosted) and a new decision tree is created with the new weights. This procedure is repeated until many trees, typically 1000, are created. The boosting used in this dissertation is known as the AdaBoost algorithm [56] and is defined as follows: Let x_i represent the set of discriminating variables for the i^{th} event and $\mathcal{F}_m(x_i)$ be the value returned by the m^{th} decision tree for the set x_i, where $\mathcal{F}$ returns 1 if the event lands on a signal leaf and -1 if the event lands on a background leaf. Let also, $\mathcal{F}_i^{\text{true}}$ be the true origin of the event, where $\mathcal{F}_i^{\text{true}}$ returns 1 for signal and -1 for background. The AdaBoost algorithm then

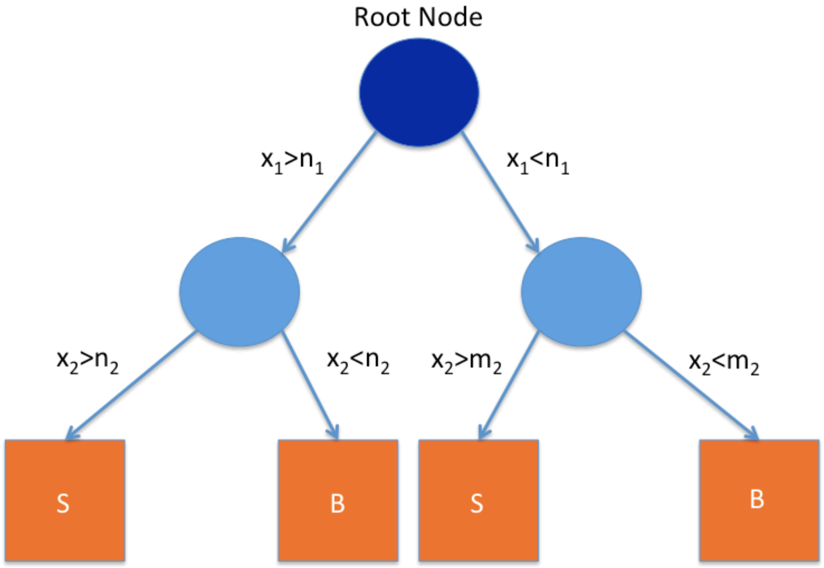


Figure 4.4.: Diagram showing the principle of a decision tree. For a given variable, x_i, the value giving the best separation, $n_i(m_i)$, is found. This is repeated for all variables until a given signal purity is reached, or until a minimum number of event in the node is reached. A category is then assigned to the leaf depending on whether signal or background is the dominant contribution.

transforms the weight, w_i^m, assigned to the i^{th} event by the m^{th} decision tree through

$$w_i^{m+1} = w_i^m e^{\alpha_m I(\mathcal{F}_i^{\text{true}} \neq \mathcal{F}_m(x_i))}, \tag{4.1}$$

where

$$\alpha_m = \ln\left(\frac{1-E_m}{E_m}\right) \tag{4.2}$$

and

$$E_m = \frac{\sum_i w_i I(\mathcal{F}_i^{\text{true}} \neq \mathcal{F}_m(x_i))}{\sum_i w_i}. \tag{4.3}$$

The function $I(\mathcal{F}_i^{\text{true}} \neq \mathcal{F}_m(x_i))$ returns 1 if $\mathcal{F}_i^{\text{true}} \neq \mathcal{F}_m(x_i)$ else 0 is returned. This ensures that only mis-classified events are boosted. The value of the BDT response for a given event is then evaluated as

$$\mathcal{F}(x_i) = \sum_m \alpha_m \mathcal{F}_m(x_i). \tag{4.4}$$

Requirements can then be placed solely on the multi-variate classifier to obtain the required signal to background ratio.

4.2.2. Datasets

The optimisation of the multi-variate selection was performed with datasets produced with the reconstruction containing updated alignment and calibrations[3]. In total, $1.0\,\mathrm{fb}^{-1}$ of LHCb data collected during 2011 data-taking has been used for multivariate optimisations. The selection at trigger level required that events have passed the hadron trigger as TOS or are selected independently of the $B_s^0 \to \phi\phi$ candidate on any trigger at the L0 stage. In addition, events were required to pass the HLT1 trigger selecting all tracks from L0 as TOS. At the HLT2 stage, events were required to pass any of the topological multivariate BDT triggers or the inclusive ϕ trigger as TOS. The details of the triggers used may be found in Section 2.5.

Simulated events were created under the same conditions as the data itself. The parameters used to simulate the distributions of helicity angles and decay time are shown in Table 4.3. In total, 1 million simulated events were generated for each LHCb dipole magnet polarity, giving 2 million simulated events in total.

Parameter	Value
$\lvert A_0\rvert^2$	0.348
$\lvert A_\perp\rvert^2$	0.365
$\lvert A_\parallel\rvert^2$	0.287
δ_0 (rad)	0.0
$\delta_\parallel$ (rad)	2.71
$\delta_\perp$ (rad)	2.39
$\phi^{s s\bar{s}}$ (rad)	0.0

Table 4.3.: Parameters used for $B_s^0 \to \phi\phi$ Monte Carlo production, where polarisation fractions are taken from the values measured by the CDF collaboration [27].

Requirement	Preselection value	Optimised value
$\lvert M_{K^+K^-K^+K^-} - M_{B_s^0}^{PDG}\rvert$ [MeV/c^2]	< 300	
K IP χ^2	> 16	
K p_T [MeV/c]	> 500	
Min. DLL$_{K\pi}$	> -5	> -3
ϕ vertex χ^2 per NDF	< 25	
ϕ^1 p_T × ϕ^2 p_T [(GeV/c)^2]	> 2	
$\lvert M_{K^+K^-} - M_{\phi}^{PDG}\rvert$ [MeV/c^2]	< 25	< 15
B_s^0 vertex χ^2 per NDF	< 15	
B_s^0 FD χ^2	> 100	
B_s^0 IP χ^2	< 25	
BDT		> 0.1

Table 4.4.: Preselection and optimised requirements used in the multivariate selection.

4.2.3. BDT Training Method

The TMVA package [58] is used to train BDT variables. The signal sample is defined by the simulated $B_s^0 \to \phi\phi$ events that pass the preselections given in Table 4.4. The background sample was obtained from LHCb data that passes the preselections given in Table 4.4, where at least one ϕ candidate lies in the K^+K^- invariant mass sidebands, defined by ($20\,\text{MeV}/c^2 < \lvert M_{K^+K^-} - M_{\phi}^{PDG}\rvert < 25\,\text{MeV}/c^2$). The emphasis in choice of variables used in the BDT has been on reducing the bias on the decay time. Therefore, the B_s^0 impact parameter χ^2 and the B_s^0 flight distance χ^2 have not been used in the BDT training. The variables used to create the BDT isolation variable were:

- $\ln(p_T)$ of the B_s^0 candidate,
- B_s^0 vertex χ^2 per degree of freedom (DOF),
- cosine of the angle between the momentum of the particle and the direction of flight from the best primary vertex (PV) to the decay vertex,
- B_s^0 pseudo-rapidity, η,
- minimum $\ln(p_T)$ of each track,
- maximum track χ^2 per DOF.

---

[3] Reconstruction version 12, selected with DaVinci version 32r2.

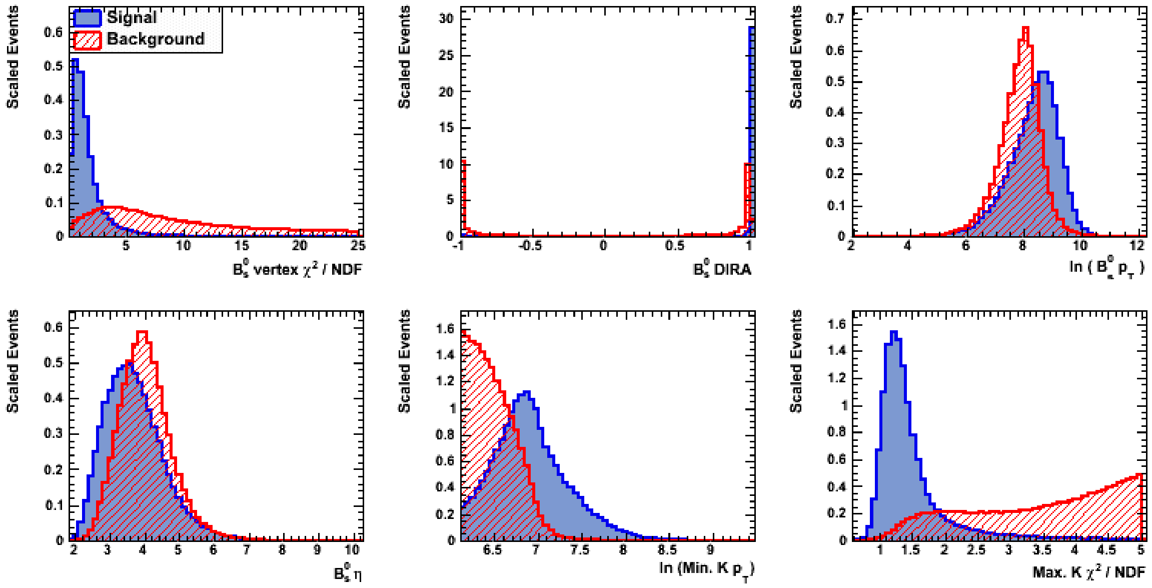


Figure 4.5.: Distributions of BDT training variables from simulated signal and background data samples, where background candidates are defined by at least one ϕ candidate in the $\mathrm{K^+K^-}$ invariant mass sidebands, $(20\,\mathrm{MeV}/c^2 < |M_{\mathrm{K^+K^-}} - M_\phi^{PDG}| < 25\,\mathrm{MeV}/c^2)$. Signal and background histograms are scaled to have the same area.

The distributions of the variables used in BDT training from signal and background samples are shown in Figure 4.5.

In addition to the BDT incorporating the adaptive boost, the TMVA package also provides a BDT method incorporating a gradient boost (BDTG); designed to achieve more robustness in the presence of outliers and mislabelled data points through the use of an alternate boosting function. Also, a BDT method incorporating transformations that decorrelate the input variables was trained (BDTD). The interested reader is directed to reference [58] for further information on the additional decorrelation transformations and boosting functions. It can be clearly seen that the nominal BDT provides the highest significance, hence this was the one used for final selections. Figure 4.6 shows the signal efficiencies, purities and significances of the range of BDT cuts for the three types of trained BDTs, where 1000 signal events and 15000 background events are assumed, representing the size of the preselected data sample obtained by applying the selections in Table 4.4.

It is important when using multivariate methods, that overtraining does not occur. Overtraining becomes evident if large differences are seen between the BDT response of training and test samples. Such overtraining is usually the result of too few events

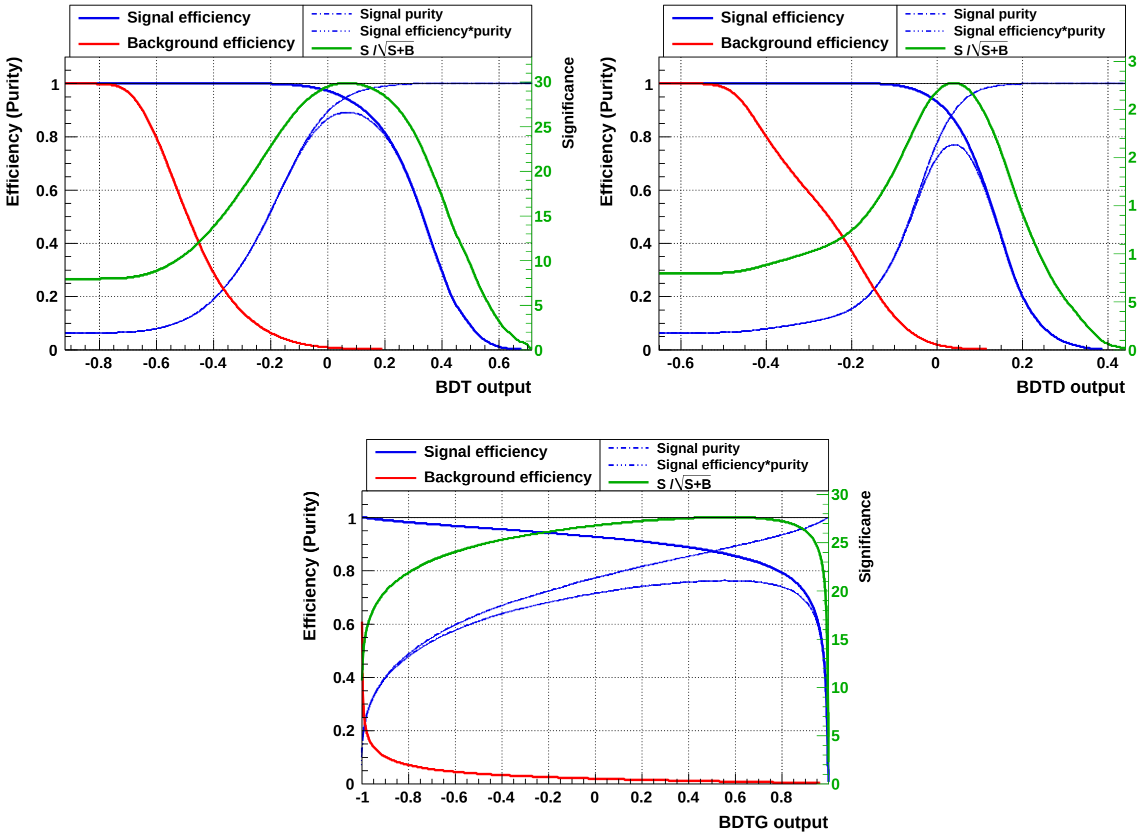


Figure 4.6.: Signal efficiency (blue, solid line), background efficiency (red, solid line), signal purity (long, dotted line), signal efficiency multiplied by purity (short, dotted line) and significance (green, solid line) of the BDT (left), BDGD (centre) and BDTG (right), where there are 1000 signal events and 15000 background events assumed.

being present in the training sample or by providing too many degrees of freedom to the multi-variate classifier. The separation between signal and background samples is shown in Figure 4.7, where the signal and background components consist of 50000-event samples. As can be seen, the BDT response works well with no signs of overtraining. Correlation of the input variables is less important for the case of BDTs, due to the robustness provided by boosting. The correlation matrix between the input variables is shown in Figure 4.8. The expected correlation is seen between the K and B_s^0 kinematic variables, all other variables are found to be uncorrelated.

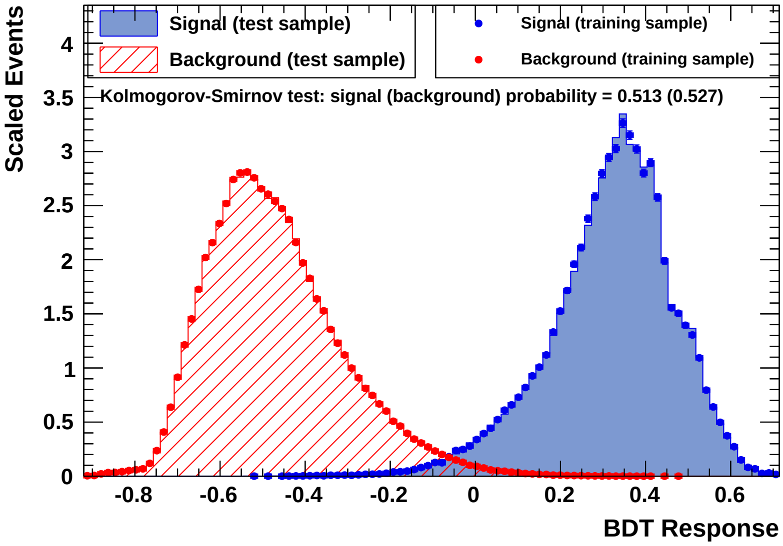


Figure 4.7.: BDT response to signal and background samples consisting of 50 thousand events each. Good separation is observed with no sign of overtraining.

4.2.4. Optimisation

In addition to the BDT isolation variable, the ϕ invariant mass and $\mathrm{DLL}_{K\pi}$ are required to further separate signal events from background events[4]. The standard figure of merit used in optimisations, $S/\sqrt{S+B}$, is not the ideal choice for this as it does not take in to account other factors that affect the sensitivity to the CP-violating phase. These factors include:

- The B_s^0 decay time resolution,
- flavour tagging performance,
- shape of the background distribution (in both time and angular variables).

The most important factors concerning the sensitivity to CP violation are the tagging performance and the variation of signal versus background with decay time. As such, a different FoM is used to account for these factors. To account for purity variations over decay time, events were split into six decay time bins of equal width spanning the time range -3.0 to $12.0\,\mathrm{ps}$. The figure of merit (Q) used for optimisation is therefore given

---

[4] The $\mathrm{DLL}_{K\pi}$ variable was not used in the BDT training as this is poorly modelled in simulated events.

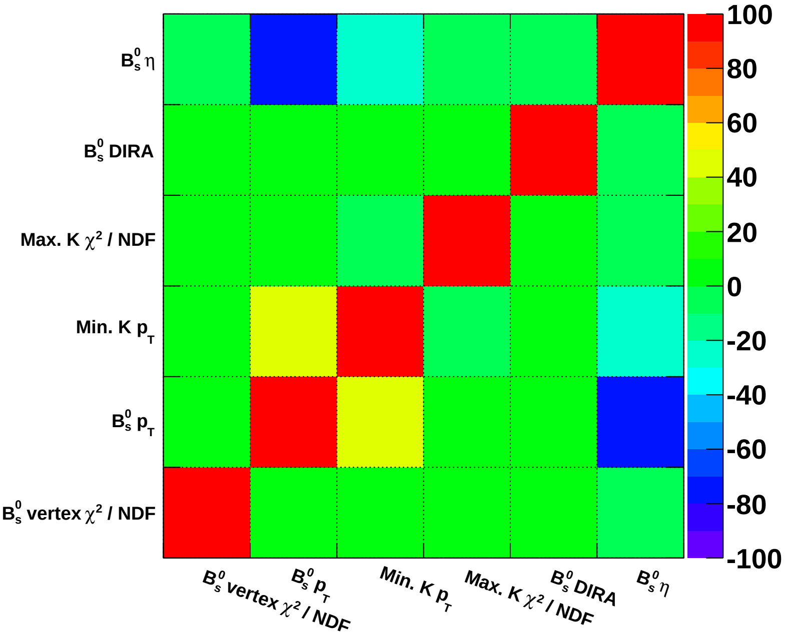


Figure 4.8.: Correlation matrix for the input variables to the multi-variate classifier, where correlation is provided in percent.

by [59]

$$Q = \sum_{k} (f^k)^2 \sum_{l=1}^{N_s^k + N_b^k} (1 - 2 \cdot \omega_l)^2 \cdot \rho_l^s, \tag{4.5}$$

where N_s^k and N_b^k are the number of signal events and background events in the kth decay time bin, respectively; ω_l is the mistag probability of the l^{th} event; f_l^k is the local purity factor (defined as $S/(S+B)$ in the kth decay time bin) and ρ_l^s is the s-weight of the event.

The BDT response was added to $1.0\,\text{fb}^{-1}$ of LHCb data, triggered and reconstructed as described in Section 4.2.2 and passing the preselections given in Table 4.4. This dataset was then s-weighted according to the reconstructed B_s^0 mass using the $_s\mathcal{P}lot$ method [54] to unfold the data into signal and background components for the cut optimisation. The values of the BDT, DLL$_{K\pi}$ and ϕ mass range used to obtain the final selection was then optimised using the CROP package [55]. The optimised cut values are given in Table 4.4. The efficiencies of the final selections in the multi-variate optimisation are shown in Table 4.5, where it can bee seen that an improved efficiency is achieved compared with that found in the cut-based requirement optimisation. Note that the efficiencies quoted from simulation are relative to simulated events that have been fully reconstructed by the LHCb detector, passing the nominal trigger requirements with a $\mathrm{K^+K^-K^+K^-}$ invariant mass within $500\,\mathrm{MeV}/c^2$ of the PDG B_s^0 mass. The signal to background ratio is calculated from $1\,\mathrm{fb}^{-1}$ of 2011 data to be 4.25, in the four-kaon invariant mass range $5200 < m_{\mathrm{K^+K^-K^+K^-}} < 5550$ MeV/c^2.

Variable	Efficiency (%)
K IP χ^2	92.26 ± 0.10
K p_{T}	96.81 ± 0.07
Min. $\mathrm{DLL}_{K\pi}$	98.75 ± 0.04
ϕ vertex χ^2 per NDF	98.59 ± 0.05
ϕ^1 p_{T} $\times$ ϕ^2 p_{T}	100
$\lvert M_\phi - M_\phi^{PDG}\rvert$	84.33 ± 0.14
B_s^0 vertex χ^2 per NDF	100
B_s^0 IP χ^2	98.59 ± 0.05
BDT	98.96 ± 0.04
Total	72.86 ± 0.17

Table 4.5.: Offline selection efficiencies for individual requirements from the multi-variate optimisation calculated from simulated $B_s^0 \to \phi\phi$ events.

4.3. Summary

Two different optimisations have been performed to separate the $B_s^0 \to \phi\phi$ signal from the relatively large backgrounds. The two methods have been the traditional cut-based optimisation and a multi-variate classifier in the form of a BDT. As expected, the BDT-based method has been found to have an improved efficiency compared with that found in the cut-based method. These efficiencies have been calculated from simulated events to be $(72.86 \pm 0.17)\,\%$ and $(60.36 \pm 0.19)\,\%$ for the multi-variate and cut-based methods, respectively. The multi-variate and cut-based optimisations achieve signal to background ratios of 4.25 and 11.7, respectively, in the four-kaon invariant mass range $5200 < m_{\mathrm{K^+K^-K^+K^-}} < 5550\,\mathrm{MeV}/c^2$ calculated from $1\,\mathrm{fb}^{-1}$ of 2011 data.

Chapter 5.

Measurement of the Polarisation Amplitudes and Triple Product Asymmetries in $B_s^0 \to \phi\phi$ Decays

“Time you enjoy wasting, was not wasted”
— John Lennon

This section describes in detail the measurements of the polarisation fractions ($|A_0|^2$, $|A_\parallel|^2$, $|A_\perp|^2$), strong phase difference ($\cos\delta_\parallel$) and triple product asymmetries that are accessible from determining the helicity angles, defined in Section 1.4.2. The measurement was published in Physics Letters B during 2012 [60].

5.1. Dataset

The decay time-integrated measurements were performed using approximately $1\,\text{fb}^{-1}$ of LHCb data collected during 2011[1]. The analysis used the cut-based selection, defined in Table 4.1. The selection at trigger level required that events passed the hadron, electron and muon triggers at the L0 stage. In addition, events were required to pass the HLT1 triggers selecting all tracks from L0 and the corresponding muon HLT1 trigger. At the HLT2 stage, events were required to pass any of the topological BDT triggers or the

---

[1] Samples were produced with version 29r2 of the DaVinci application and made use of version 12 of the reconstruction.

inclusive ϕ trigger. The requirement that events merely pass the trigger is not optimal as this does not necessarily mean that the $B_s^0 \to \phi\phi$ candidate was responsible. The use of this instead of the requirement that the $B_s^0 \to \phi\phi$ candidate alone would have triggered the event (TOS) allowed for a larger data sample to be analysed. Therefore, in the measurement of the polarisation fractions and strong phase difference, a simultaneous fit was performed for events selected as TOS at each trigger stage and those which were not, denoted TOS and non-TOS, respectively. Simulated events used for the purposes of acceptance corrections were created under the same conditions as the data itself. The parameters used to simulate the distributions of helicity angles and decay time are shown in Table 4.3.

To determine the signal yield an unbinned maximum likelihood fit was performed to the $K^+K^-K^+K^-$ invariant mass distribution. The $K^+K^-K^+K^-$ invariant mass of the $B_s^0 \to \phi\phi$ signal component was modelled by two Gaussian functions with a common mean. The width of the first Gaussian was measured from data to be 13.9 ± 0.6 MeV/c^2. The relative fraction, f_2, and width, $\sigma_{B_s^0,2}$, of the second Gaussian are fixed to 0.215 and 29.5 MeV/c^2 respectively, where values have been obtained from simulation. The background was described by an exponential function, with associated slope α_{BKG}. Figure 5.1 shows the four-kaon invariant mass distribution for selected events. A signal yield of 801 ± 29 events was found when fitting the PDF described above to the data. A goodness-of-fit test based on the point-to-point dissimilarity method [61], yielded a p-value of 0.62, indicating that the fit result is in good agreement with the data distribution.

5.2. Polarisation Amplitudes

5.2.1. Angular Distributions

The time-dependent differential decay rate for the $B_s^0 \to \phi\phi$ mode (derived in detail in Appendix A) can be written as

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_1 \mathrm{d}\cos\theta_2 \mathrm{d}\Phi \mathrm{d}t} \propto F(t, \theta_1, \theta_2, \Phi) = \sum_{i=1}^{6} K_i(t) f_i(\theta_1, \theta_2, \Phi), \tag{5.1}$$

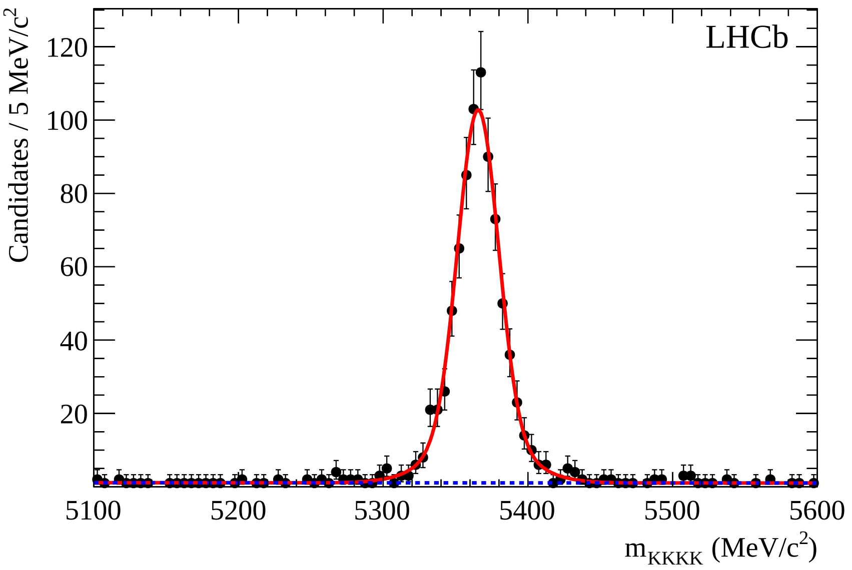


Figure 5.1.: Invariant $K^+K^-K^+K^-$ mass distribution for selected $B_s^0 \to \phi\phi$ candidates. A fit of a double Gaussian signal component together with an exponential background (dotted line) is superimposed.

where the θ_1, θ_2 and Φ are the helicity angles defined in Figure 1.7, and angular functions $f_i(\theta_1, \theta_2, \Phi)$ are given by [62]

$$\begin{aligned}
f_1(\theta_1,\theta_2,\Phi) &= 4\cos^2\theta_1\cos^2\theta_2, \\
f_2(\theta_1,\theta_2,\Phi) &= \sin^2\theta_1\sin^2\theta_2(1+\cos 2\Phi), \\
f_3(\theta_1,\theta_2,\Phi) &= \sin^2\theta_1\sin^2\theta_2(1-\cos 2\Phi), \\
f_4(\theta_1,\theta_2,\Phi) &= -2\sin^2\theta_1\sin^2\theta_2\sin 2\Phi, \\
f_5(\theta_1,\theta_2,\Phi) &= \sqrt{2}\sin 2\theta_1\sin 2\theta_2\cos\Phi, \\
f_6(\theta_1,\theta_2,\Phi) &= -\sqrt{2}\sin 2\theta_1\sin 2\theta_2\sin\Phi. \qquad (5.2)
\end{aligned}$$

The time-dependent functions $K_i(t)$ are given by [63]

$$\begin{aligned}
K_1(t) &= \frac{1}{2}|A_0|^2[(1+\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm L}t} + (1-\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm H}t} \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi^{ss\bar{s}}],\\
K_2(t) &= \frac{1}{2}|A_\parallel|^2[(1+\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm L}t} + (1-\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm H}t} \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi^{ss\bar{s}}],\\
K_3(t) &= \frac{1}{2}|A_\perp|^2[(1-\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm L}t} + (1+\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm H}t} \mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi^{ss\bar{s}}],\\
K_4(t) &= |A_\parallel||A_\perp|[\pm e^{-\Gamma_s t}\{\sin\delta_1\cos(\Delta m_s t) - \cos\delta_1\sin(\Delta m_s t)\cos\phi^{ss\bar{s}}\}\\
&\quad -\frac{1}{2}(e^{-\Gamma_{\rm H}t} - e^{-\Gamma_{\rm L}t})\cos\delta_1\sin\phi^{ss\bar{s}}],\\
K_5(t) &= \frac{1}{2}|A_0||A_\parallel|\cos(\delta_2-\delta_1)\\
&\quad [(1+\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm L}t} + (1-\cos\phi^{ss\bar{s}})e^{-\Gamma_{\rm H}t} \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi^{ss\bar{s}}],\\
K_6(t) &= |A_0||A_\perp|[\pm e^{-\Gamma_s t}\{\sin\delta_2\cos(\Delta m_s t) - \cos\delta_2\sin(\Delta m_s t)\cos\phi^{ss\bar{s}}\}\\
&\quad -\frac{1}{2}(e^{-\Gamma_{\rm H}t} - e^{-\Gamma_{\rm L}t})\cos\delta_2\sin\phi^{ss\bar{s}}], \qquad (5.3)
\end{aligned}$$

where the upper of the ± or ∓ signs refers to the B_s^0 meson and the lower refers to a $\overline{B}_s^0$ meson. Here, $\Gamma_{\rm L}$ and $\Gamma_{\rm H}$ are the decay widths of the light and heavy B_s^0 mass eigenstates,[2] Δm_s is the B_s^0 oscillation frequency, $\delta_1 = \arg(A_\perp/A_\parallel)$ and $\delta_2 = \arg(A_\perp/A_0)$ are *CP*-conserving strong phases, and $\phi^{ss\bar{s}}$ is the weak *CP*-violating phase. It was assumed that the weak phase is zero in line with the SM expectation [22]. The quantities $\Gamma_{\rm H}$ and $\Gamma_{\rm L}$ correspond to the observables $\Delta\Gamma_s = \Gamma_{\rm L} - \Gamma_{\rm H}$ and $\Gamma_s = (\Gamma_{\rm L} + \Gamma_{\rm H})/2$.

Integrating over time and ignoring the initial flavour of the B_s^0 meson, equation 5.1 becomes

$$F(\theta_1,\theta_2,\Phi) = \sum_{i=1}^{6} K_i^{\rm TI} f_i(\theta_1,\theta_2,\Phi), \qquad (5.4)$$

where

$$\begin{aligned}
K_1^{\rm TI} &= |A_0|^2/\Gamma_{\rm L}, &(5.5)\\
K_2^{\rm TI} &= |A_\parallel|^2/\Gamma_{\rm L}, &(5.6)\\
K_3^{\rm TI} &= |A_\perp|^2/\Gamma_{\rm H}, &(5.7)\\
K_4^{\rm TI} &= 0, &(5.8)\\
K_5^{\rm TI} &= |A_0||A_\parallel|\cos(\delta_\parallel)/\Gamma_{\rm L}, &(5.9)\\
K_6^{\rm TI} &= 0. &(5.10)
\end{aligned}$$

---

[2] Units are adopted such that $\hbar = 1$.

The strong phase difference is given by $\delta_\parallel \equiv \delta_2 - \delta_1 = \arg(A_\parallel/A_0)$, and the time integration assumes uniform time acceptance. Ignorance of the production flavour assumes an equal number of B_s^0 and $\bar{B}_s^0$ mesons are produced. The effect of the assumption of no production asymmetry was found from simulated events to introduce no significant uncertainty.

5.2.2. Fit Details and Constraints

The form of the PDF, $P(m, \theta_1, \theta_2, \Phi)$, used in the fitting of the data explicitly parametrises the background component resulting in

$$P(m, \theta_1, \theta_2, \Phi) = f_s \cdot F(\theta_1, \theta_2, \Phi) \cdot G(m) + (1 - f_s) \cdot \tilde{F}(\theta_1, \theta_2, \Phi) \cdot \tilde{G}(m), \qquad (5.11)$$

where m is the $K^+K^-K^+K^-$ invariant mass, f_s is the signal fraction, $F(\theta_1, \theta_2, \Phi)$ is the $B_s^0 \to \phi\phi$ time-integrated angular distribution, described in Section 5.2.1, $G(m)$ is the double Gaussian B_s^0 mass distribution, described in Section 5.1, $\tilde{F}(\theta_1, \theta_2, \Phi)$ is the flat distribution parametrising the angular dependence of the combinatoric background, and $\tilde{G}(m)$ is the exponential function used to describe mass dependence of the combinatoric background. The requirements placed on the $B_s^0 \to \phi\phi$ candidates were found from simulated events to allow only a negligible fraction of B-related backgrounds to contribute to the final dataset. The parameters found in the functions forming the PDF in equation 5.11 are listed in Table 5.1, along with how each parameter was treated in the fit.

Function	Parameter	Type	Value
$F(\theta_1, \theta_2, \Phi)$	$\|A_0\|^2$, $\|A_\perp\|^2$	Free	
	$\|A_\parallel\|^2$	Constrained	$1 - \|A_0\|^2 - \|A_\perp\|^2$
	$\cos\delta_\parallel$	Free	
	Γ_s, $\Delta\Gamma_s$	Constrained	LHCb measurement [52]
$G(m)$	$m_{B_s^0}$	Free	
	$\sigma_{B_s^0}$	Free	
	$\sigma_{B_s^0,2}$	Fixed	29.5 MeV/c^2
	f_2	Fixed	0.215
$\tilde{G}(m)$	α_{BKG}	Free	

Table 5.1.: Summary of the parameters contained in the functions of equation 5.11 with associated fit treatment.

The kaon impact parameter selections required to separate signal from background, and the small dataset sample create a difficult environment to measure the lifetimes of the physical B_s^0 mass eigenstates, Γ_s and $\Delta\Gamma_s$. However this difficulty can be easily overcome through the use of Gaussian constraints to the LHCb measurements obtained from the $B_s^0 \to J/\psi\,\phi$ decay of [52]

$$\Gamma_s = 0.657 \pm 0.009(\text{stat}) \pm 0.008(\text{syst})\,\text{ps}^{-1}, \tag{5.12}$$

$$\Delta\Gamma_s = 0.123 \pm 0.029(\text{stat}) \pm 0.011(\text{syst})\,\text{ps}^{-1}. \tag{5.13}$$

The constraint was applied taking in to account the correlation between Γ_s and $\Delta\Gamma_s$ of $\rho(\Gamma_s, \Delta\Gamma_s) = -0.30$.

5.2.3. Angular Acceptance

As explained in Section 3.3, efficiencies that are not dependent on fitted parameters may be corrected for with the use of so-called acceptance weights. These weights were determined from fully selected simulated events for the case of the efficiency as a function of helicity angles. The one-dimensional efficiencies of the helicity angles are shown in Figure 5.2. The decrease in efficiency seen as $\cos\theta_{1,2}$ approach ±1 is largely due to the p_{T} requirements imposed to separate the $B_s^0 \to \phi\phi$ signal from the background.

As mentioned in Section 5.1, TOS and non-TOS events were fitted simultaneously. This allowed for separate angular acceptance weights to be used for each dataset. The weights used compared to those found from combining both datasets are shown in Table 5.2.

Angular function	Weights (triggered)	Weights (TOS)	Weights (non-TOS)
f_1	0.946 ± 0.009	0.931 ± 0.010	0.943 ± 0.011
f_2	1.029 ± 0.014	1.037 ± 0.015	1.030 ± 0.017
f_3	1.025 ± 0.012	1.032 ± 0.013	1.028 ± 0.014
f_4	0.019 ± 0.019	0.032 ± 0.020	-0.001 ± 0.023
f_5	-0.004 ± 0.014	0.007 ± 0.016	-0.020 ± 0.017
f_6	-0.001 ± 0.012	0.004 ± 0.013	-0.012 ± 0.015

Table 5.2.: Angular acceptance weights determined from fully selected simulated events, used for the measurement of the polarisation amplitudes and strong phase difference.

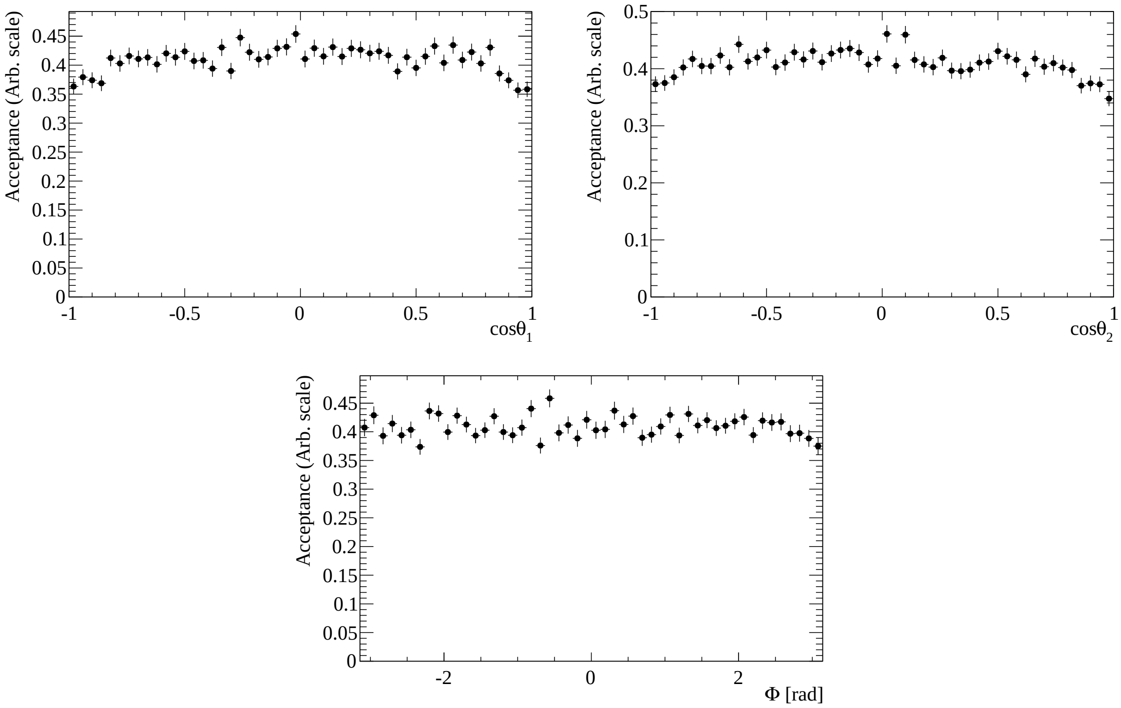


Figure 5.2.: Angular acceptance for the $\cos\theta_1$ (top-left), $\cos\theta_2$ (top-right), and Φ (bottom) helicity angles obtained from fully selected simulated events.

5.2.4. Results

The results of the simultaneous data fit to the PDF in equation 5.11 are shown in Table 5.3, where the signal fraction of the TOS sample, f_s^{TOS}, was allowed to differ from that of the non-TOS sample, $f_s^{\mathrm{non-TOS}}$.

The distribution of the data in each of the helicity angles along with the associated fit result is shown in Figure 5.3. It can be seen from Figure 5.3 that a good quality fit is obtained in the projections onto each helicity angle. A goodness-of-fit test based on the point-to-point dissimilarity method [61], yielded a p-value of 0.45, indicating that good agreement is seen between the data and the fit result.

5.2.5. Systematic Uncertainties

A number of uncertainties need to be accounted for that affect the accuracy of the measurement. These include the uncertainties on the angular acceptance arising from the limited quantity of simulated events with which it is determined, the effect of ignoring

Parameter	Measurement
$\lvert A_0\rvert^2$	0.365±0.022
$\lvert A_\perp\rvert^2$	0.291±0.024
$\lvert A_\parallel\rvert^2 = 1-(\lvert A_0\rvert^2+\lvert A_\perp\rvert^2)$	0.344±0.024
$\cos(\delta_\parallel)$	−0.844±0.068
$m_{B_s^0}$ (MeV/c^2)	5365.3±0.6
$\sigma_{B_s^0}$ (MeV/c^2)	14.2±0.6
α_{BKG} ([MeV/c^2]$^{-1}$)	0.0023±0.0005
f_s^{TOS}	0.953±0.012
$f_s^{\mathrm{non-TOS}}$	0.930±0.012

Table 5.3.: Measured polarization amplitudes and strong phase difference. The uncertainties are statistical only. The sum of the squared amplitudes is constrained to unity. The correlation coefficient between $|A_0|^2$ and $|A_\perp|^2$ is −0.47.

the low efficiency at small B_s^0 decay times that arises from requirements on the impact parameter of the kaon tracks both in the trigger and in the subsequent selections. In addition, the effect of pollution from S-wave, explained in detail in the next section, is non-negligible.

5.2.5.1. S-wave Pollution

The B_s^0 decay into the $K^+K^-K^+K^-$ final state can proceed via combinations of intermediate vector (ϕ) and scalar ($f_0(980)$) resonances and scalar non-resonant K^+K^- pairs. Thus the total decay amplitude is a coherent sum of P-wave (vector-vector), S-wave (vector-scalar) and SS-wave (scalar-scalar) contributions. This extends the differential decay rate seen in equation 5.1 from a sum of 6 terms to a sum of 15 terms, with an additional two amplitudes, $A_{S(SS)}$, and two additional CP-conserving strong phases, $\delta_{S(SS)}$. The derivation and form of these additional terms is explained in detail in Appendix A. Under the same assumptions used to arrive at the time-integrated form in equations 5.5-5.10, the inclusion of S-wave requires the extension of equation 5.4 through

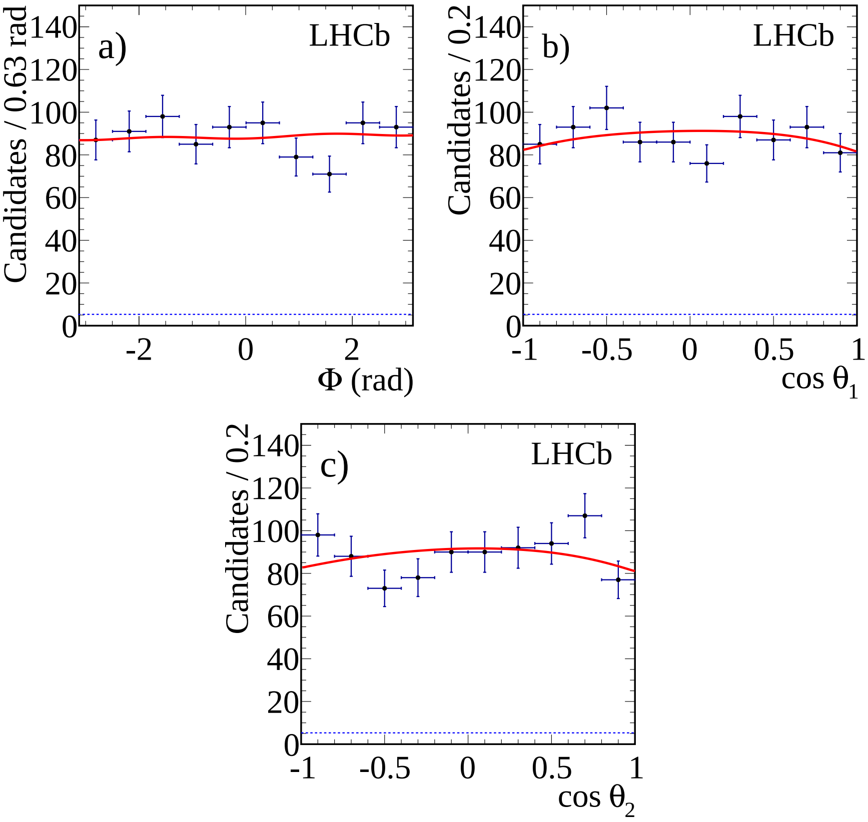


Figure 5.3.: Angular distributions for (a) Φ, (b) $\cos\theta_1$ and (c) $\cos\theta_2$ of $B_s^0 \to \phi\phi$ events with the fit projections for signal and background superimposed for the total fitted PDF (solid line) and background component (dotted line).

$F(\theta_1, \theta_2, \Phi) \to F(\theta_1, \theta_2, \Phi) + F_{S-\text{wave}}(\theta_1, \theta_2, \Phi)$, where

$$\begin{aligned} F_{S-\text{wave}}(\theta_1, \theta_2, \Phi) = {} & \frac{4}{9}|A_{SS}|^2 \tau_L + \frac{4}{3}|A_S|^2 \tau_H (\cos\theta_1 + \cos\theta_2)^2 \\ & + \frac{8}{3}|A_0||A_{SS}|\tau_L \cos(\delta_{SS}) \cos\theta_1 \cos\theta_2 \\ & + \frac{4\sqrt{2}}{3}|A_\parallel||A_{SS}|\tau_L \cos(\delta_2 - \delta_1 - \delta_{SS}) \sin\theta_1 \sin\theta_2 \cos\Phi \\ & + \frac{-4\sqrt{2}}{\sqrt{3}}|A_\perp||A_S|\tau_L \sin(\delta_2 - \delta_S) \sin\theta_1 \sin\theta_2 (\cos\theta_1 + \cos\theta_2) \sin\Phi. \end{aligned} \tag{5.14}$$

A fit to the m_{KK} invariant mass indicated an S-wave component of less than 1 % with uncertainties of order 1 %. It is for this reason that a 2 % contribution was considered as a systematic uncertainty. The uncertainty due to the pollution of S-wave has been calculated through the use of simplified simulations. In the generation of simplified simulations, it has been assumed that $\delta_S = \pi/2$ and $\delta_{SS} = \pi$. It has also been assumed that 1.7 % of the total 2 % S-wave arises from the A_S contribution and 0.3 % arises from the A_{SS} contribution. The difference in fitted values found from simulated datasets due to ignoring various levels of S-wave is shown in Table 5.4 along with the values of the physics parameters used for the generation of simplified simulations.

Parameter	Gen. value	Change with 2 % S-wave	Change with 5 % S-wave
$\lvert A_0\rvert^2$	0.357	0.007	0.019
$\lvert A_\perp\rvert^2$	0.299	0.005	0.016
$\lvert A_\parallel\rvert^2$	0.344	0.012	0.035
$cos\delta_\parallel$	-0.866	0.001	0.003

Table 5.4.: Difference in fitted values found from simulated datasets due to ignoring various levels of S-wave.

5.2.5.2. Angular Acceptance Uncertainty

The basis of the systematic error due to the acceptance correction was obtained by coherently varying the weights by ±1σ using the weights covariance matrix in Table 5.5. In order to vary the weights, a matrix (U) is found to diagonalise the covariance matrix (C), i.e. satisfying the equation

	f_1	f_2	f_3	f_4	f_5	f_6
f_1	8.71714e-05	-4.55602e-05	-5.79748e-05	-1.0125e-06	5.78018e-05	1.69601e-06
f_2	-4.55602e-05	0.000188967	-1.70504e-05	4.92157e-06	8.21257e-05	8.42591e-07
f_3	-5.79748e-05	-1.70504e-05	0.000138672	6.10737e-07	2.03456e-05	9.24192e-07
f_4	-1.0125e-06	4.92157e-06	6.10737e-07	0.00035188	3.57457e-06	4.08374e-05
f_5	5.78018e-05	8.21257e-05	2.03456e-05	3.57457e-06	0.000210417	4.97971e-06
f_6	1.69601e-06	8.42591e-07	9.24192e-07	4.08374e-05	4.97971e-06	0.00015444

Table 5.5.: Covariance matrix of the normalisation weights

$$C = U \cdot D \cdot U^T, \tag{5.15}$$

where D is the diagonalised matrix. The square roots of the eigenvalues of the diagonalised matrix are taken to be the uncorrelated uncertainty. A random number is then generated for each weight with a Gaussian constraint centred on zero with a width of the uncorrelated uncertainty. This variation is then converted back to account for correlations through the equation

$$\Delta\vec{M} = U\Delta\vec{D} \tag{5.16}$$

where $\Delta\vec{M}$ is the correlated variation and $\Delta\vec{D}$ is the uncorrelated variation. The correlated variation is then added to the central values of the normalisation weights and a fit is made to 10000 simulated events. This is repeated 1000 times and the width of a Gaussian fit to the distribution of the bias on each parameter is taken to be the systematic uncertainty for that parameter. Figure 5.4 shows the results of the aforementioned study. The systematic uncertainties taken are summarised in Table 5.6.

Parameter	Width
$\|A_0\|^2$	0.007
$\|A_\perp\|^2$	0.006
$\|A_\parallel\|^2$	0.006
$cos\delta_\parallel$	0.028

Table 5.6.: Systematic uncertainties due to the variation of acceptance weights with statistical errors.

5.2.5.3. Effect of Time Acceptance

The B_s^0 decay time acceptance may be parametrised by the functional form

$$\varepsilon(t) = \frac{a(1-ct)}{1+(dt)^{-b}}, \tag{5.17}$$

where a, b, c, d are parameters to be fitted. The decay time acceptance function along with a corresponding fit may be found in Figure 5.5, where the acceptance has been determined from fully selected, simulated events. The decrease in efficiency at larger decay times is known to be due to the reduced reconstruction efficiency of such events in the LHCb VELO.

The fit results to the parametrisation shown in equation 5.17 are displayed in Table 5.7. This fit function was then included in a decay time-dependent PDF. Toy datasets are

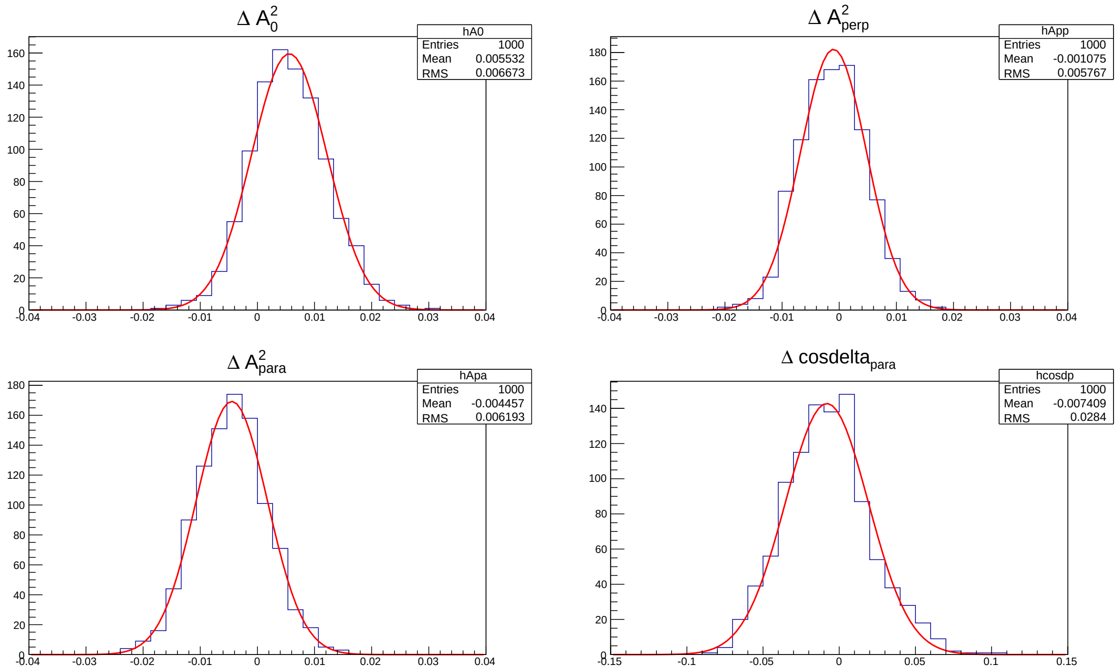


Figure 5.4.: Variation of fitted parameters when normalisation weights are varied according to a Gaussian distribution of 1σ around each weight as described in the text.

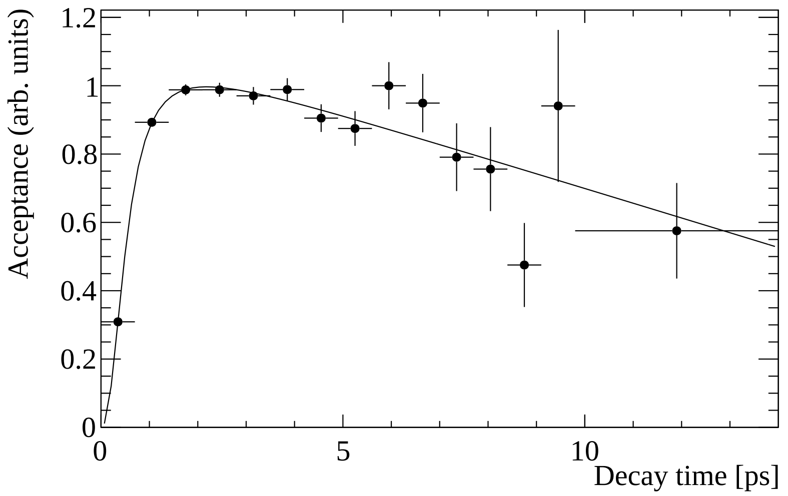


Figure 5.5.: Decay time acceptance for fully selected simulated $B_s^0 \rightarrow \phi\phi$ events, using the cut-based selection method.

generated with and without the decay time acceptance. These were then fitted in the same time-independent method as the data itself. The largest difference between the toy results with and without time acceptance was then taken to be the systematic error.

Parameter	Value
a	10.85 ± 0.36
b	2.26 ± 0.14
c	0.038 ± 0.006
d	1.86 ± 0.08

Table 5.7.: Fit result for the parametrisation of the decay time acceptance used for the generation of systematic uncertainties for the time-integrated analysis.

$\phi^{ss\bar{s}}$ Gen. Value	$\Delta\lvert A_0\rvert^2$	$\Delta\lvert A_\perp\rvert^2$	$\Delta\lvert A_\parallel\rvert^2$	$\Delta\cos\delta_\parallel$
0.0 rad	-0.004	0.006	-0.002	0.007
0.5 rad	-0.006	0.006	0.000	0.007
-0.5 rad	-0.001	0.003	-0.002	-0.004
Syst. uncertainty	0.006	0.006	0.002	0.007

Table 5.8.: Systematic uncertainties as a result of time acceptance for polarisation amplitudes and strong phases.

5.2.5.4. Additional Uncertainties

The angular parametrisation of the combinatoric background used in the calculation of the central result has been assumed to be flat. The justification of this is due to the relatively small number of background candidates present in the data sample. The distributions of the helicity angles from candidates in the B_s^0 sidebands, defined as having a four-kaon invariant mass greater than 60 MeV/c^2 from the PDG B_s^0 mass, is shown in Figure 5.6. It can be seen for the $\cos\theta_{1,2}$ angles that the flat background does not describe the data perfectly, therefore a systematic uncertainty has been derived from the difference between the nominal fit and the use of a $3 \times 3 \times 3$ bin histogram to describe the background. These differences are given in Table 5.9.

As has been mentioned in Section 5.1, extra care is needed due to the choice of trigger requirement. It has been found that different fractions of TOS events are seen in the data sample (42 %) compared to the simulated sample (56 %). To account for this, the difference between the nominal fit (using separate acceptance weights for TOS and non-TOS events) and the fit using one set of acceptance weights for all candidates has been taken as a systematic uncertainty. This uncertainty is shown in Table 5.9.

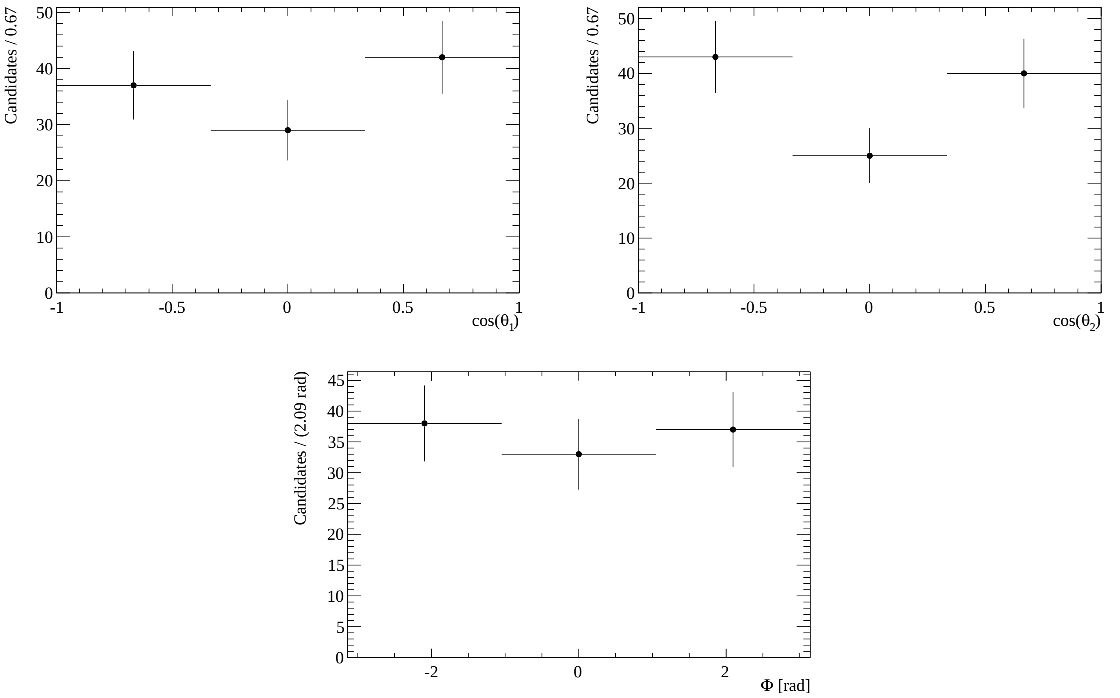


Figure 5.6.: Distribution in each helicity angle of candidates in the B_s^0 sidebands, defined as having a four-kaon invariant mass greater than 60 MeV/c^2 from the PDG B_s^0 mass.

Source	$\sigma^{\text{syst.}}(\lvert A_0\rvert^2)$	$\sigma^{\text{syst.}}(\lvert A_\perp\rvert^2)$	$\sigma^{\text{syst.}}(\lvert A_\parallel\rvert^2)$	$\sigma^{\text{syst.}}(\cos\delta_\parallel)$
Simulation TOS fraction	0.003	0.002	0.001	0.004
Background model	0.001	-	0.001	0.003

Table 5.9.: Systematic errors on the measured polarisation amplitudes and strong phases arising from the background model and treatment of the trigger.

5.3. Triple Product Asymmetries

5.3.1. T-odd Observables

As discussed in Section 1.4.2, triple product asymmetries exploit the angular distributions observed in $P \to VV$ decays to isolate interference terms between CP-odd and CP-even amplitudes. The detailed explanation of the correspondence of T-violating triple product asymmetries and CP-violating phases may be found in Appendix B.

A T-odd asymmetry, A_U, is defined as the difference in the number of decays involving positive and negative values of $\sin 2\Phi$:

$$A_U \equiv \frac{\Gamma(\sin 2\Phi > 0) - \Gamma(\sin 2\Phi < 0)}{\Gamma(\sin 2\Phi > 0) + \Gamma(\sin 2\Phi < 0)}. \tag{5.18}$$

Similarly A_V is defined as

$$A_V \equiv \frac{\Gamma(\mathrm{sign}(\cos\theta_1 \cos\theta_2)\sin\Phi > 0) - \Gamma(\mathrm{sign}(\cos\theta_1 \cos\theta_2)\sin\Phi < 0)}{\Gamma(\mathrm{sign}(\cos\theta_1 \cos\theta_2)\sin\Phi > 0) + \Gamma(\mathrm{sign}(\cos\theta_1 \cos\theta_2)\sin\Phi < 0)}. \tag{5.19}$$

The advantages of triple product asymmetries in this form are that it is possible to observe CP violation without the need for the observation of the B_s^0 decay time or the initial flavour of the B_s^0 meson.

5.3.2. Fit Method

In order to measure the triple product asymmetries in $B_s^0 \to \phi\phi$ decays, two independent datasets were created for the case of each triple product, giving four in total. For each triple product, the $U(V) > 0$ dataset and $U(V) < 0$ dataset were fitted simultaneously to the $K^+K^-K^+K^-$ invariant mass to extract the asymmetry. The PDF used for fitting was of the form

$$P(m) = \sum_{i\in\{+,-\}} f_i^S G(m) + f_i^B \tilde{G}(m), \tag{5.20}$$

where

$$f_+^{S(B)} = \frac{1}{2}(A_{U(V)}^{S(B)} + 1), \tag{5.21}$$

$$f_-^{S(B)} = \frac{1}{2}(1 - A_{U(V)}^{S(B)}). \tag{5.22}$$

A superscript $S(B)$ indicates a signal (background) component and N is the total number of events in the dataset. The functions $G(m)$ and $\tilde{G}(m)$ represent the same double Gaussian and exponential functions that were used in Section 5.2.2. The parameters found in the functions forming the PDF in equation 5.20 are listed in Table 5.10, along with how each parameter was treated in the fit.

Function	Parameter	Type	Value
$f_{\pm}^{S(B)}$	$A_{U(V)}^{S(B)}$	Free	
	$N^{S(B)}$	Free	
$G(m)$	$m_{B_s^0}$	Free	
	$\sigma_{B_s^0}$	Free	
	$\sigma_{B_s^0,2}$	Fixed	$29.5\,\mathrm{MeV}/c^2$
	f_2	Fixed	0.215
$\tilde{G}(m)$	α_{BKG}	Free	

Table 5.10.: Summary of the parameters contained in the functions of equation 5.20 with associated fit treatment for the fit to obtain the triple product asymmetries.

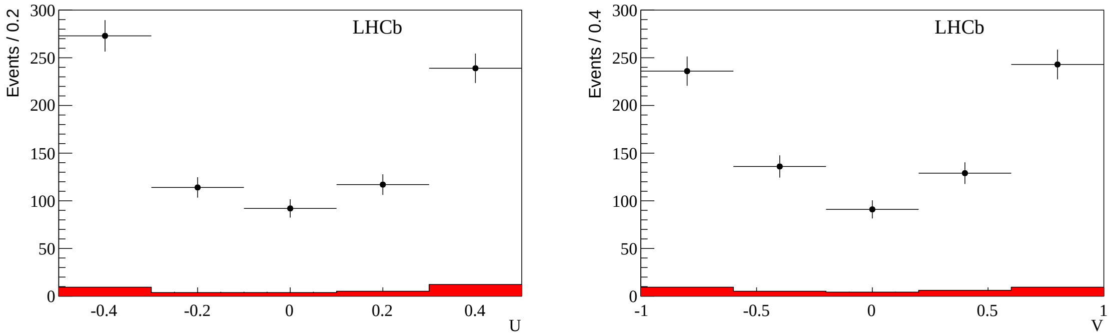


Figure 5.7.: Distributions of the U and V observables for the $B_s^0 \to \phi\phi$ data in the mass range $5286.6 < M(B_s^0) < 5446.6\ MeV/c^2$. The distribution for the background taken from the mass sidebands and normalized to the same mass range is shown in red.

5.3.3. Results

The measured distributions of the U and V observables for $B_s^0 \to \phi\phi$ decays are shown in Figure 5.7 for the mass range $5286.6 < M(B_s^0) < 5446.6 MeV/c^2$.

The fits to the invariant mass of the individual datasets used to extract the triple product asymmetries are shown in Figure 5.8. The triple product asymmetries were found from the simultaneous fits measured found to be

$$A_U = -0.055 \pm 0.036, \tag{5.23}$$

$$A_V = 0.010 \pm 0.036. \tag{5.24}$$

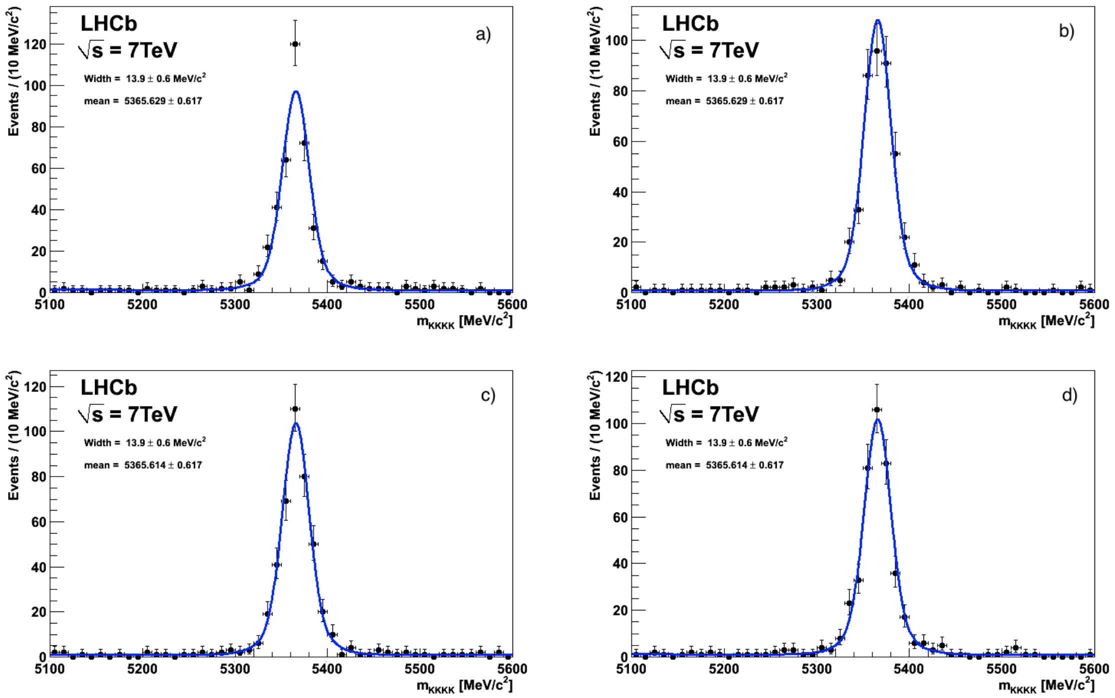


Figure 5.8.: Distributions of the $K^+K^-K^+K^-$ invariant mass for the regions a) $U > 0$, b) $U < 0$, c) $V > 0$, d) $V < 0$ for the $B_s \to \phi\phi$ decay mode. The fit is overlaid.

5.3.4. Systematic Uncertainties

The main sources of systematic uncertainty for the case of the triple product asymmetries are the decay time and angular acceptances, as found for the measurements of the polarisation amplitudes. The effect of the mass model is also known to be significant.

5.3.4.1. Angular Acceptance Uncertainty

In order to quantify the effect on the triple product uncertainties due to the angular acceptance, simplified simulations were generated with and without the effects of angular acceptance. The parameters used to generate the simplified simulations are shown in Table 5.11. A fit was made to the U and V distributions, with 1000 datasets being generated from each fit, giving 4000 datasets in total. The difference between the A_U and A_V parameters with and without the effect of angular acceptance was calculated for the case of A_U and A_V separately and the mean of a Gaussian fit to the distributions was determined. This procedure was repeated for three different values of $\phi^{ss\bar{s}}$ as no measurement of this parameter had been performed at the time. The biases are given for

Parameter	Value
$\lvert A_0\rvert^2$	0.346
$\lvert A_\perp\rvert^2$	0.320
$\lvert A_\parallel\rvert^2$	0.334
δ_1 (rad)	0.01
δ_2 (rad)	2.63
Δm_s (ps^{-1})	17.8
$\phi^{s\bar{s}s}$ (rad)	{0,1.57,-1.57}

Table 5.11.: Physics parameters used in the generation of simplified simulations used to determine the systematic uncertainty in triple product asymmetry measurements.

the three different values of $\phi^{ss\bar{s}}$ in Table 5.12. The largest average value of the bias was then taken as the systematic uncertainty for both asymmetries.

$\phi^{ss\bar{s}}$ value (rad)	ΔA_U	ΔA_V
0.0	0.001	0.000
1.57	0.007	0.010
-1.57	-0.018	-0.008
Average	0.009	0.006

Table 5.12.: Biases on triple products due to angular acceptance corrections. The Average magnitude of the biases is also provided.

5.3.4.2. Effect of Time Acceptance

The same simplified simulation datasets created with and without the effect of the efficiency as a function of B_s^0 decay time that were used to calculate the systematic uncertainties in Section 5.2.5.3 were also used to calculate the systematic uncertainty on the triple product asymmetries. The difference between fitted results with and without the effects of the efficiency as a function of B_s^0 decay time is shown in Table 5.13, where three different generation values of $\phi^{ss\bar{s}}$ have been compared. The largest value was taken as the systematic uncertainty.

$\phi^{ss\bar{s}}$ Gen. Value	ΔA_U	ΔA_V
0.0 rad	-0.006	0.003
0.5 rad	0.003	-0.014
-0.5 rad	0.002	-0.002
Syst. uncertainty	0.006	0.014

Table 5.13.: Systematic uncertainties as a result of time acceptance for triple product asymmetries.

5.3.4.3. Mass Model and Calibration Uncertainty

The triple product asymmetries are calculated from mass fits. Therefore it is important to assess the effects of using different models to parametrise the four-kaon invariant mass. The alternative model used was a single Gaussian to describe the $B_s^0 \rightarrow \phi\phi$ signal and a flat function to describe the combinatoric background. Using this alternative function shifted the A_U and A_V parameters by 0.003 and 0.005, respectively.

The momenta of the kaon tracks used to calculate the four-kaon invariant mass are uncalibrated if used directly from the reconstruction. The reason calibration is necessary is mainly due to the uncertainty in the magnetic field strength. If the magnetic field is wrong by a factor, $(1+\alpha)$, the momenta of the final state tracks must be scaled by a factor of $(1+\alpha)$. Under the assumptions that the final state tracks originate from a decay that is far above threshold and $\alpha << 1$, the correction shifts the mass of the resonance according to [64]

$$\Delta m_{KKKK} = \alpha \cdot m_{KKKK}^{\text{true}} \tag{5.25}$$

$$\Rightarrow m_{KKKK} = (1+\alpha) m_{KKKK}^{\text{true}}, \tag{5.26}$$

where m_{KKKK} represents the shifted reconstructed mass and m_{KKKK}^{true} the correct reconstructed mass. The α factor was found on a run-by-run basis from $J/\psi \rightarrow \mu^+\mu^-$ decays. The mass calibration was found to shift the A_U parameter by 0.003, which is negligible compared with other systematic contributions, but was found to have no effect on the A_V parameter. Therefore no systematic uncertainty was assigned.

5.4. Final Results and Summary

The most accurate measurements of the polarisation amplitudes, strong phase difference, and triple product asymmetries in the $B_s^0 \to \phi\phi$ decay have been presented based on $1.0\,\mathrm{fb}^{-1}$ of LHCb data collected during 2011. These have been measured to be

$$\begin{aligned}
|A_0|^2 &= 0.365 \pm 0.022\,(\mathrm{stat}) \pm 0.012\,(\mathrm{syst})\,,\\
|A_\perp|^2 &= 0.291 \pm 0.024\,(\mathrm{stat}) \pm 0.010\,(\mathrm{syst})\,,\\
\cos\delta_\parallel &= -0.844 \pm 0.068\,(\mathrm{stat}) \pm 0.029\,(\mathrm{syst})\,,\\
A_U &= -0.055 \pm 0.036\,(\mathrm{stat}) \pm 0.018\,(\mathrm{syst})\,,\\
A_V &= 0.010 \pm 0.036\,(\mathrm{stat}) \pm 0.018\,(\mathrm{syst})\,.
\end{aligned}$$

The systematic uncertainties for the polarisation amplitudes and strong phase difference are summarised in Table 5.14. The corresponding uncertainties on the triple product asymmetries are summarised in Table 5.15. The dominant contributions to the systematic uncertainties consistently arise from the B_s^0 decay-time acceptance and angular acceptance.

Source	$\lvert A_0\rvert^2$	$\lvert A_\perp\rvert^2$	$\lvert A_\parallel\rvert^2$	$\cos\delta_\parallel$
S-wave component	0.007	0.005	0.012	0.001
Decay time acceptance	0.006	0.006	0.002	0.007
Angular acceptance	0.007	0.006	0.006	0.028
Trigger category	0.003	0.002	0.001	0.004
Background model	0.001	–	0.001	0.003
Total	0.012	0.010	0.014	0.029

Table 5.14.: Summary of the systematic uncertainties on the measured polarisation amplitudes and the strong phase difference.

The measured polarisation amplitudes are in agreement with the CDF measurement and the predictions of QCD factorisation shown in Section 1.4.2.2. The T-violating triple product asymmetries are found to be consistent with CP conservation but with large uncertainties. All parameters in Table 5.14 have been measured with a statistical uncertainty of around half the values reported in the previous CDF measurement [27].

Source	A_U	A_V	Final uncertainty
Angular acceptance	0.009	0.006	0.009
Decay time acceptance	0.006	0.014	0.014
Fit model	0.004	0.005	0.005
Total			0.018

Table 5.15.: Summary of the systematic uncertainties on the triple product asymmetries A_U and A_V. The total uncertainty is the quadratic sum of the larger of the two components.

Chapter 6.

First Measurement of the CP-Violating Phase in $B_s^0 \to \phi\phi$ Decays

"Sed fugit interea, fugit inreparabile tempus."
— Virgil

This Chapter explains in detail the decay time dependent measurement of the CP-violating phase in $B_s^0 \to \phi\phi$ decays. The polarisation fractions and strong phase differences were also measured. The measurement was published in Physical Review Letters during 2013 [65].

6.1. Dataset

The measurements were performed using approximately $1\,\text{fb}^{-1}$ of LHCb data collected during 2011[1]. The analysis used the multi-variate selection, defined in Table 4.4. The nominal trigger requirements were imposed. This consisted of the signal candidate causing the event to be triggered (TOS) for the hadron trigger at the L0 stage or the event passing independently of the signal candidate (TIS) on any L0 trigger line. The TOS requirement was imposed on the HLT1 trigger selecting all tracks from L0. At the HLT2 stage, the TOS requirement was imposed on the topological BDT triggers or the inclusive ϕ trigger.

---

[1] Samples were produced with version 29r2 of the DaVinci application and made use of version 12 of the reconstruction.

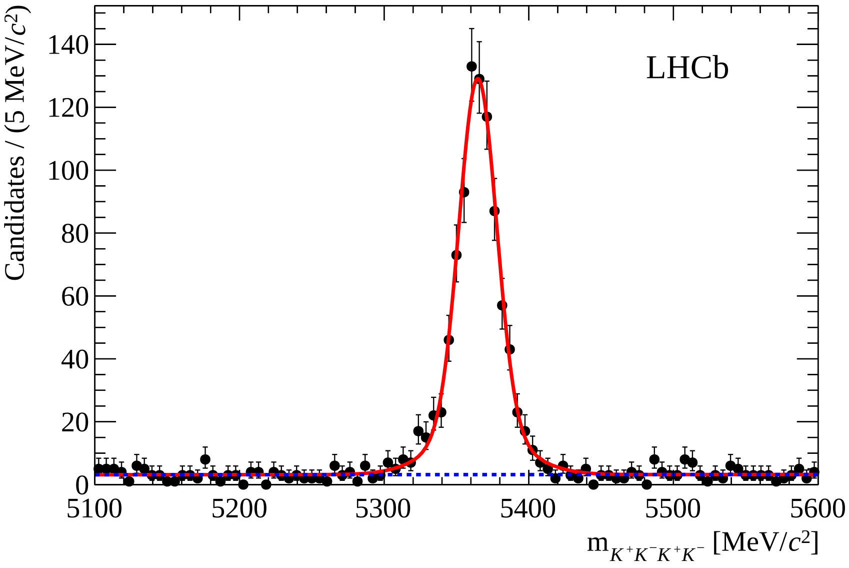


Figure 6.1.: Invariant $K^+K^-K^+K^-$ mass distribution for selected $B_s^0 \to \phi\phi$ candidates. A fit of a double Gaussian signal component together with an exponential background (dotted line) is superimposed.

Simulated events used for the purposes of acceptance corrections were created under the same conditions as the data itself. The parameters used to simulate the distributions of helicity angles and decay time are shown in Table 4.3.

To determine the signal yield an unbinned maximum likelihood fit was performed with the same model as used in Section 5. The width of the first Gaussian is measured from data to be 12.9 ± 0.5 MeV/c^2. The relative fraction, f_2, and width, $\sigma_{B_s^0,2}$, of the second Gaussian were fixed to 0.215 and 29.5 MeV/c^2 respectively, where values were obtained from simulation. The background was described by an exponential function, with decay constant, $\alpha_{\rm BKG}$. Figure 6.1 shows the four-kaon invariant mass distribution for selected events. A signal yield of 880 ± 31 events was found when fitting the PDF described above to the data. A goodness-of-fit test based on the point-to-point dissimilarity method [61], yielded a p-value of 0.39, indicating that the fit result is in good agreement with the data distribution.

6.2. Angular and Time-Dependent Decay Rates

In contrast to the measurement presented in Section 5, where the presence of an S-wave component was treated as a systematic uncertainty, the decay time dependent measurement was performed accounting for such a contribution directly in the angular fit. As has been discussed in Section 5.2.5.1, accounting for the full S-wave contribution in the angular fit extends equation 5.1 to

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_1 \mathrm{d}\cos\theta_2 \mathrm{d}\Phi \mathrm{d}t} \propto F(t,\theta_1,\theta_2,\Phi) = \sum_{i=1}^{15} K_i(t) f_i(\theta_1,\theta_2,\Phi), \tag{6.1}$$

where the time-dependent functions, $K_i(t)$, can be written as [62]

$$K_i(t) = N_i e^{-\Gamma_s t}[c_i \cos(\Delta m_s t) + d_i \sin(\Delta m_s t) + a_i \cosh(\tfrac{1}{2}\Delta\Gamma_s t) + b_i \sinh(\tfrac{1}{2}\Delta\Gamma_s t)]. \tag{6.2}$$

The angular ($f_i(\theta_1,\theta_2,\Phi)$) and decay time-dependent functions are given for the 15 terms in Table 6.1, where the parameters have been defined in Section 5. The full derivation of these functions is given in Appendix A.

6.2.1. S-wave Coupling

In previous discussions on the subject of the S-wave, the difference between the S-wave and P-wave line-shapes in terms of the invariant mass has been ignored. Clearly, if two resonances are far apart, then the interference between them must be small as a consequence.

The polarisation amplitudes and strong phases contain a dependence on the $\mathrm{K}^+\mathrm{K}^-$ invariant mass, $\mu = m_{\mathrm{K}^+\mathrm{K}^-}$. All P-wave amplitudes contain the same dependence on μ. The effect of using a finite $m_{\mathrm{K}^+\mathrm{K}^-}$ range, $[\mu_l, \mu_h]$, can be incorporated through making the substitutions [66]

$$|A_i(\mu)|^2 \to \int_{\mu_l}^{\mu_h} |A_i(\mu)|^2 \mathrm{d}\mu \qquad \text{for} \;\; i \in \{\parallel, \perp, 0, S, SS\}, \tag{6.3}$$

$$A_i^*(\mu)A_j(\mu) \to \int_{\mu_l}^{\mu_h} A_i^*(\mu)A_j(\mu)\mathrm{d}\mu \qquad \text{for} \;\; i \neq j. \tag{6.4}$$

The μ-dependence in amplitudes can be parametrised by two normalised complex functions, $g(\mu)$ and $h(\mu)$, for the P-wave and S-wave amplitudes, respectively, such that

$$A_i(\mu) = a_i g(\mu) \qquad \text{for} \;\; i \in \{\parallel, \perp, 0\}, \tag{6.5}$$

$$A_j(\mu) = a_j h(\mu) \qquad \text{for} \;\; j \in \{S, SS\}, \tag{6.6}$$

i	N_i	a_i	b_i	c_i	d_i	f_i
1	$\|A_0\|^2$	1	$-\cos\phi^{ss\bar{s}}$	0	$\sin\phi^{ss\bar{s}}$	$4\cos^2\theta_1\cos^2\theta_2$
2	$\|A_\parallel\|^2$	1	$-\cos\phi^{ss\bar{s}}$	0	$\sin\phi^{ss\bar{s}}$	$\sin^2\theta_1\sin^2\theta_2(1+\cos 2\Phi)$
3	$\|A_\perp\|^2$	1	$\cos\phi^{ss\bar{s}}$	0	$-\sin\phi^{ss\bar{s}}$	$\sin^2\theta_1\sin^2\theta_2(1-\cos 2\Phi)$
4	$\|A_\parallel\|\|A_\perp\|$	0	$-\cos\delta_1\sin\phi^{ss\bar{s}}$	$\sin\delta_1$	$-\cos\delta_1\cos\phi^{ss\bar{s}}$	$-2\sin^2\theta_1\sin^2\theta_2\sin 2\Phi$
5	$\|A_\parallel\|\|A_0\|$	$\cos(\delta_{2,1})$	$-\cos(\delta_{2,1})\cos\phi^{ss\bar{s}}$	0	$\cos(\delta_{2,1})\sin\phi^{ss\bar{s}}$	$\sqrt{2}\sin 2\theta_1\sin 2\theta_2\cos\Phi$
6	$\|A_0\|\|A_\perp\|$	0	$-\cos\delta_2\sin\phi^{ss\bar{s}}$	$\sin\delta_2$	$-\cos\delta_2\cos\phi^{ss\bar{s}}$	$-\sqrt{2}\sin 2\theta_1\sin 2\theta_2\sin\Phi$
7	$\|A_{SS}\|^2$	1	$-\cos\phi^{ss\bar{s}}$	0	$\sin\phi^{ss\bar{s}}$	$\frac{4}{9}$
8	$\|A_S\|^2$	1	$\cos\phi^{ss\bar{s}}$	0	$-\sin\phi^{ss\bar{s}}$	$\frac{4}{3}(\cos\theta_1+\cos\theta_2)^2$
9	$\|A_S\|\|A_{SS}\|$	0	$-\sin(\delta_S-\delta_{SS})\sin\phi^{ss\bar{s}}$	$\cos(\delta_{SS}-\delta_S)$	$-\sin(\delta_{SS}-\delta_S)\cos\phi^{ss\bar{s}}$	$\frac{8}{3\sqrt{3}}(\cos\theta_1+\cos\theta_2)$
10	$\|A_0\|\|A_{SS}\|$	$\cos\delta_{SS}$	$-\cos\delta_{SS}\cos\phi^{ss\bar{s}}$	0	$\cos\delta_{SS}\sin\phi^{ss\bar{s}}$	$\frac{8}{3}\cos\theta_1\cos\theta_2$
11	$\|A_\parallel\|\|A_{SS}\|$	$\cos(\delta_{2,1}-\delta_{SS})$	$-\cos(\delta_{2,1}-\delta_{SS})\cos\phi^{ss\bar{s}}$	0	$\cos(\delta_{2,1}-\delta_{SS})\sin\phi^{ss\bar{s}}$	$\frac{4\sqrt{2}}{3}\sin\theta_1\sin\theta_2\cos\Phi$
12	$\|A_\perp\|\|A_{SS}\|$	0	$-\cos(\delta_2-\delta_{SS})\sin\phi^{ss\bar{s}}$	$\sin(\delta_2-\delta_{SS})$	$-\cos(\delta_2-\delta_{SS})\cos\phi^{ss\bar{s}}$	$-\frac{4\sqrt{2}}{3}\sin\theta_1\sin\theta_2\sin\Phi$
13	$\|A_0\|\|A_S\|$	0	$\sin\delta_S\sin\phi^{ss\bar{s}}$	$\cos\delta_S$	$\sin\delta_S\cos\phi^{ss\bar{s}}$	$\frac{8}{\sqrt{3}}\cos\theta_1\cos\theta_2$ $\times(\cos\theta_1+\cos\theta_2)$
14	$\|A_\parallel\|\|A_S\|$	0	$-\sin(\delta_{2,1}-\delta_S)\sin\phi^{ss\bar{s}}$	$\cos(\delta_{2,1}-\delta_S)$	$-\sin(\delta_{2,1}-\delta_S)\cos\phi^{ss\bar{s}}$	$\frac{4\sqrt{2}}{\sqrt{3}}\sin\theta_1\sin\theta_2$ $\times(\cos\theta_1+\cos\theta_2)\cos\Phi$
15	$\|A_\perp\|\|A_S\|$	$\sin(\delta_2-\delta_S)$	$\sin(\delta_2-\delta_S)\cos\phi^{ss\bar{s}}$	0	$-\sin(\delta_2-\delta_S)\sin\phi^{ss\bar{s}}$	$-\frac{4\sqrt{2}}{\sqrt{3}}\sin\theta_1\sin\theta_2$ $\times(\cos\theta_1+\cos\theta_2)\sin\Phi$

Table 6.1.: Coefficients of the time-dependent terms and angular functions used in equation 5.1. Amplitudes are defined at $t = 0$.

where a_k denote real coefficients and

$$\int_{\mu_l}^{\mu_h} |g(\mu)|^2 \mathrm{d}\mu := \int_{\mu_l}^{\mu_h} |h(\mu)|^2 \mathrm{d}\mu = 1. \tag{6.7}$$

We can also parametrise the interference between the normalised S-wave and P-wave amplitudes as

$$\int_{\mu_l}^{\mu_h} g^*(\mu) h(\mu) \mathrm{d}\mu \equiv C_{SP} e^{i\theta_{SP}}, \tag{6.8}$$

where C_{SP} is a real number in the range $[0, 1]$ and can be viewed as an effective coupling parameter between the P-wave and S-wave line shapes. The form of $g(\mu)$ is known to be a Breit-Wigner with width and mean denoted by Γ_ϕ and μ_ϕ, respectively. In the calculation of the C_{SP} factor, PDG values of the width and mean of the ϕ resonance were used. The form of the S-wave line-shape is known to have a small μ-dependence close to the ϕ mass. It is for this reason that a flat model is used to describe the S-wave [66]. The explicit functional forms of the line-shapes used were

$$g(\mu) = \sqrt{\frac{\Gamma_\phi/2}{\Delta\lambda}} \cdot \frac{1}{\mu - \mu_\phi + i\Gamma_\phi/2}, \tag{6.9}$$

$$h(\mu) = \sqrt{\frac{1}{\Delta\mu}}, \tag{6.10}$$

where

$$\Delta\lambda = \tan^{-1}\frac{2(\mu_h - \mu_\phi)}{\Gamma_\phi} - \tan^{-1}\frac{2(\mu_l - \mu_\phi)}{\Gamma_\phi}, \tag{6.11}$$

$$\Delta\mu = \mu_h - \mu_l. \tag{6.12}$$

From these line-shapes, it is straightforward to derive

$$C_{SP} \cdot e^{-i\theta_{SP}} = \sqrt{\frac{\Gamma_\phi}{2\Delta\mu\Delta\lambda}} \ln\frac{\mu_h - \mu_\phi + i\Gamma_\phi/2}{\mu_l - \mu_\phi + i\Gamma_\phi/2}. \tag{6.13}$$

From equation 6.13, the real and imaginary components were computed to obtain

$$C_{SP} = \sqrt{\frac{\Gamma_\phi}{2\Delta\mu\Delta\lambda}} \sqrt{(\ln(Z)^2 + \eta^2}, \tag{6.14}$$

$$\theta_{SP} = -\arg(\ln Z + i\eta), \tag{6.15}$$

where

$$Z = \sqrt{\frac{(\mu_h - \mu_\phi)^2 + (\Gamma_\phi/2)^2}{(\mu_l - \mu_\phi)^2 + (\Gamma_\phi/2)^2}}, \tag{6.16}$$

$$\eta = \arg(\mu_h - \mu_\phi + i\Gamma_\phi/2) - \arg(\mu_l - \mu_\phi + i\Gamma_\phi/2). \tag{6.17}$$

Therefore, for a ϕ mass window of $30\,\mathrm{MeV}/c^2$ centred on the nominal ϕ mass, the derivation described in this section led to a C_{SP} parameter of 0.45 that was multiplied by the interference terms between the ϕf_0 S-wave and the other amplitudes. The factor multiplied by the P-wave and $f_0 f_0$ interferences was thus 0.45^2.

An additional factor of 0.45^2 was also needed on part of the diagonal $|A_s(t)|^2$ term. From the derivation of this term, it can be seen that the two cases of ϕf_0 and $f_0\phi$ S-wave should be treated separately. In this case, the term in the PDF is of the form

$$\begin{aligned}
&(A_s(t)g(\mu_1)h(\mu_2)\cos\theta_1 + A_s(t)g(\mu_2)h(\mu_1)\cos\theta_2) \times (A_s(t)g(\mu_1)h(\mu_2)\cos\theta_1 \\
&+ A_s(t)g(\mu_2)h(\mu_1)\cos\theta_2)^* = |A_s(t)g(\mu_1)h(\mu_2)|^2\cos^2\theta_1 + |A_s(t)g(\mu_2)h(\mu_1)|^2\cos^2\theta_2 \\
&+ |A_s(t)|^2[g(\mu_1)h^*(\mu_1)][h(\mu_2)g^*(\mu_2)]\cos\theta_1\cos\theta_2 \\
&+ |A_s(t)|^2[g(\mu_2)h^*(\mu_2)][h(\mu_1)g^*(\mu_1)]\cos\theta_1\cos\theta_2
\end{aligned} \tag{6.18}$$

where μ_i denotes the invariant mass of the i^{th} resonance. It can be seen in the last two terms of equation 6.18 that the masses are allowed to differ. Therefore, it is in these last two terms that a factor of 0.45^2 was required after integration over μ_1 and μ_2.

6.3. Identification of B_s^0 Flavour at Production

It is important to be able to determine the initial flavour of the B_s^0 meson as this allows access a larger amount of terms sensitive to *CP* violation. Figure 6.2 describes the principles behind the different flavour tagging algorithms applicable to the $B_s^0 \to \phi\phi$ decay. These may be separated in to two categories, opposite-side (OS) tagging, described in detail in Section 6.3.1, and same-side kaon tagging, described in detail in Section 6.3.2.

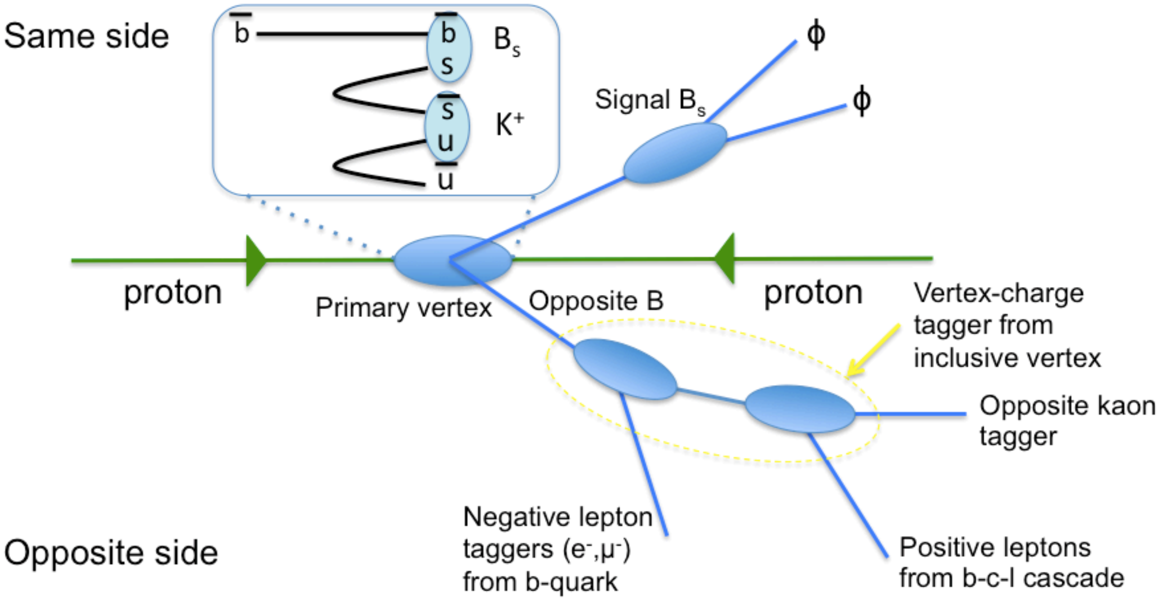


Figure 6.2.: Diagram of the principles behind the flavour tagging algorithms.

6.3.1. Opposite-Side Flavour Tagging

The principle behind opposite-side flavour tagging is to identify the flavour of the hadron that results from the $\bar{b}(b)$-quark produced in association with the signal $b(\bar{b})$-quark. In order to determine the flavour of the opposite-side b-quark, the OS flavour-tagging algorithms use the charge of the lepton in semi-leptonic decays, the charge of the kaon from the $b \to c \to s$ decay chain or the charge of the inclusive secondary vertex reconstructed from the b-hadron decay products [67].

6.3.1.1. Single-Particle OS Tagging Algorithms

The OS tagging algorithms exploiting the charge of a particle in the decay of an opposite-side b-quark identify such particles using the general properties of B decays. These are that in general particles from B decays will have a large impact parameter significance (IP/σ_{IP}) with respect to the primary vertex and large transverse momentum, p_{T}. To reject particles coming from other primary interactions, impact parameter significance selections are also imposed with respect to such pile-up primary vertices $(IP^{PU}/\sigma_{IP}^{PU})$. Particle identification information from the calorimeters, RICH detectors, and muon stations are also used in the form of differences in the likelihoods of the different mass hypotheses, denoted by $\mathrm{DLL}_{K\pi}$, $\mathrm{DLL}_{p\pi}$, $\mathrm{DLL}_{e\pi}$ and $\mathrm{DLL}_{\mu\pi}$. The selection requirements for the OS muon, electron and kaon tagging algorithms are given in Table 6.2. If multiple

Tagger	Min. p_{T} [GeV/c]	Min. p [GeV/c]	Min. (IP/σ_{IP})	PID requirements	Min. $(IP^{PU}/\sigma_{IP}^{PU})$
μ	1.2	2.0	-	$\mathrm{DLL}_{\mu\pi} > 2.5$	3.0
e	1.0	2.0	2.0	$\mathrm{DLL}_{e\pi} > 4.0$	3.0
K	0.8	5.9	4.0	$\mathrm{DLL}_{K\pi} > 6.5$, $\mathrm{DLL}_{Kp} > -3.5$	4.7

Table 6.2.: Selection requirements used to isolate particles used for single-particle OS flavour tagging [67].

candidates are found from the same single-particle flavour-tagging algorithm, then the candidate with the highest transverse momentum is used to infer the flavour of the signal B_s^0 meson.

6.3.1.2. Vertex Charge OS Tagging Algorithm

The initial vertex constructed by the algorithm, combines two tracks with $p_{\mathrm{T}} > 0.15\,\mathrm{GeV}/c$ and $IP/\sigma_{IP} > 2.5$ and assumes a pion mass hypothesis. A good quality vertex fit is imposed and combinations compatible with the K_{S}^0 mass are rejected to reduce the large K_{S}^0 background. For each candidate, the probability of originating from a B decay is estimated from the vertex fit quality and kinematic information. The candidate with the highest probability is used. Tracks that are compatible with originating from the vertex, but not from the primary vertex, are added to form the final inclusive vertex. The inclusive vertex is then required to have total momentum $> 10\,\mathrm{GeV}/c$, total $p_{\mathrm{T}} > 1.5\,\mathrm{GeV}/c$, invariant mass $> 0.5\,\mathrm{GeV}/c^2$, and the sum $IP/\sigma_{IP} > 10.0$. The charge of the inclusive vertex, Q_{IV} is then calculated as

$$Q_{IV} = \frac{\sum_i Q_i {p_{\mathrm{T}}}_i^{\kappa}}{\sum_i {p_{\mathrm{T}}}_i^{\kappa}}, \tag{6.19}$$

where Q_i is the charge of the i^{th} track, ${p_{\mathrm{T}}}_i$ is the transverse momentum of the i^{th} track and $\kappa = 0.4$ optimises the tagging power.

6.3.2. Same-Side Kaon Flavour Tagging

When a signal B_s^0 meson is formed, there is an associated $\bar{s}$-quark that 50% of the time forms a charged kaon, the flavour of which will allow for the identification of the flavour of the signal B_s^0 meson. This principle is exploited by the same-side kaon tagger [68].

In order to select the charged kaon associated with the fragmentation of the B_s^0 meson, a minimum set of criteria is imposed that is the same as for the OS K tagger, defined in Table 6.2. Tracks associated to the signal candidate are excluded. Tagging tracks are required to be outside a conical volume around the B candidate and decay products, defined by a minimum polar angle. As the fragmentation kaon is often located close to the signal B_s^0 meson in phase space, requirements on the maximum difference in pseudo-rapidity and the angle between the tagging track and the B_s^0 momentum are imposed. In addition, a maximum value of the impact parameter significance is imposed with respect to the B_s^0 production vertex, to ensure the tagging particle arises from the fragmentation.

As in the case of single-particle OS taggers, in the event of multiple particles satisfying requirements, the one with the highest p_{T} is used to determine the flavour of the signal B_s^0 meson.

6.3.3. Mistag Probabilities and Combination of Different Flavour-Tagging Algorithms

For all flavour-tagging algorithms, there is a probability for the associated tag to be incorrect. This can be estimated on an event-by-event basis from the properties of the event and of the algorithm itself. The estimation of this so-called mistag rate is performed using a neural network that uses the signal B_s^0 transverse momentum, the number of pile-up vertices, the number of tracks preselected as tagging particles, and the kinematics of the tagging particles[2]. For the case of opposite-side flavour-tagging algorithms, the neural network is trained using simulated $B^+ \to J/\psi \mathrm{K}^+$ events, whereas for the same-side kaon flavour-tagging algorithm, the neural network is trained using simulated $B_s^0 \to D_s^- \pi^+$ events.

---

[2] For the case of the inclusive vertex algorithm, the kinematics of the tracks associated to the vertex are used as inputs to the neural network.

In a large number of events, it is the case that more than one flavour-tagging algorithm provides a decision. The mistag rate must then be re-computed to reflect this. The combined probability that the B_s^0 meson contains a b-quark, $P(b)$, is calculated as

$$P(b) = \frac{p(b)}{p(\bar{b})}, \tag{6.20}$$

where

$$p(b) = \prod_i \left(\frac{1+d_i}{2} - d_i(1-\eta_i) \right), \tag{6.21}$$

$$p(\bar{b}) = \prod_i \left(\frac{1-d_i}{2} + d_i(1-\eta_i) \right). \tag{6.22}$$

In equations 6.21 and 6.22, d_i and η_i represent the decision and mistag probability of the i^{th} flavour tagging algorithm, respectively. The combined decision and mistag rate are

$$d = \begin{cases} -1, & \text{if } P(b) > P(\bar{b}) \\ +1, & \text{otherwise} \end{cases} \tag{6.23}$$

$$\eta = \begin{cases} 1 - P(b), & \text{if } P(b) > P(\bar{b}) \\ 1 - P(\bar{b}). & \text{otherwise} \end{cases} \tag{6.24}$$

The method for combining the information from multiple flavour-tagging algorithms has been used not only to combine the different OS flavour-tagging responses, but also, where necessary to combine the information from the OS and SS kaon algorithms.

6.3.4. Calibration of Flavour-Tagging Algorithms

The mistag probability found as the output of the neural network requires calibration to be accurate with respect to the data. The calibration is performed using decays that are self-tagged, thus giving a data-driven method of measuring the mistag fraction. The relation between the measured mistag probability, ω, and the calculated mistag probability, η, is given by

$$\omega = p_0 + p_1 \cdot (\eta - < \eta_c >), \tag{6.25}$$

where p_0 and p_1 are calibration parameters found from fits to the control channels, and $< \eta_c >$ is the average mistag rate in the data sample being considered.

In the case of the OS flavour-tagging algorithms, the control channel used is $B^+ \to J/\psi \mathrm{K}^+$, as the charge of the kaon allows for the flavour of the B meson to be known and hence the mistag fraction to be measured. The measured mistag rate versus predicted mistag rate in $B^+ \to J/\psi \mathrm{K}^+$ decays is shown in Figure 6.3, along with a corresponding fit to find the calibration parameters.

For the case of the SS kaon flavour-tagging algorithm, the $B_s^0 \to D_s^- \pi^+$ decay mode has been used to determine the calibration parameters. The use of a B_s^0 decay mode to determine the calibration parameters is challenging due to the fast oscillations of B_s^0 mesons. The challenges of such a method are that the decay time resolution must be known accurately in addition to the knowledge of the efficiency as a function of B_s^0 decay time, as the measured mistag rate is evaluated from a fit to the decay time of the form

$$A(t) = (1 - 2\omega) \exp(-\frac{1}{2} \left(\Delta m_s \sigma_t\right)^2) \cos(\Delta m_s t)/ \cosh(\Delta\Gamma_s t/2), \tag{6.26}$$

where σ_t is the decay-time resolution. The advantage of this method is that the calibration parameters may be obtained from an unbinned fit in equation 6.26 in addition to the binned fit method used to determine the calibration parameters of the OS taggers. The measured mistag rate versus predicted mistag rate in $B_s^0 \to D_s^- \pi^+$ decays is shown in Figure 6.3, along with corresponding fits to find the calibration parameters.

The use of OS and SS kaon flavour tagging algorithms has required that four distinct datasets be fitted simultaneously. These correspond to untagged events, events tagged solely by the OS tagging algorithms, events tagged solely by the SSK tagging algorithm and events where both types of tagging algorithm provided a decision. In the case of a combined decision, a combined mistag probability is provided through the method described in Section 6.3.3. This combined mistag probability was calculated using calibrated OS and SSK mistag probabilities.

Systematic uncertainties on the OS tagging algorithm mainly arise due to differences in the calibration parameters obtained from different run periods and also from differences in the calibration parameters depending on the flavour of the signal B meson. Systematic uncertainties on the SS kaon tagging algorithm calibration parameters mainly consist of uncertainties of the decay time resolution and the difference between the results of the binned and unbinned fit to determine the calibration parameters. Table 6.3 shows

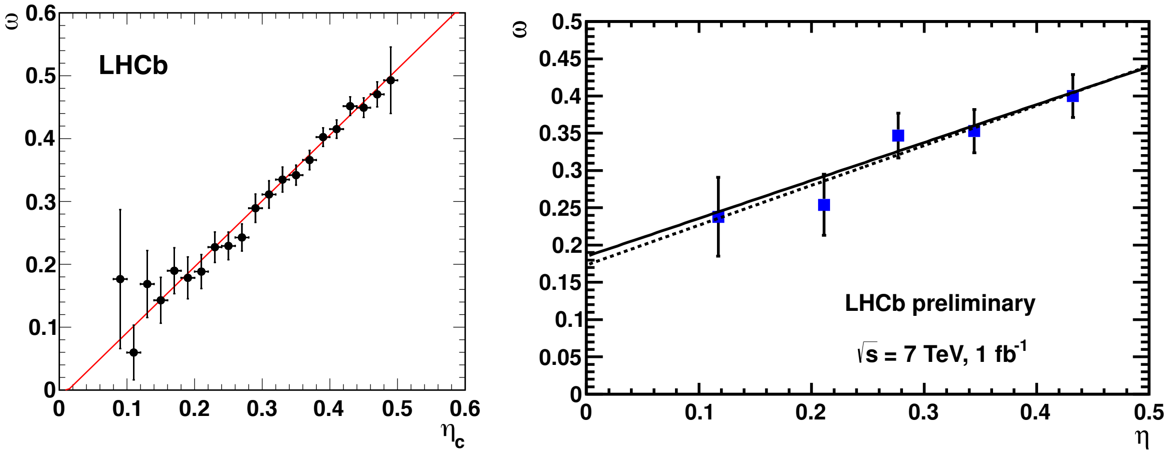


Figure 6.3.: Predicted mistag rate versus measured mistag rate in the $B^+ \to J/\psi \mathrm{K}^+$ decay (left) and $B_s^0 \to D_s^- \pi^+$ (right), used to determine the calibration parameters for the OS and SS flavour tagging algorithms, respectively [67, 68]. The solid lines show fits to the data points using equation 6.25, while the dotted line shows the result of an unbinned fit to determine the calibration parameters.

the calibration parameters applied to the $B_s^0 \to \phi\phi$ data. Uncertainties contain both statistical and systematic errors. The effect on the time-dependent $B_s^0 \to \phi\phi$ analysis from the uncertainty on the calibration parameters is accounted for directly during fitting with the use of Gaussian constraints.

Dataset	p_0	p_1	$< \eta_c >$
OS tagged	0.392 ± 0.008	1.035 ± 0.023	0.391
SSK tagged	0.350 ± 0.017	0.51 ± 0.16	0.324
Both tagged	0.0 ± 0.025	1.0	0.0

Table 6.3.: Summary of the calibration parameters defined in equation 6.25 for the different tagging categories [15].

6.4. Decay Time Resolution

The sensitivity on $\phi^{s\bar{s}s\bar{s}}$ is greatly improved through the precise measurement of the B_s^0 decay time. The difference between the decay time calculated from the reconstructed momentum and the exact decay time of truth-matched events from simulation is shown in Figure 6.4. The result of a double Gaussian fit to the decay time residuals distribution

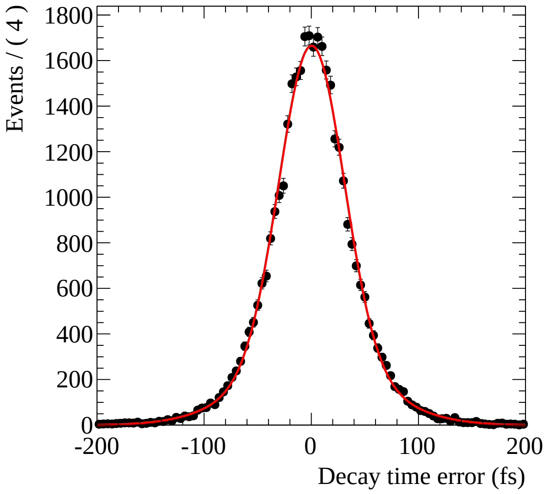


Figure 6.4.: Difference between reconstructed decay time and the exact decay time of truth-matched events from simulation. Also plotted is the result of a double Gaussian fit.

yielded widths of $\sigma_1 = 30.0 \pm 0.4$ fs and $\sigma_2 = 61.7 \pm 1.0$ fs for the first and second Gaussian, respectively. The fraction of the first Gaussian was found to be 0.727 ± 0.015. In fitting for $\phi^{ss\bar{s}}$, the important quantity is the error on ϕ_s. This can be related to the time dilution ($\mathcal{D}_t$) through

$$\sigma(\sin\phi_s) \propto \frac{1}{\mathcal{D}_t}, \tag{6.27}$$

where for a given resolution function $R(\delta_t)$, $\mathcal{D}_t$ is defined through [69]

$$\mathcal{D}_t(\nu) = \frac{2}{\sqrt{\pi}} \int_0^\infty R(\delta_t)\cos(-\nu\delta_t)\mathrm{d}\delta_t, \tag{6.28}$$

where ν is the oscillation frequency of the amplitude and δ_t is the error on the decay time. For a sum of Gaussian resolution functions, $\mathcal{D}_t$ reduces to [70]

$$\mathcal{D}_t = \sum_j f_j \exp(-\Delta m_s^2 \sigma_{j,t}^2/2), \tag{6.29}$$

where f_j is the fraction of the j^{th} Gaussian, Δm_s is the B_s^0 oscillation frequency and $\sigma_{j,t}$ is the width of the j^{th} Gaussian. Therefore, equating dilutions, the equivalent single Gaussian resolution applied during fitting was 39.7 fs.

6.5. Decay Time Acceptance

As has been explained in Section 5, the detector reconstruction, trigger and offline requirements introduce acceptance effects on the proper time of the B_s^0 meson. The acceptance correction was then calculated by taking the ratio of fully selected, simulated events with those from generator level. The B_s^0 decay time acceptance was accounted for directly in fitting through the direct use of the histogram shown in Figure 6.5. It can be seen in Figure 6.5 that systematic fluctuations appear at around 2.5 ps and 3.5 ps. In order to account for this, a fit was made to the time acceptance of the form

$$\varepsilon(t) = \frac{a(1-ct)}{e+(dt)^{-b}}, \tag{6.30}$$

where a, b, c, d and e are parameters to be fitted. The parameters that go in to the fit are displayed in Table 6.4. The fit is also displayed in Figure 6.5. The difference between the fitted acceptance and the histogram was taken as the systematic uncertainty due to the time acceptance.

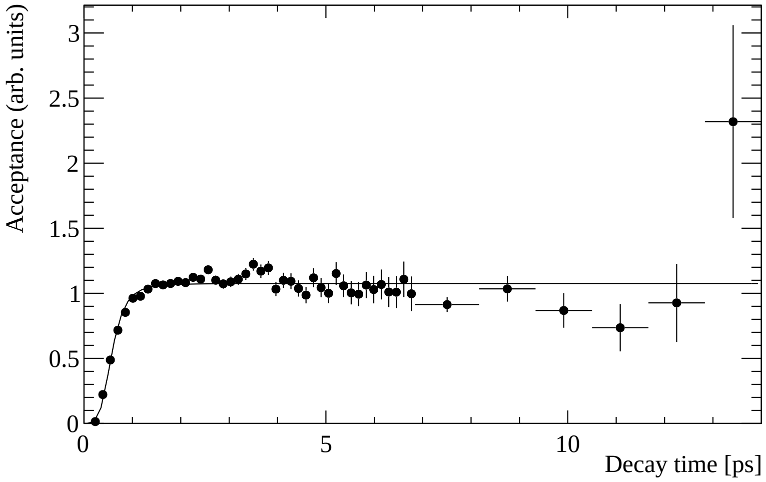


Figure 6.5.: Time acceptance including effects from the detector, trigger and selection for simulated $B_s^0 \to \phi\phi$ decays. The nominal fit used the histogram directly. The fitted time acceptance was used to determine the systematic uncertainty.

Parameter	Value
a	1.94 ± 0.15
b	4.20 ± 0.07
c	0 (fixed)
d	1.51 ± 0.03
e	1.81 ± 0.14

Table 6.4.: Fit result for the nominal parametrisation of the proper time acceptance in the $B_s^0 \to \phi\phi$ decay.

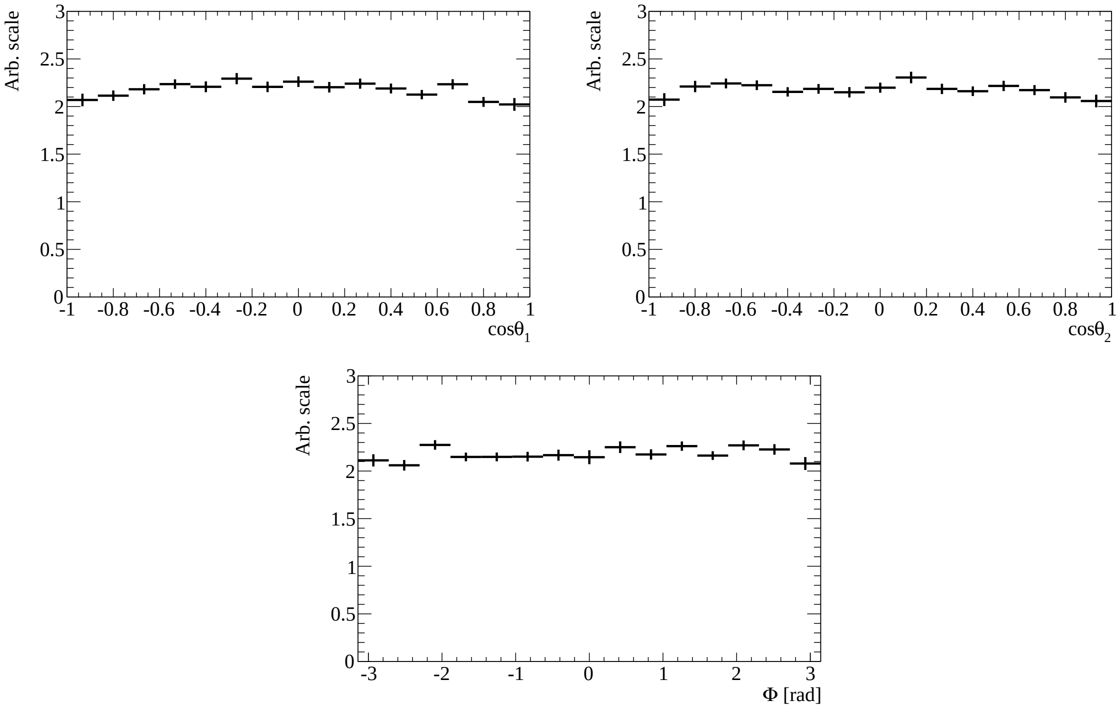


Figure 6.6.: Angular acceptance for the $\cos\theta_1$ (top-left), $\cos\theta_2$ (top-right), and Φ (bottom) helicity angles obtained from fully selected simulated events.

6.6. Angular Acceptance

The multi-variate selection requirements used for the decay-time dependent measurement introduce efficiencies that depend on the helicity angles. These efficiencies, calculated from fully-selected simulated events are shown in Figure 6.6, where the drop in efficiency as $\cos\theta_{1,2}$ approaches ±1 is explained by the kinematic cuts imposed on the kaon tracks.

The method of the angular acceptance weights, used to correct for angular acceptances, was the same as used in the decay time integrated analysis detailed in Section 5.2.3. The weights used to describe the acceptance in Figure 6.6 are given in Table 6.5. The efficiencies shown in Figure 6.6 appear relatively flat, therefore no large deviations from 1 are observed for the cases of the orthogonal angular functions (f_1, f_2, f_3, f_7 and f_8), and no large deviations from zero are observed for the case of the functions describing interferences.

Angular Term	Weights
f_1	0.971 ± 0.007
f_2	1.021 ± 0.007
f_3	1.024 ± 0.005
f_4	0.002 ± 0.009
f_5	-0.021 ± 0.009
f_6	0.013 ± 0.007
f_7	1.001 ± 0.005
f_8	0.982 ± 0.009
f_9	-0.021 ± 0.021
f_{10}	0.011 ± 0.012
f_{11}	0.008 ± 0.013
f_{12}	-0.013 ± 0.011
f_{13}	-0.016 ± 0.021
f_{14}	0.016 ± 0.017
f_{15}	-0.016 ± 0.013

Table 6.5.: Normalisation weights used for angular acceptance systematic uncertainties for the $B_s^0 \to \phi\phi$ decay.

6.7. Fit Details and Ingredients

The ${}_s\mathcal{P}lot$ method, discussed in Section 3.4, was used to initially assign s-weights to the data, based on the fit to the four-kaon invariant mass, described in Section 6.1. This then enabled the fit to be performed to the distributions of decay time and helicity angles using only the signal component. The PDF, $P(t, \theta_1, \theta_2, \Phi)$, used in the fitting of the data was then of the form

$$P(t, \theta_1, \theta_2, \Phi) = G(t) \otimes F(t, \theta_1, \theta_2, \Phi), \tag{6.31}$$

where $F(t, \theta_1, \theta_2, \Phi)$ contains the decay time-dependent angular distribution of the $B_s^0 \to \phi\phi$ decay, described in Section 6.2, and $G(t)$ is a Gaussian function describing the time resolution that was convolved with the decay time-dependent angular distribution.

6.7.1. Parameters and External Inputs

In addition to the external inputs described in Section 6.3 related to the calibration of the flavour tagging, the values of the B_s^0-$\overline{B}_s^0$ mixing frequency and physical eigenstate decay rates are also needed due to the small size of the dataset. The physical eigenstate decay rates, parametrised by Γ_s and $\Delta\Gamma_s$, were constrained to the LHCb measurements obtained from the $B_s^0 \to J/\psi\,\phi$ decay of [15]

$$\Gamma_s = 0.663 \pm 0.005(\text{stat}) \pm 0.006(\text{syst})\,\text{ps}^{-1}, \tag{6.32}$$

$$\Delta\Gamma_s = 0.100 \pm 0.016(\text{stat}) \pm 0.003(\text{syst})\,\text{ps}^{-1}. \tag{6.33}$$

The constraint was applied taking in to account the correlation between Γ_s and $\Delta\Gamma_s$ of $\rho(\Gamma_s, \Delta\Gamma_s) = -0.30$. The value of the B_s^0-$\overline{B}_s^0$ mixing frequency was constrained to an LHCb measurement in $B_s^0 \to D_s^- \pi^+$ decays of [71]

$$\Delta m_s = 17.725 \pm 0.041(\text{stat}) \pm 0.026(\text{syst})\,\text{ps}^{-1}. \tag{6.34}$$

The parameters found in the functions forming the PDF in equation 6.31 are summarised in Table 6.6, along with how each parameter was treated in the fit. The terms of $F(t, \theta_1, \theta_2, \Phi)$ involving $|A_{SS}|^2$ were neglected in nominal results as the size of the dataset was considered too small to reliably determine the associated parameters. The effect of neglecting such a contribution was thus treated as a systematic uncertainty.

Function	Parameter	Type	Value
$F(t,\theta_1,\theta_2,\Phi)$	$\phi^{s s\bar{s}}$	Free	
	$\lvert A_0\rvert^2$	Free	
	$\lvert A_\perp\rvert^2$	Free	
	$\lvert A_\parallel\rvert^2$	Constrained	$1-\lvert A_0\rvert^2-\lvert A_\perp\rvert^2$
	$\lvert A_S\rvert^2$	Free	
	δ_1	Free	
	δ_2	Free	
	δ_S	Free	
	Γ_s, $\Delta\Gamma_s$	Constrained	LHCb measurement [15]
	Δm_s	Constrained	LHCb measurement [71]
	p_0^{OS}, p_1^{OS}	Constrained	LHCb measurement [15]
	η^{OS}	Fixed	LHCb measurement [15]
	p_0^{SSK}, p_1^{SSK}	Constrained	LHCb measurement [15]
	η^{SSK}	Fixed	LHCb measurement [15]
	p_0^{BOTH}	Constrained	LHCb measurement [15]
	η^{BOTH}, p_1^{BOTH}	Fixed	LHCb measurement [15]
$G(t)$	σ_t	Fixed	40 fs

Table 6.6.: Summary of the parameters contained in the functions of equation 6.31 with associated fit treatment.

6.8. Fit Results

The fit results to s-weighted distributions of the helicity angles and B_s^0 decay time are shown in Table 6.7 for parameters that are either left free or are constrained. A point estimate is not provided for the $\phi^{s s\bar{s}}$ parameter as in Figure 6.8, the likelihood is only parabolic to between 1σ and 2σ. Therefore, a confidence region is quoted instead of a point estimate. In Table 6.7, the 68 % confidence level is quoted for $\phi^{s s\bar{s}}$. This contains statistical uncertainties only. The s-weighted distributions of the helicity angles and the B_s^0 decay time are shown in Figure 6.7. Fit components corresponding to the *CP*-even, *CP*-odd and *S*-wave terms in the $B_s^0 \to \phi\phi$ PDF are overlaid. It is worth reminding that the choice of which ϕ meson is used to determine θ_1 and which is used to determine θ_2 was randomised, hence any structure seen in one projection that is not in the other is purely a result of statistical fluctuations. This randomisation has no impact on the physics parameters due to the symmetry of the PDF in the $\cos\theta_1$ and $\cos\theta_2$ observables.

The likelihood profile for the *CP*-violating weak phase $\phi^{ss\bar{s}}$, shown in Figure 6.8, is not parabolic. This is due to the relatively small number of events used to determine this parameter.

Parameter	Fitted Value
$\phi^{ss\bar{s}}$ (rad) 68% CL	$(-2.37, -0.92)$
$\lvert A_\perp \rvert^2$	0.358 ± 0.046
$\lvert A_0 \rvert^2$	0.329 ± 0.033
$\lvert A_S \rvert^2$	$0.016^{+0.024}_{-0.012}$
δ_S (rad)	$0.65^{+0.89}_{-1.65}$
δ_1 (rad)	2.19 ± 0.44
δ_2 (rad)	-1.47 ± 0.48
Γ_s (ps^{-1})	0.660 ± 0.008
$\Delta\Gamma_s$ (ps^{-1})	0.106 ± 0.017
Δm_s (ps^{-1})	17.74 ± 0.05
p_0^{OS}	0.001 ± 0.008
p_1^{OS}	1.000 ± 0.023
p_0^{SSK}	0.001 ± 0.017
p_1^{SSK}	0.989 ± 0.159
p_0^{BOTH}	0.002 ± 0.025

Table 6.7.: Raw fit results to the s-weighted dataset. Note that the value of the predicted mistag rate in the dataset has been pre-calibrated, therefore the central values of the calibration parameters are quoted relative to perfect calibration, i.e. $p_0 = 0$ and $p_1 = 1$.

6.9. Systematic Uncertainties

The dominant sources of systematic uncertainty include the angular and B_s^0 decay time acceptances, in addition to the knowledge of the *S*-wave. The uncertainties due to the external inputs, described in Section 6.7 were accounted for directly in the fit with the use of Gaussian constraints. This therefore increases the statistical uncertainties accordingly.

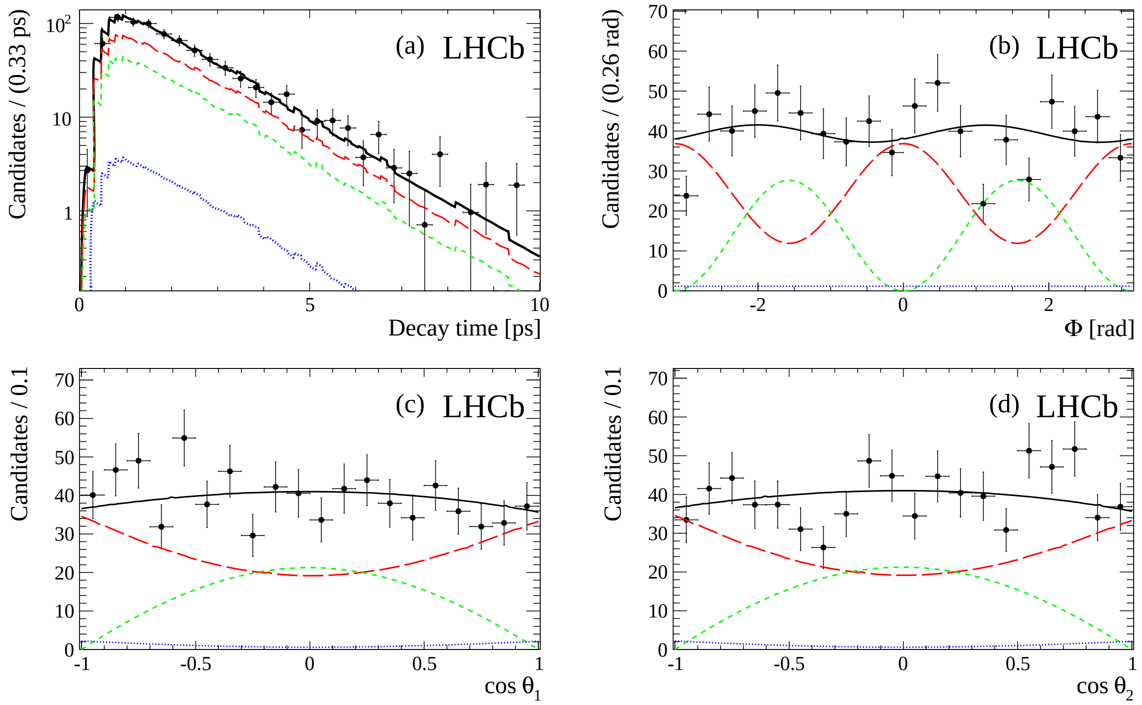


Figure 6.7.: Fit projections and s-weighted data distributions for the $B_s^0 \to \phi\phi$ (a) decay time, (b) helicity angle Φ and the cosine of the helicity angles (c) $\cos\theta_1$ and (d) $\cos\theta_2$. The data are marked as points, while the solid lines represent the projections of the best fit. The *CP*-even *P*-wave, the *CP*-odd *P*-wave and *S*-wave components are shown by the long dashed, short dashed and dotted lines, respectively.

6.9.1. Angular Acceptance

Two sources of uncertainty were considered relating to the angular acceptance. These are the statistical uncertainty on the normalisation weights, arising from the limited number of simulated events used to determine the acceptance weights given in Table 6.5, and also the kinematic disagreement between data and simulated distributions.

The method of accounting for the uncertainty on the angular acceptance weights was identical to that used in the decay time-integrated measurement, described in Section 5.2.5.2. With this method, angular acceptance weights were picked at random from Gaussian distributions each having a width equal to the decorrelated error. The decorrelated variation was then added to the central values of the acceptance weights and a fit was made to 10000 simulated events. This was repeated 1000 times and the width of a Gaussian fit to the distribution of the bias on each parameter was taken to be the systematic uncertainty for that parameter.

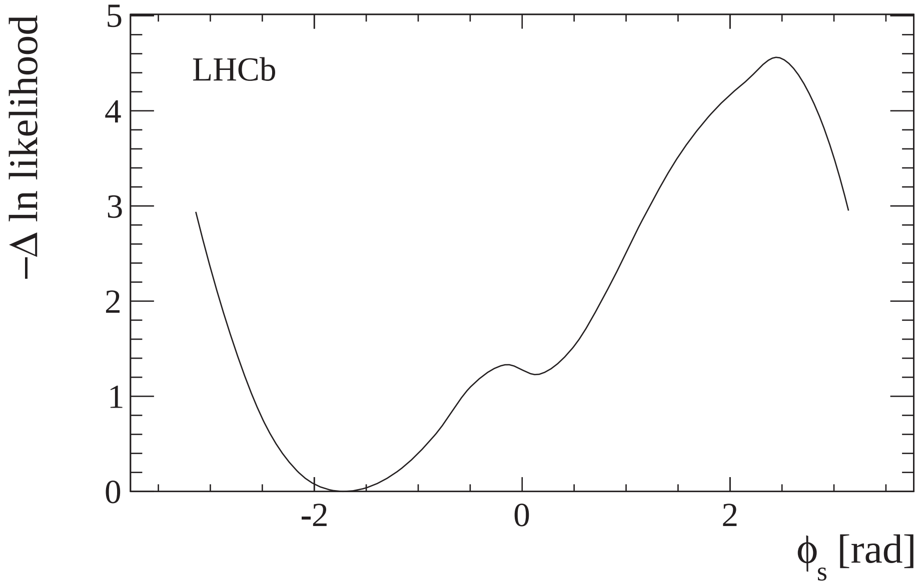


Figure 6.8.: Negative Δln likelihood scan of $\phi^{ss\bar{s}}$. Only the statistical uncertainty is included.

In addition to the statistical uncertainty on acceptance weights, differences are known to exist in the kinematic distributions of particles between data and simulation. The full comparison between data and simulation is given in Appendix E. From the full comparison, it can be seen that the four largest differences between data and simulation are from the maximum kaon p_{T}, the B_s^0 p_{T}, the B_s^0 η and the maximum track χ^2 per NDF distributions. To account for these differences, acceptance weights were recalculated using re-weighted simulated events and further fits to the data performed. The recalculated weights are shown in Table 6.8. The largest difference in the central values shown in Table 6.9 are taken as systematic uncertainties.

A summary of the systematic uncertainties is given in Table 6.9.

6.9.2. Decay Time Acceptance

The difference between the fit result obtained using the decay time acceptance in the form of a histogram and the fit result from the decay time acceptance in the form of the parametrisation given in equation 6.30 was used to determine the systematic uncertainty. The reason for this can be seen in Figure 6.5, where structures are seen at $\sim 2.5\,\mathrm{ps}$ and $\sim 3.5\,\mathrm{ps}$. The systematic uncertainties due to decay time acceptance are shown in Table 6.10.

Angular Term	Weights (K p_{T})	Weights (B_s^0 p_{T})	Weights (B_s^0 η)	Weights (Track χ^2)
f_1	0.977	0.969	0.970	0.958
f_2	1.016	1.023	1.019	1.028
f_3	1.020	1.025	1.023	1.037
f_4	−0.003	0.001	0.002	0.006
f_5	−0.028	−0.022	−0.021	−0.026
f_6	0.012	0.012	0.013	0.012
f_7	0.999	1.000	1.001	1.002
f_8	0.988	0.983	0.987	0.975
f_9	−0.018	−0.023	−0.020	−0.018
f_{10}	0.011	0.009	0.010	0.021
f_{11}	0.008	0.010	0.010	0.017
f_{12}	−0.011	−0.012	−0.013	−0.021
f_{13}	−0.019	−0.018	−0.018	−0.014
f_{14}	0.014	0.013	0.013	0.012
f_{15}	−0.021	−0.021	−0.015	−0.019

Table 6.8.: Angular acceptance weights used for angular acceptance systematic uncertainties due to kinematic re-weighting.

6.9.3. *S*-wave

As discussed in Section 6.2.1, the S-wave and the P-wave components are described by differing parametrisations in terms of $m_{\mathrm{K^+K^-}}$. Therefore, the distributions in $m_{\mathrm{K^+K^-}}$ provide a first order estimate of the total S-wave present. Figure 6.9 shows the result of a background-subtracted fit to the $m_{\mathrm{K^+K^-}}$ distribution where a Flatté function [72] has been used to model the S-wave $m_{\mathrm{K^+K^-}}$ distribution and a relativistic Breit-Wigner [73] of width fixed to $G_0 = 4.26\ \mathrm{MeV}/c^2$ [10], convoluted with a Gaussian resolution function has been used to model the P-wave $m_{\mathrm{K^+K^-}}$ distribution. The choice of a Flatté function assumes that the S-wave originates from an f_0 resonance. Therefore, ϕf_0 and $f_0 f_0$ are used to refer to the contributions from the A_S and A_{SS} amplitudes, respectively. The Flatté function [72] is given by

$$F(m_{K^+K^-}) = \left|\frac{1}{(m_{f_0}^2 - m_{K^+K^-}^2) - i m_{f_0}(g_K \Gamma_{KK} + g_\pi \Gamma_{\pi\pi})}\right|^2, \qquad (6.35)$$

Parameter	Diff. (K p_{T})	Diff. (B_s^0 p_{T})	Diff. (B_s^0 η)	Diff. (B_s^0 η)	Systematic uncertainty (RW)	Systematic uncertainty (stat)
$\phi^{s s \bar{s}}$ (rad)	0.01	0.00	0.00	0.02	0.02	0.02
$\|A_0\|^2$	0.004	0.001	0.000	0.007	0.007	0.002
$\|A_\perp\|^2$	0.003	0.000	0.000	0.004	0.004	0.003
$\|A_S\|^2$	0.0001	0.0000	0.0000	0.0004	0.0004	0.0004
δ_2 (rad)	0.03	0.00	0.01	0.01	0.03	0.05
δ_1 (rad)	0.01	0.00	0.00	0.02	0.02	0.04
δ_S (rad)	0.03	0.01	0.01	0.02	0.03	0.06

Table 6.9.: Systematic uncertainties due to the recalculation of acceptance weights with reweighted (RW) simulation and due to the statistical uncertainties on the angular acceptance weights.

Parameter	Uncertainty
$\phi^{s s \bar{s}}$ (rad)	0.09
$\|A_0\|^2$	0.003
$\|A_\perp\|^2$	0.005
$\|A_S\|^2$	0.001
δ_2 (rad)	0.02
δ_1 (rad)	0.02
δ_S (rad)	0.05

Table 6.10.: Systematic uncertainties due to time acceptance for the decay time-dependent measurement of $B_s^0 \to \phi\phi$.

where $\Gamma_{KK(\pi\pi)} = \sqrt{|m^2_{K^+K^-}/4 - m^2_{K(\pi)}|}$, m_{f_0} is the mass of the f_0 meson, and $m_{K(\pi)}$ is the kaon(pion) mass. The kaon(pion) couplings, $g_{K(\pi)}$, were fixed to the LHCb measured values [74]. The relativistic Breit-Wigner function [73] is given by

$$B(m_{K^+K^-}) = \frac{m_{K^+K^-} m_\phi G}{(m_\phi^2 - m^2_{K^+K^-})^2 + m_\phi^2 G^2}, \tag{6.36}$$

where $G = G_0 \frac{m_\phi}{m_{K^+K^-}} \left(\sqrt{\frac{m^2_{K^+K^-} - 4m_K^2}{m_\phi^2 - 4m_K^2}} \right)^3$, G_0 is the natural ϕ width, and m_ϕ is the mass of the ϕ meson. The fit yields a $2.0 \pm 1.1\%$ S-wave fraction including both ϕf_0 and $f_0 f_0$ S-wave components. The fit parameters of the line-shapes are given in Table 6.11.

Parameter	Value
Breit-Wigner mean m_ϕ (MeV/c^2)	1019.7 ± 0.1
Mass resolution σ_ϕ (MeV/c^2)	1.1 ± 0.2
Flatté mean m_{f_0} (MeV/c^2)	980 (fixed)
Flatté coupling g_K (MeV)	597 (fixed)
Flatté coupling g_π (MeV)	199 (fixed)

Table 6.11.: Fitted parameters of the *P*-wave and *S*-wave line-shapes.

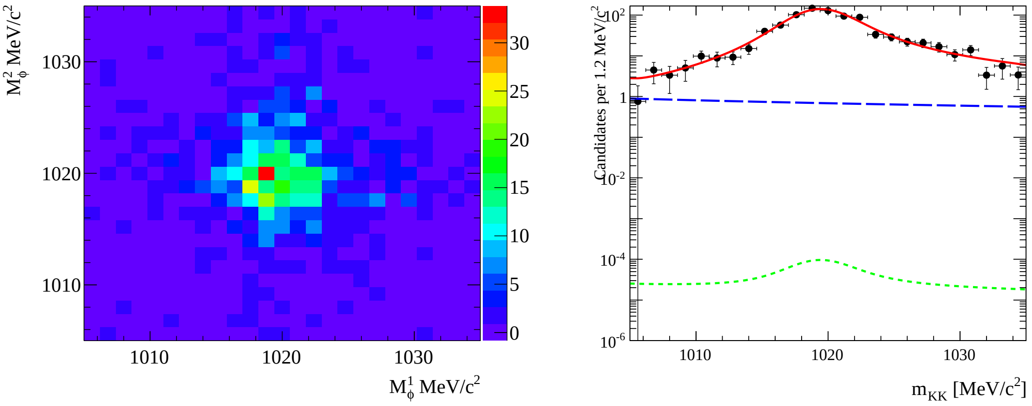


Figure 6.9.: $m^1_{K^+K^-}$ versus $m^2_{K^+K^-}$ distribution for $B_s^0 \to \phi\phi$ events (left) and fit projections on to $m^1_{K^+K^-}$ (right). The total fit is shown in red, the ϕf_0 component is shown in green and the $f_0 f_0$ component is shown in blue.

The shapes of the fitted *S*-wave components are shown in Figure 6.10. In the nominal result, an s-weighted fit was chosen with only the $B_s^0 \to \phi f_0$ component contributing. This was chosen as the fits to $m_{K^+K^-}$ invariant mass indicate that a total of around 2 % *S*-wave was expected. It is physically expected that $|A_S| \sim |A_{SS}|^2$, which would therefore mean a negligible contribution from the A_{SS} amplitude.

Table 6.12 shows the results of fits to the dataset allowing successively larger numbers of *S*-wave parameters to float in the fit. As can be seen from Table 6.12, the $B_s^0 \to f_0 f_0$ *S*-wave component fits to large values but with large statistical uncertainties. However this does not affect the *P*-wave parameters to a large degree. Figure 6.11 shows the likelihood scan of the $\phi^{ss\bar{s}}$ parameter for the three different levels of *S*-wave. In all cases, the addition of *S*-wave terms causes small effects on *P*-wave physics parameters. The difference between the physics parameters for the case of the nominal fit and the fit with the full *S*-wave component free was included as a systematic uncertainty.

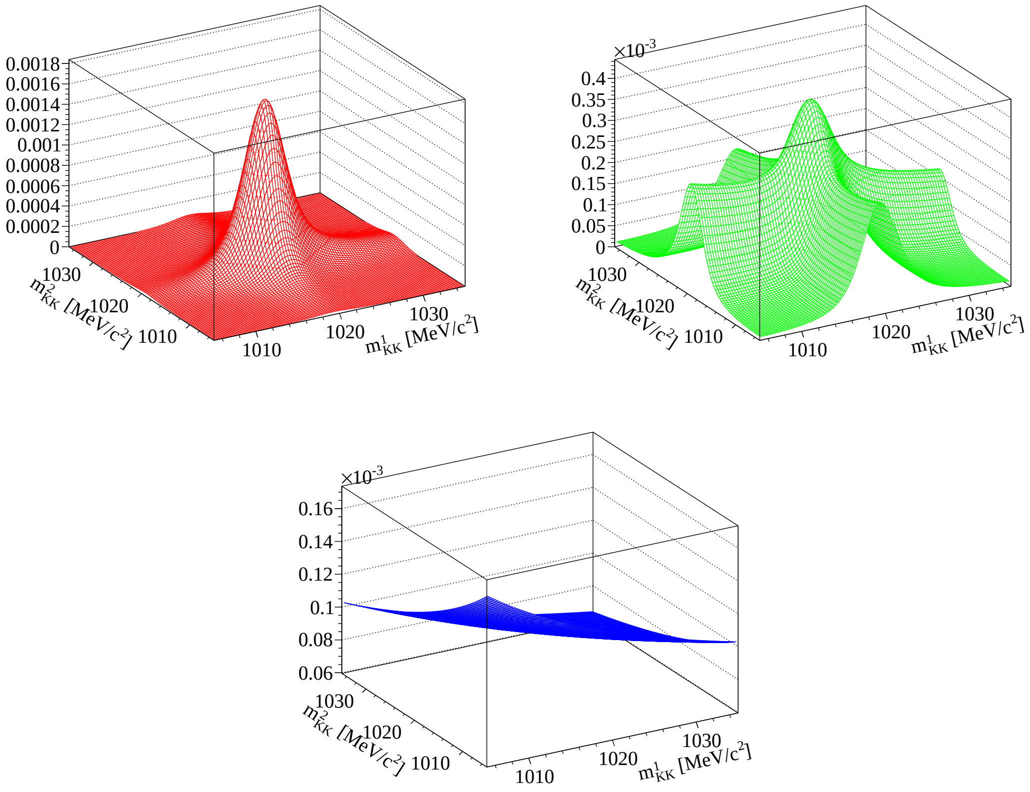


Figure 6.10.: Fitted shapes of the *P*-wave (red), ϕf_0 *S*-wave (green) and $f_0 f_0$ *S*-wave (blue). Note the scale of the z-axis is arbitrary.

6.9.4. Mass Model

The systematic effect of the mass model was taken in to account by using an alternative s-weighting based on a single Gaussian model compared to the double Gaussian described in Section 6.1. The single Gaussian was found to have a width of $(15.5 \pm 0.5)\,\mathrm{MeV}/c^2$ and a mean of $(5364.6 \pm 0.6)\,\mathrm{MeV}/c^2$. The fit used in the calculation of s-weights for systematic uncertainties is shown in Figure 6.12. The difference between the fitted parameters in the s-weighted fit for these s-weights and those of the nominal s-weights are taken as systematic uncertainties. The effect on the physics parameters used as systematic uncertainties is given in Table 6.13.

Parameter	$\phi\phi$ fit	$\phi\phi$ and ϕf_0 fit	$\phi\phi$, ϕf_0 and $f_0 f_0$ fit
$\phi^{ss\bar{s}}$ (rad)	$(-2.35, -1.00)$	$(-2.37, -0.92)$	$(-2.14, -0.76)$
$\|A_0\|^2$	0.335 ± 0.028	0.329 ± 0.033	0.342 ± 0.031
$\|A_\perp\|^2$	0.355 ± 0.040	0.358 ± 0.046	0.343 ± 0.044
$\|A_S\|^2$	0 (fixed)	$0.016^{+0.024}_{-0.012}$	$0.008^{+0.015}_{-0.007}$
$\|A_{SS}\|^2$	0 (fixed)	0 (fixed)	0.090 ± 0.055
δ_1 (rad)	2.12 ± 0.39	2.19 ± 0.44	2.08 ± 0.43
δ_2 (rad)	-1.59 ± 0.40	-1.47 ± 0.48	-1.43 ± 0.46
δ_S (rad)	0 (fixed)	$0.65^{+0.89}_{-1.65}$	$0.43^{+0.95}_{-1.19}$
δ_{SS} (rad)	0 (fixed)	0 (fixed)	-2.43 ± 0.49
Γ_s (ps^{-1})	0.660 ± 0.008	0.660 ± 0.008	0.660 ± 0.008
$\Delta\Gamma_s$ (ps^{-1})	0.105 ± 0.016	0.106 ± 0.017	0.107 ± 0.016
Δm_s (ps^{-1})	17.73 ± 0.05	17.74 ± 0.05	17.73 ± 0.05

Table 6.12.: Comparison of the different s-weighted fit results allowing various components of the *S*-wave to float in the fit.

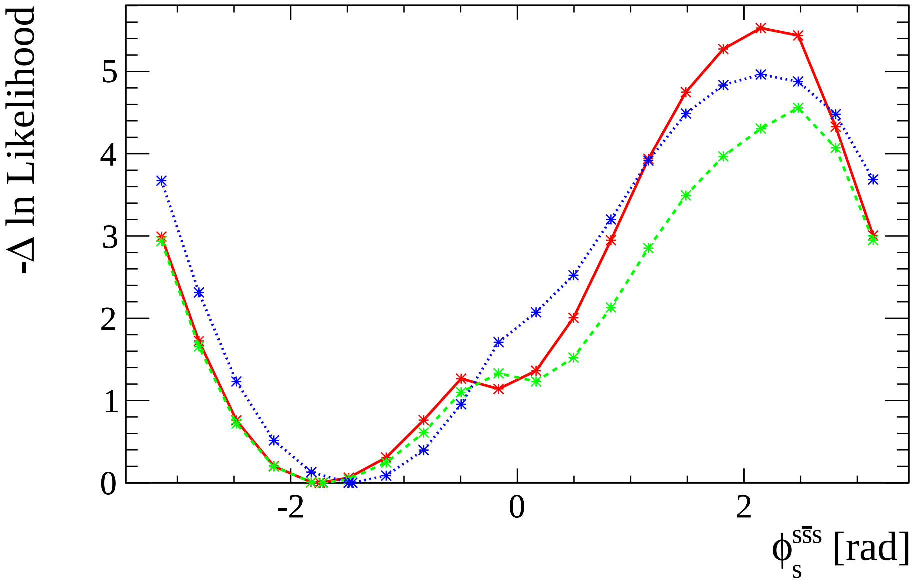


Figure 6.11.: Likelihood scan for the $\phi^{ss\bar{s}}$ parameter obtained from the *P*-wave only fit (red solid), the *P*-wave in addition to the $B_s^0 \to \phi f_0$ component (green dashed) and the *P*-wave in addition to all *S*-wave (blue dotted).

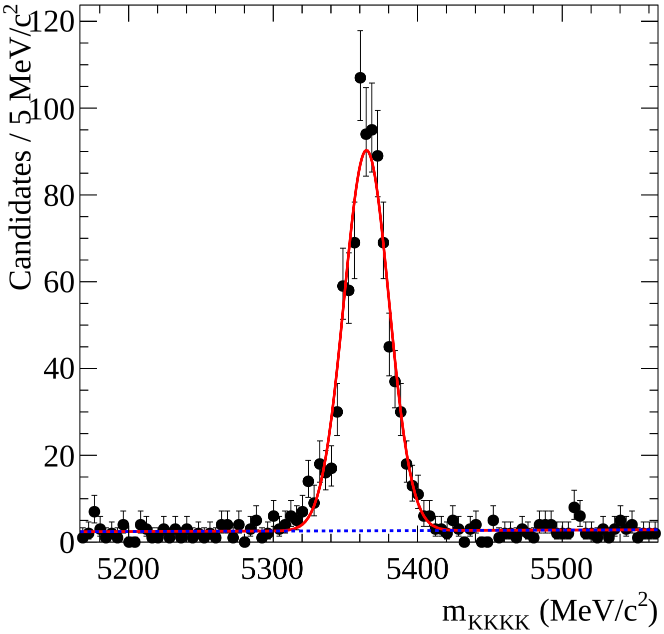


Figure 6.12.: Single Gaussian signal mass fit used to generate s-weights for the mass model systematic uncertainty.

Parameter	Uncertainty
$\phi^{s\bar{s}s}$ (rad)	0.02
$\lvert A_0 \rvert^2$	0.004
$\lvert A_\perp \rvert^2$	0.002
$\lvert A_S \rvert^2$	0.003
δ_2 (rad)	0.01
δ_1 (rad)	0.02
δ_S (rad)	0.10

Table 6.13.: Systematic uncertainties due to the signal model used for s-weight generation.

6.9.5. Peaking Background

The major source of background expected for the $B_s^0 \to \phi\phi$ decay is that of the $B^0 \to \phi K^{*0}$ decay, in the case that a pion is mis-identified as a kaon being reflected in to the B_s^0 mass peak. In order to quantify how many candidates we would expect to find in our data, 1.5 million $B^0 \to \phi K^{*0}$ simulated events were reconstructed and selected in an identical way to the data itself. Table 6.14 shows the amount of simulated events that pass each stage of the selection. The distributions in K^+K^- and $K^+K^-K^+K^-$ invariant

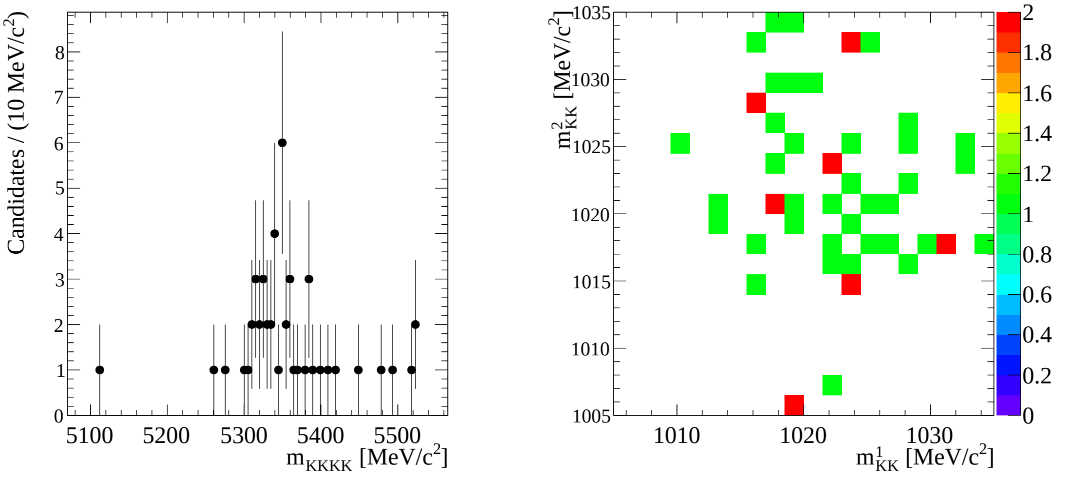


Figure 6.13.: Distributions of $m_{\mathrm{K^+K^-K^+K^-}}$ (left) and $m_{K^+K^-}$ versus $m_{K^+K^-}$ (right) distributions for $B^0 \to \phi K^{*0}$ MC events passing all selection stages.

masses for the events that pass all selection stages are shown in Figure 6.13, where it can be seen that mis-identified events peak at just below the B_s^0 mass. The branching

Selection stage	# Simulated events passed
Generation	$1,500,000$
Stripping and reconstruction	6612
Pre-selection and trigger	117
Final cuts	51

Table 6.14.: Number of $B^0 \to \phi K^{*0}$ simulated events passing each selection stage.

ratio of $B^0 \to \phi K^{*0}$ has been previously measured to be $(9.8 \pm 0.6) \times 10^{-6}$ [10]. The number of $B^0 \to \phi K^*(892)$ decays expected to be produced in 1.0 fb^{-1} of pp collision data inside the LHCb detector in the $KKK\pi$ final state is given through

$$N_{B^0\to\phi K^{*0}} = 2 \times \mathcal{L}_{\mathrm{Int}} \times \sigma_{b\bar{b}}^{\mathrm{acc}} \times f_d \times \mathcal{B}(B^0 \to \phi K^{*0}) \times \mathcal{B}(\phi \to KK) \times \mathcal{B}(K^{*0} \to K\pi), \tag{6.37}$$

where f_d is the probability for a b-quark to produce a B^0 meson, $\mathcal{L}_{\mathrm{Int}}$ is the integrated luminosity and $\sigma_{b\bar{b}}^{\mathrm{acc}}$ is the cross section of hadrons produced in the LHCb acceptance. The values used are shown in Table 6.15. When these are input into equation 6.37, this gave rise to 192069 events produced in the $\mathrm{K^+K^-K^+}\pi^-$ final state inside the LHCb acceptance. Scaling the mis-identified simulated events according to the number expected in $1.0\,\mathrm{fb}^{-1}$ yields approximately 6 events in the dataset or 0.7 % of the signal yield. This is then taken to be a negligible number, hence no systematic uncertainty is assigned.

Parameter	Value
$\sigma_{b\bar{b}}^{\mathrm{acc}}$	$(75.3 \pm 5.4 \pm 13.0)$ µb [75]
f_d	0.399 ± 0.011 [10]
$\mathcal{B}(B^0 \to \phi K^{*0})$	$(9.8 \pm 0.6) \times 10^{-6}$ [10]
$\mathcal{B}(\phi \to \mathrm{K}^+\mathrm{K}^-)$	0.489 ± 0.005 [10]
$\mathcal{B}(K^{*0} \to K\pi)$	0.667 [10]

Table 6.15.: Reference values used in equation 6.37.

6.9.6. Decay Time Resolution

It has been previously found in time-dependent studies of the $B_s^0 \to J/\psi\,\phi$ decay that simulated events provide a time resolution that is 5 fs smaller than resolutions obtained from data-driven methods. Therefore, the differences between the results obtained with the nominal time resolution and a resolution increased by 5 fs were taken as systematic uncertainties. This results in uncertainties of $\sigma(\phi^{s s \bar{s}}) = 0.01\,\mathrm{rad}$ and $\sigma(|A_S|^2) = 0.0001$.

6.9.7. Fit Bias

The negative log-likelihood (NLL) method contains an inherent bias that arises when the dataset size is not sufficiently large enough to obtain a Gaussian likelihood. Biases are accounted for in the systematic uncertainties through the use of simplified simulation studies. For these studies, datasets of the same size as the signal, with the same tagging efficiencies as the data are used. For the case of tagged events, an average mistag rate of 0.374 was used. The same physics parameters as the nominal s-weighted fit result are used to generate simulated datasets. In total, one thousand simulations are performed. The mean values of the Gaussian fits to the pull distributions are used to assign the systematic uncertainty. The pull distributions obtained from the simplified simulations are given in Figure 6.14. The resulting uncertainties are given in Table 6.16.

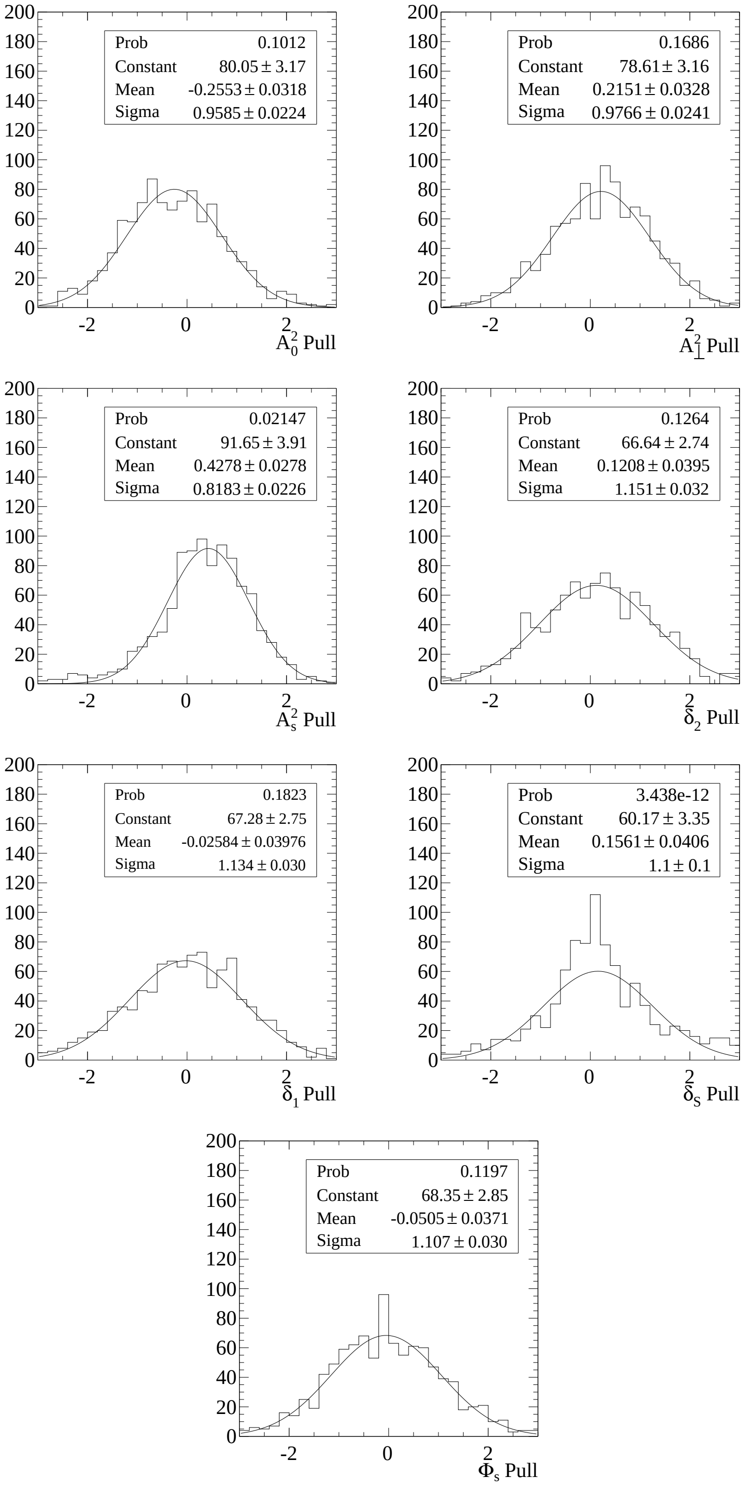


Figure 6.14.: Pull distributions from simplified simulations used to obtain the systematic uncertainty related to fit bias of the NLL method.

Parameter	Uncertainty
$\phi^{ss\bar{s}}$	0.02
$\|A_0\|^2$	0.006
$\|A_\perp\|^2$	0.006
$\|A_s\|^2$	0.003
δ_2 (rad)	0.06
δ_1 (rad)	0.01
δ_S (rad)	0.21

Table 6.16.: Systematic uncertainties due to fit bias.

6.10. Feldman Cousins Coverage Correction

In the case that low numbers of events are used in the NLL method to make measurements, log-likelihoods are not parabolic which may give rise to so-called under-coverage. The method of Feldman and Cousins that corrects for this under-coverage has been explained in Section 3.2. The basic principle of the Feldman Cousins method relies on the creation of toys at each point in a data likelihood scan. When toy datasets are generated for the scanned parameter at that point, a distribution of the difference in log likelihoods (DLL) is calculated between the fit with the parameter of interest fixed and free. The confidence level value was then calculated from the fraction of DLLs that fall below the DLL found from data.

The toys generated for the Feldman Cousins scan are fitted in the same way as the data in the nominal result. A scan was performed at 20 points in the 2σ window around the central value, with 500 toy datasets generated at each point. In the scan, three confidence levels were computed:

1. The Feldman Cousins scan as described previously (shown as red solid line in plots).
2. The confidence level from the data itself, calculated as the error function of the DLL found in data (shown as green dotted line in plots).
3. The theoretical confidence level, calculated in the same way as the CL from data with the exception that the DLL is assumed to be Gaussian around the measured central value with same width as the parabolic error measured in data (shown as blue dashed line in plots).

The result of the Feldman Cousins scan is shown in Figure 6.15. The Feldman

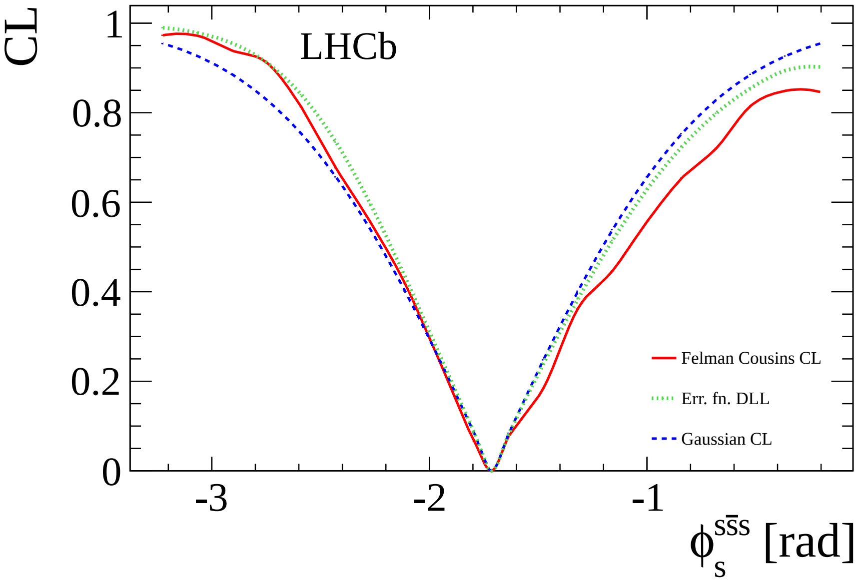


Figure 6.15.: Result of the $\phi^{ss\bar{s}}$ Feldman Cousins scan. The Feldman Cousins scan is shown by the red solid line. The CL curve found from the error function of the data likelihood scan is shown by the green dotted line and the CL found as the error function of a parabolic likelihood curve is shown by the blue dashed line.

Cousins scan is found to agree with the data quite well on the left side of the central value, while under-coverage is seen on the right side. The difference between the Feldman Cousins scan and the data confidence level, shown in Figure 6.15, was used to determine the coverage correction for the upper and lower limits of the 68 % CL. As the overall systematic uncertainty on $\phi^{ss\bar{s}}$ is small in relation to the statistical precision, the systematic uncertainty was added in quadrature to the upper and lower uncertainties of $\phi^{ss\bar{s}}$. This gives a coverage corrected 68 % CL of $[-2.46, -0.76]$ rad.

6.11. Final Results and Summary

The first decay time-dependent measurement in the $B_s^0 \to \phi\phi$ decay has been performed. This constitutes the first decay time-dependent measurement of a $b \to s\bar{s}s$ transition in the B_s^0 system. Results are summarised in Table 6.17. The *CP*-violating phase is

found to have a 68 % confidence level of $[-2.46, -0.76]$ rad. This does not contain the SM prediction, though the measurement is highly limited by statistical uncertainties. The probability of the SM hypothesis is found to be 16 %. Polarisation amplitudes are found to be consistent with those reported in Chapter 5, though the effect of allowing

Parameter	Value	$\sigma_{\text{stat.}}$	$\sigma_{\text{syst.}}$
$\phi^{s\bar{s}s}$ (rad) (68 % CL)		$[-2.46, -0.76]$	
$\|A_0\|^2$	0.329	0.033	0.017
$\|A_\perp\|^2$	0.358	0.046	0.018
$\|A_S\|^2$	0.016	$^{+0.024}_{-0.012}$	0.009
δ_1 (rad)	2.19	0.44	0.12
δ_2 (rad)	-1.47	0.48	0.10
δ_S (rad)	0.65	$^{+0.89}_{-1.65}$	0.33

Table 6.17.: Main fit results with statistical and systematic uncertainties. A 68 % statistical confidence interval is quoted for $\phi^{s\bar{s}s}$.

Parameter	Mass model	Angular acc.	Sim. re-weighting	Fit bias	Time res.	Time acc.	S-wave	Total
$\phi^{s\bar{s}s}$ (rad)	0.02	0.02	0.02	0.02	0.01	0.09	0.20	0.22
$\|A_0\|^2$	0.004	0.002	0.007	0.006	0.000	0.003	0.013	0.017
$\|A_\perp\|^2$	0.002	0.003	0.004	0.006	0.000	0.005	0.015	0.018
$\|A_S\|^2$	0.003	0.0004	0.0004	0.003	0.0001	0.001	0.008	0.009
δ_2 (rad)	0.01	0.05	0.03	0.06	0.00	0.02	0.04	0.10
δ_1 (rad)	0.02	0.04	0.02	0.01	0.00	0.02	0.11	0.12
δ_S (rad)	0.10	0.06	0.03	0.21	0.00	0.05	0.22	0.33

Table 6.18.: Summary of systematic uncertainties.

$\phi^{s\bar{s}s}$ to differ from zero causes the central values of the polarisation amplitudes to differ and the uncertainties to be larger than those reported in Chapter 5.

Systematic uncertainties are summarised in Table 6.18. The dominant sources of systematic uncertainties are found to be from decay time and angular acceptances, in addition to the knowledge of the S-wave. However, the dominant sources of uncertainty still remain significantly less than the statistical uncertainties.

Chapter 7.

Implications and Future Prospects

"We know what we are, but we know not what we may be."
— William Shakespeare

The measurements made with $B_s^0 \to \phi\phi$ events have in all cases been dominated by statistical uncertainties. The relatively suppressed branching ratio combined with detector efficiencies limit the dataset sizes available for measurements. The $B_s^0 \to \phi\phi$ decay is considered to be a golden mode for measuring *CP* violation in the upgraded LHCb detector, which will provide a much larger dataset size for more precise measurements. It is therefore the purpose of this Chapter to describe the upgraded LHCb detector and the corresponding improvements in detector performance, in addition to expectations of the precision of measurements made with the upgraded detector.

The LHCb detector has been performing remarkably well in the challenging environment of proton-proton collisions. However improvements to the detector have been planned and are due to be implemented in the long shutdown commencing in 2018.

The purpose of the Upgrade is to allow the LHCb detector to operate at a higher luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and an average number of visible interactions per crossing, μ, of 5 [76]. This will allow for an annual integrated luminosity of $5\,\mathrm{fb}^{-1}$ and greatly improves on the design luminosity of the current LHCb detector of $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

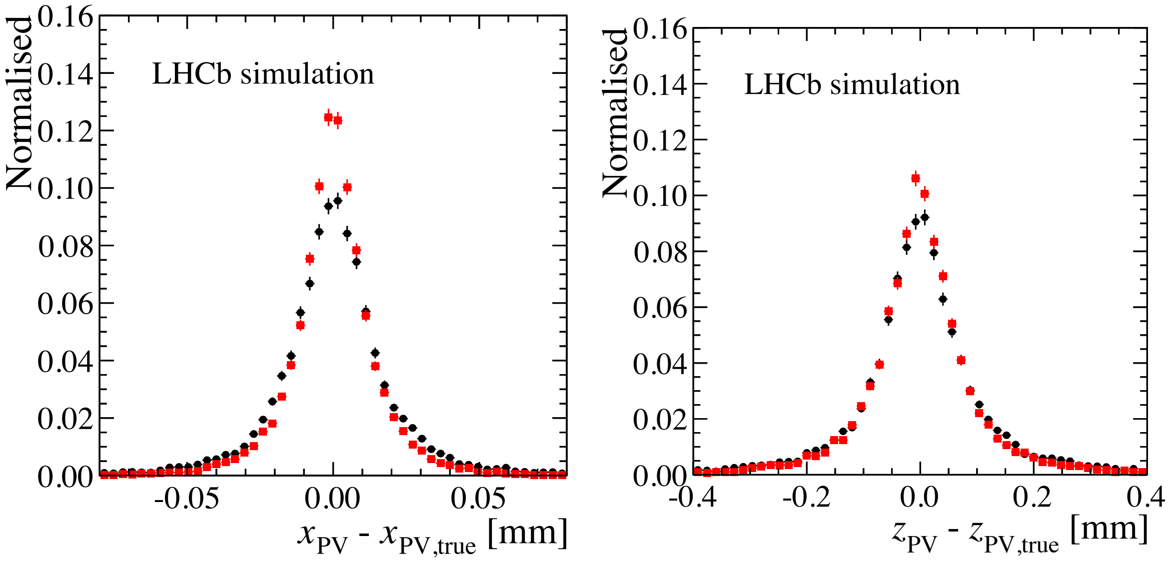


Figure 7.1.: Primary vertex resolution in the x (left) and z (right) directions. The red and black points show the performance of the upgraded VELO and current VELO, respectively [77].

7.1. Upgrade Components

The major focal points of the Upgrade are centred on the trigger system, in addition to the tracking stations, RICH detectors and vertex locator. All sub-detectors will have readout electronics upgraded, in order to cope with the requirement of a 40 MHz readout rate, corresponding to the LHC bunch crossing rate.

7.1.1. Vertex Locator

In order to cope with the higher occupancies associated with the luminosity of collisions in the upgrade and allow for a 40 MHz readout rate, a redesigned VErtex LOcator (VELO) is required. This redesign includes new readout architecture and silicon pixel sensors replacing the silicon strip sensors in the current VELO.

The upgraded VELO will replace the silicon strip sensors described in Section 2.3.2 with 12 silicon sensors in each L-shaped half station. The VELO will be comprised of 52 half-stations. Each silicon sensor will consist of a 256 × 256 array of pixels, with each pixel having an area of $55 \times 55\,\mu\text{m}^2$. The performance offered by this design is shown in Figure 7.1 in terms of primary vertex resolution, where a clear improvement is seen for the case of the upgraded VELO.

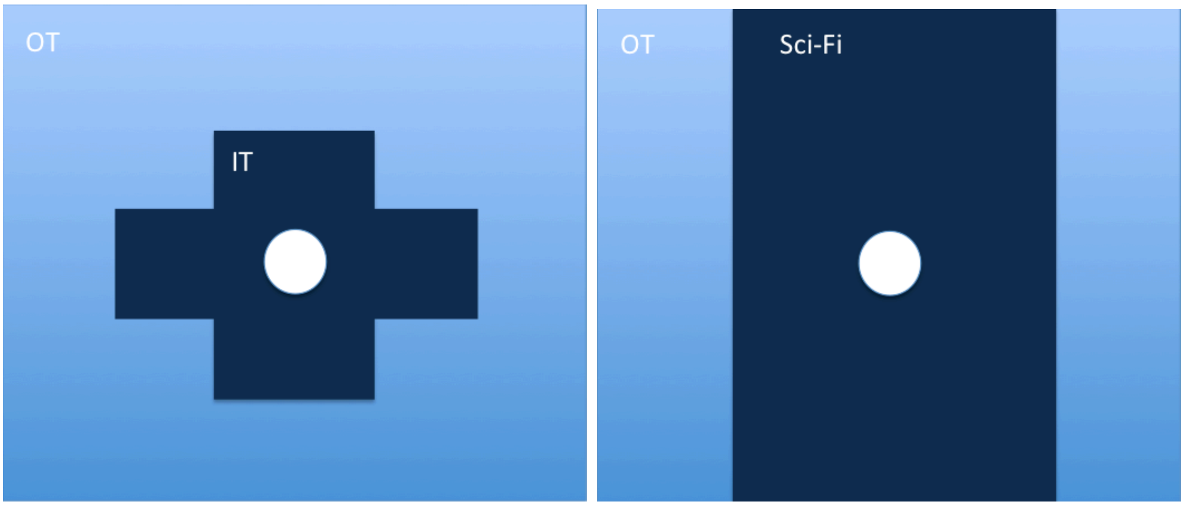


Figure 7.2.: Comparison of layouts in a tracking station between the current LHCb detector (left) and upgraded LHCb detector (right).

7.1.2. Tracking Stations

The luminosities provided in the LHCb upgrade give rise to occupancies that are much greater than 25 % for the regions of the current outer tracker that are located closest to the LHC beam axis. In order to remove this limitation, it is proposed that the current inner tracker and part of the outer tracker are replaced with a new scintillating fibre tracker, as depicted in Figure 7.2. The central principle of a scintillating fibre tracker involves the collection of photons produced by a charged particle traversing the fibre through the use of silicon photo-multipliers. Layers will be arranged at the same stereo angle rotation as found in the current trackers and will provide a spatial hit resolution of 60-100 µm. This will be achieved with the scintillating fibres having a diameter of 250 µm and with spatial position of the fibres known to within 6 µm.

7.1.3. RICH Detectors

The overall structure of the RICH detectors is to remain the same as for current data taking, however some major changes are required in order to read out the detector at 40 MHz in the higher luminosities of the Upgrade environment.

The Hybrid Photon Detectors (HPDs), described in Section 2.4.1, must be replaced by Multi-anode Photo Multiplier Tubes (MaPMTs), which are able to be read out at 40 MHz. In addition, the aerogel radiator, which is expected to produce poor performance at the

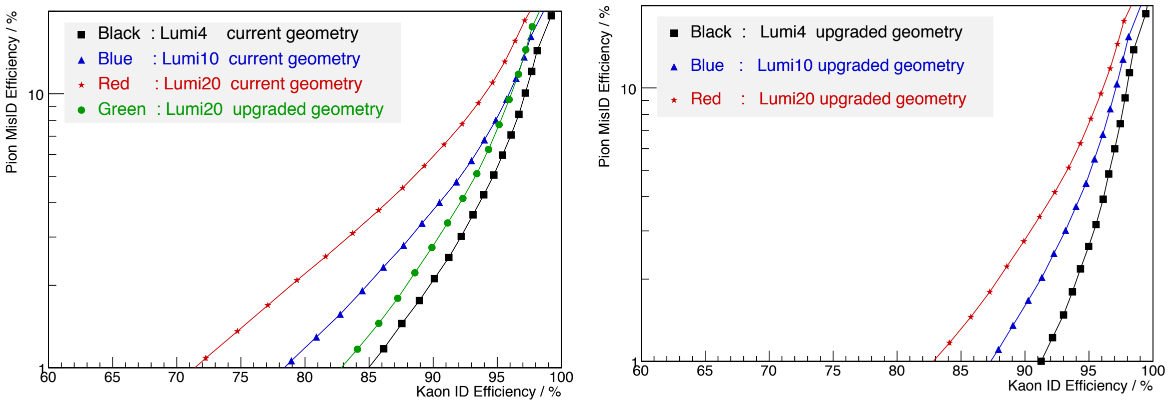


Figure 7.3.: PID performance in terms of kaon efficiency versus pion mid-identification rate for the current RICH sub-detector geometry compared to the optimised Upgrade geometry (left), and performance of the optimised Upgrade geometry in different beam conditions (right) [78].

luminosities of the upgraded LHCb experiment, will be removed. As the Cherenkov rings produced by interactions in the aerogel radiator have a larger radius than those produced in RICH 1 gas, the optical layout of the RICH 1 detector will be re-optimised to direct the Cherenkov photons on to a smaller total area of MaPMTs.

The particle identification (PID) performance of the current RICH detector is compared to that proposed for the Upgrade in Figure 7.3. In the comparisons, three different values of the luminosity are considered, corresponding to $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (Lumi4), $1 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (Lumi10), and $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (Lumi20). Plots are created by changing the value of the $\mathrm{DLL}_{K\pi}$ variable in simulated $B_s^0 \to \phi\phi$ events. It can be clearly seen in the plots that the upgraded LHCb detector offers better performance, especially in the luminosities of the Upgrade environment, described by Lumi20.

7.1.4. Trigger

In the current LHCb detector, the size of the Event Filter Farm (EFF), that processes the accepted L0 candidates, limits the L0 accept rate to 1 MHz. This in turn has a substantial impact on the trigger efficiency in hadronic B decays. The $B_s^0 \to \phi\phi$ Low Level Trigger (LLT) efficiency as a function of LLT accept rate is shown in Figure 7.4. A substantial part of the motivation for the LHCb Upgrade is to increase this efficiency through the enlargement of the EFF. It can be seen from Figure 7.4 that substantial increases in $B_s^0 \to \phi\phi$ efficiency will be gained from an increase in EFF size and hence an increase in LLT accept rate.

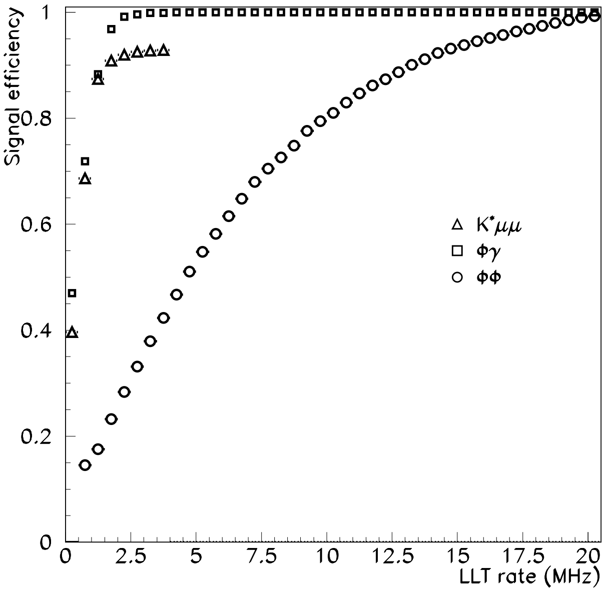


Figure 7.4.: LLT efficiency as a function of LLT accept rate for the case of $B_s^0 \to \phi\phi$, $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\mu\mu$ decays in upgrade conditions [79].

It is foreseen in the upgraded High Level Trigger (HLT), that the output of the first level of the HLT will be stored in a buffer that will allow for the full event reconstruction in the second level to be performed at a later time. This will then allow for a reconstruction to be performed that is very similar to that used in the final analyses, as improved calibrations of the particle identification and tracking can be incorporated.

The upgraded LHCb trigger is expected to use an EFF that is ten times the size of the current EFF. Initially, the EFF will be increased in size by a factor of five. Assuming an average HLT processing time of 20 ms per event, this would allow the output rate of the LLT to increase to 5 MHz. However, as the luminosity will also increase by a factor of five to $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, the selection requirements present in the LLT will remain roughly the same as 2011 conditions, which will mean a similar total trigger efficiency in the $B_s^0 \to \phi\phi$ decay of $\sim 29\,\%$ [79]. When the size of the EFF is then doubled so that there is a factor ten increase with respect to 2011 conditions, the trigger selections will be able to be loosened, providing a trigger efficiency of $\sim 50\,\%$ [79].

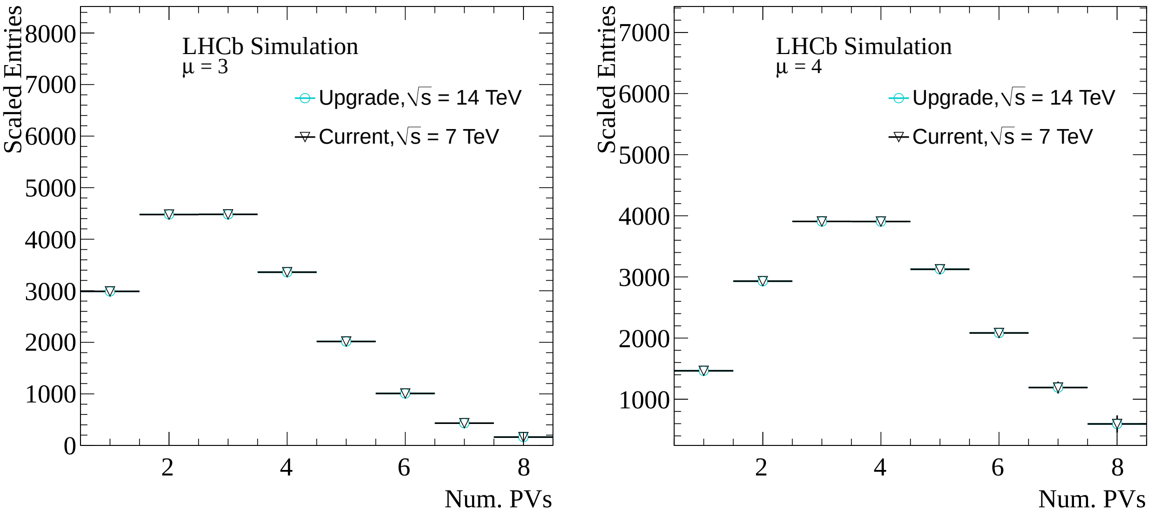


Figure 7.5.: Comparison of the distribution of the number of visible primary vertices for $\mu = 3$ and 4 as validation of the Poisson re-weighting.

7.2. Upgrade Performance

The simulation of the upgraded LHCb detector allows for the full reconstruction to be tested with every sub-detector included. Simulated event samples have been generated with two different beam configurations with an average number of proton-proton interactions per bunch crossing, ν, of 3.8 and 7.6. This corresponds to luminosities of 1×10^{33} and $2 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$, respectively. This then allows performance plots to be given as a function of the average number of visible proton-proton interactions per bunch crossing, μ. This is achieved through Poisson re-weighting according to the number of visible primary vertices in a given event. The validation of this procedure is given in Figure 7.5, where a Poisson distribution is seen for the number of visible primary vertices centred on the required value of μ. The average number of visible proton-proton interactions per bunch crossing is related to the average number of proton-proton interactions by the ratio of the difference in the total and elastic cross sections to the total cross section. This gives $\mu \sim 0.7\nu$.

The performance of the upgraded LHCb detector is compared to the current LHCb detector in terms of overall efficiencies in Section 7.2.1, efficiencies as a function of helicity angles and B_s^0 decay time in Section 7.2.2, and B_s^0 decay time and vertex resolution in Section 7.2.3.

7.2.1. Efficiency Comparison

Reconstruction and selection efficiencies directly affect the overall size of the datasets used to make measurements. It is therefore important to verify that the upgraded LHCb detector does not suffer losses in efficiency when compared to the current LHCb detector.

Future analyses of the $B_s^0 \to \phi\phi$ decay will use looser pre-selections than those applied to the dataset used in the measurement of the CP-violating phase, given in Table 4.4. Therefore for the purposes of selection efficiencies and distributions describing the performance in this Section, a looser selection has been used and is shown in Table 7.1.

Requirement	Preselection value
$\lvert M_{\mathrm{K^+K^-K^+K^-}} - M_{B_s^0}^{PDG}\rvert$ [MeV/c^2]	< 300
K IP χ^2	> 2.5
K p_{T} [MeV/c]	> 400
$\lvert M_{\mathrm{K^+K^-}} - M_{\phi}^{PDG}\rvert$ [MeV/c^2]	< 25
ϕ vertex χ^2 per NDF	< 15
ϕ^1 $p_{\mathrm{T}} \times \phi^2$ p_{T} [(GeV/c)^2]	> 2
B_s^0 vertex χ^2 per NDF	< 15
DOCA χ^2	< 40
$\cos\theta_{DIRA}$	> 0.99

Table 7.1.: Preselections applied to $B_s^0 \to \phi\phi$ simulated events to determine Upgrade efficiencies and performance, where IP and DOCA stand for impact parameter and distance of closest approach, respectively, and θ_{DIRA} is the angle between the B_s^0 momentum and the direction of the primary vertex to the decay vertex.

Table 7.2 shows the reconstruction efficiencies and the combined pre-selection and reconstruction efficiencies for the upgraded LHCb detector compared to the current configuration. Note that the effect of the trigger is not included since the trigger selections are not yet finalised. The upgraded LHCb detector is found to have significantly better reconstruction efficiency. The efficiency of the pre-selection has also been applied and is found to be better in the upgraded LHCb detector. The improved impact parameter resolution offered by the upgraded VELO is the main reason for the latter.

Detector configuration	Reconstruction Eff. (%)	Combined Eff. (%)
Upgraded LHCb detector, $\nu = 3.8$	22.53 ± 0.09	20.68 ± 0.09
Upgraded LHCb detector, $\nu = 7.6$	21.53 ± 0.09	19.42 ± 0.09
Current LHCb detector	17.35 ± 0.04	13.86 ± 0.03

Table 7.2.: Reconstruction efficiencies and reconstruction efficiencies combined with preselection efficiencies for $B_s^0 \to \phi\phi$ candidates in the case of the upgraded LHCb detector in two beam conditions, corresponding to $\nu = 3.8$ and 7.6, compared with the current LHCb detector configuration.

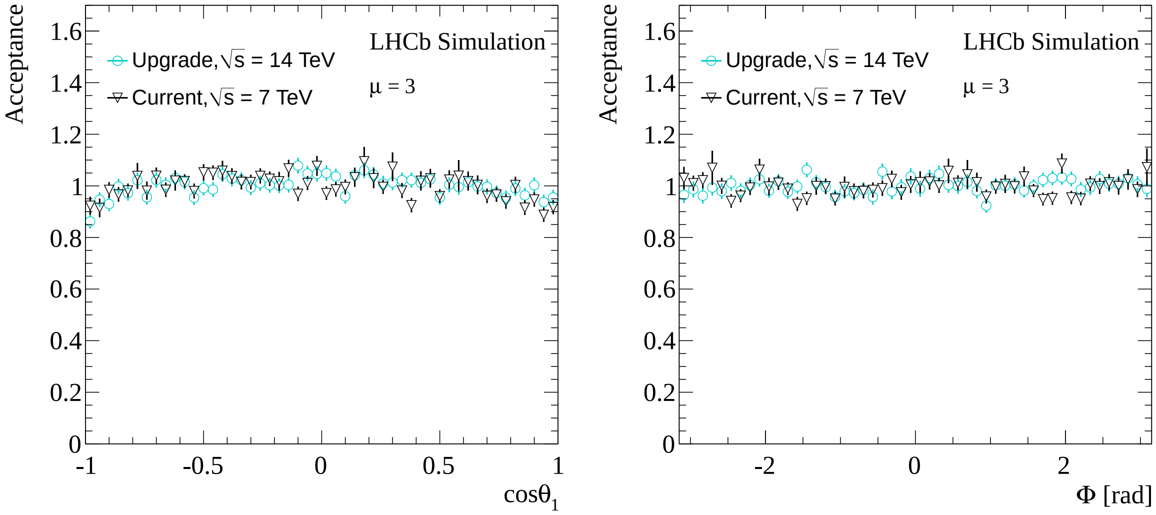


Figure 7.6.: Comparison between the upgraded and current LHCb detector of the efficiency as a function of the θ_1 (left) and Φ (right) helicity angles for $\mu = 3$.

7.2.2. Comparison of Acceptances

The detector geometry and selection requirements determine the variation of the efficiency as functions of the different decay angles. The angular acceptance of the different detector options is compared in Figure 7.6 for the θ_1 helicity angle, where a similar dependency is seen for the case of the current detector and the Upgrade configuration. The same comparison is made for the Φ helicity angle, where again a similar dependency can be seen.

A comparison of the decay time acceptance is presented in Figure 7.7. The reduced efficiencies at small decay times occur due to the requirements on the kaon impact parameter and related variables. The acceptance has a negative slope for large decay times in the case of the current LHCb detector. This is a known effect and is understood

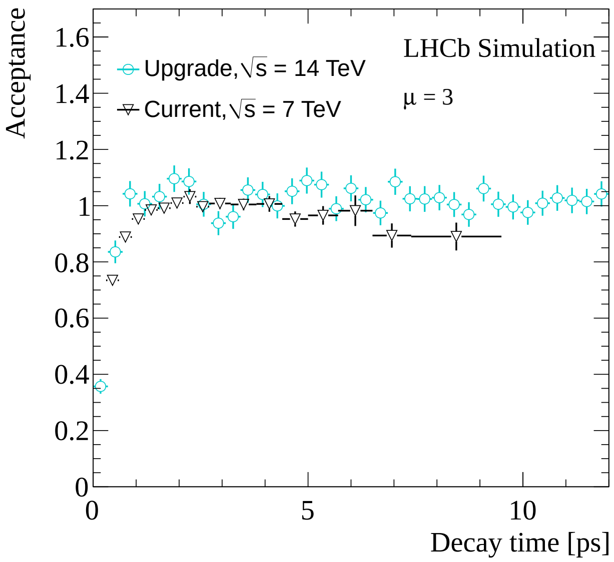


Figure 7.7.: Comparison between the upgraded and current LHCb detector of the efficiency as a function of the B_s^0 decay time for $\mu = 3$.

to be related to the geometry of the VELO. The decay time acceptance is observed to be flatter for the upgraded LHCb detector, although the statistical uncertainties become quite large for decay times greater than 5 ps. In addition, it can be seen that the decrease in efficiency at small decay times is less severe for the upgraded LHCb detector. This is expected due to the improved IP resolution in the upgraded VELO.

7.2.3. Comparison of Resolutions

The decay time is measured through $\tau = mL/p$ where m is the mass of the B_s^0 meson, L is the length the B_s^0 meson travelled before decaying and p is the B_s^0 momentum. The accuracy of the decay time is then determined through the accuracy of the decay vertex position and the accuracy of the momentum measurement. Figure 7.8 compares the difference between the true and reconstructed B_s^0 decay times, where an improvement can be seen for the case of the upgraded LHCb detector. Figures 7.9 and 7.10 compare the differences in the true and reconstructed values of the B_s^0 decay vertex position on the x-axis and the B_s^0 p_{T}, respectively.

In order to compare numerically the decay time resolution between the upgraded LHCb detector and the current detector, the histograms shown in Figure 7.8 have been fitted with the same double Gaussian model as used in Section 6.4. In addition, the same

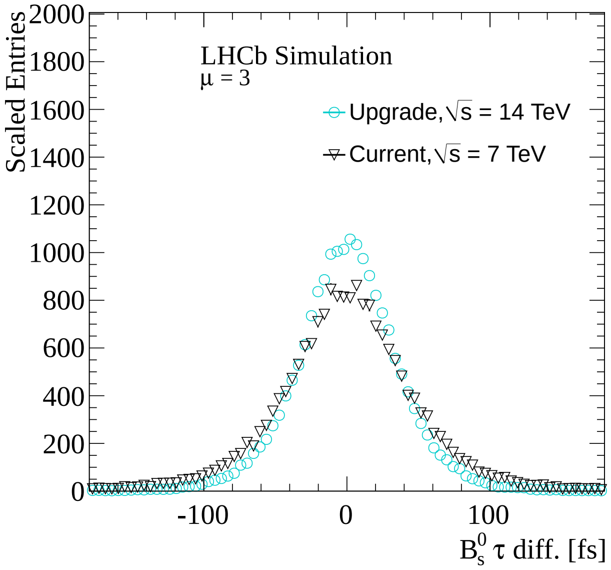


Figure 7.8.: Comparison between the upgraded and current LHCb detector of the B_s^0 decay time resolution for $\mu = 3$.

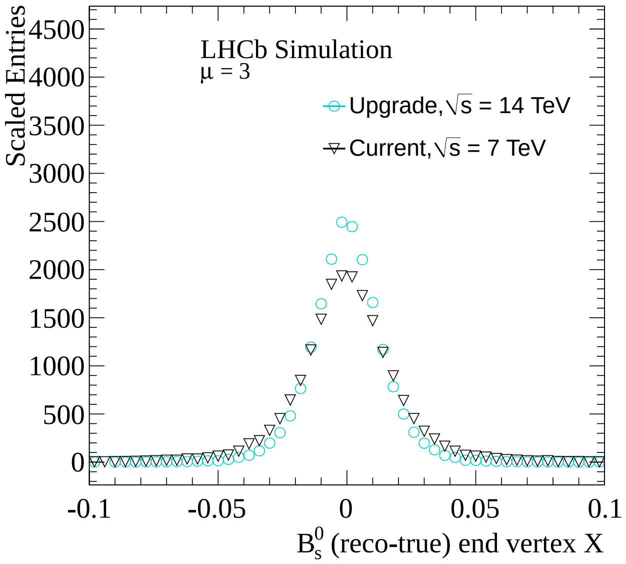


Figure 7.9.: Comparison between the upgraded and current LHCb detector of the B_s^0 decay vertex resolution in the x-axis for $\mu = 3$.

fits have been performed for the upgraded LHCb detector for $\mu = 4$ and 6. The results of the fits are shown in Table 7.3 along with the equivalent single Gaussian resolution. In all cases, the effective single Gaussian resolution is found to be significantly better in

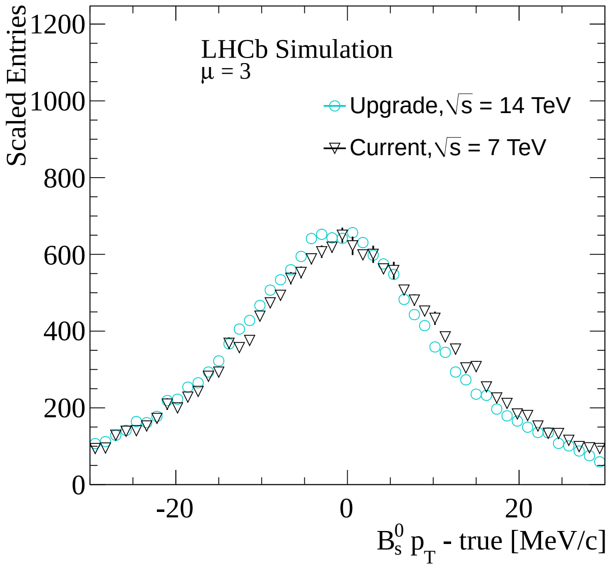


Figure 7.10.: Comparison between the upgraded and current LHCb detector of the B_s^0 p_T resolution for $\mu = 3$.

Detector configuration	σ_1(fs)	σ_2 (fs)	f	σ^eff (fs)
Upgraded LHCb detector, $\mu = 3$	29.3 ± 0.6	55.7 ± 1.9	0.76 ± 0.03	36.4
Upgraded LHCb detector, $\mu = 4$	29.4 ± 0.6	56.5 ± 1.9	0.76 ± 0.03	36.6
Upgraded LHCb detector, $\mu = 6$	29.9 ± 0.5	58.9 ± 2.0	0.78 ± 0.03	37.1
Current LHCb detector, $\mu = 3$	34.3 ± 0.9	69.6 ± 2.4	0.67 ± 0.03	46.8

Table 7.3.: Fit results to the double Gaussian decay time resolution model, where σ_1 and σ_2 denote the width of the first and second Gaussian functions, f denotes the fraction of the first Gaussian, and σ^eff denotes the effective single Gaussian resolution.

the case of the upgraded LHCb detector, though a small degradation in performance can be seen as the number of interactions per bunch crossing increases.

7.3. Expected Precision of Upgrade Measurements

The current LHCb detector has collected $3\,\mathrm{fb}^{-1}$ of data during 2011 and 2012, which has been termed Run 1. The data that will be collected from 2015 until the next long shutdown is termed Run 2, and is expected to add an additional $5\,\mathrm{fb}^{-1}$ of data. After the upgrade of the LHCb detector, it is anticipated that $5\,\mathrm{fb}^{-1}$ per year is added, which will result in $23\,\mathrm{fb}^{-1}$ of data collected by the end of Run 3 and $46\,\mathrm{fb}^{-1}$ of data collected by the end of Run 4 in 2028 [80, 81]. It is possible that LHCb will continue to run beyond 2028 until the ATLAS and CMS collaborations have each recorded $3000\,\mathrm{fb}^{-1}$. After such a run has finished, the LHCb experiment could collect as much as $100\,\mathrm{fb}^{-1}$ of data.

Figure 7.11 shows the expected sensitivity in terms of statistical precision of the CP-violating phase in $B_s^0 \to \phi\phi$ decays as a function of LHC run. The value of the uncertainty assigned to LHC Run 1 has been obtained with preliminary studies using all data collected during 2011 and 2012. It should be noted that the preliminary study of the complete LHC Run 1 dataset uses the pre-selection of Table 7.1, giving an increased selection efficiency with respect to the selections described in Chapter 4. The values of the uncertainty assigned for LHC runs 2-5 have been obtained from direct extrapolation of the Run 1 uncertainty, taking in to account the improved efficiencies of the LLT after the LHCb Upgrade has taken place [80]. It can be seen from Figure 7.11 that an upgraded LHCb experiment is capable of measuring the phase with greater precision than the current theoretical uncertainties.

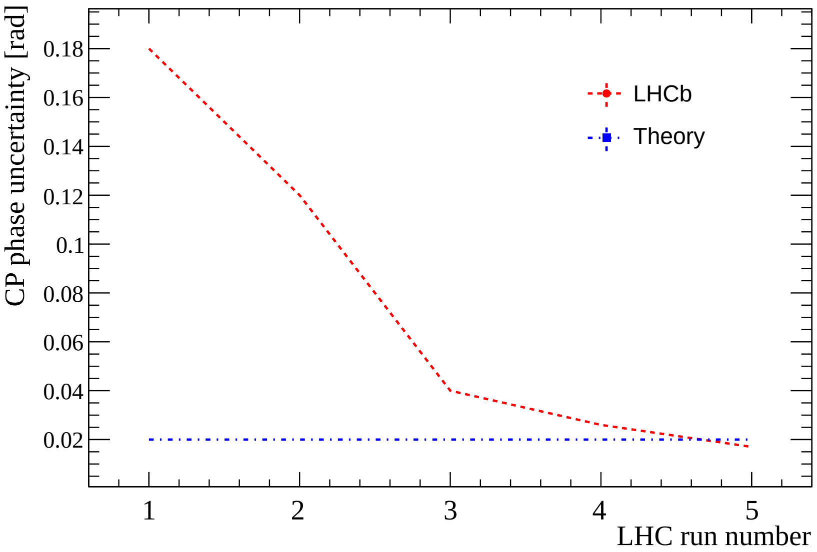


Figure 7.11.: Evolution of the experimental uncertainty on the *CP*-violating phase with LHC era. Shown for comparison is the expected theoretical uncertainty [22].

Chapter 8.

Summary, Conclusions and Outlook

"I hope that I may always desire more than I can accomplish."
— Michelangelo

The work presented in this dissertation is comprised of measurements of RICH subdetector performance during 2011 and 2012 data-taking, measurements of T violation and CP violation in the $B_s^0 \to \phi\phi$ decay with 2011 data, in addition to the studies of potential performance in the upgraded LHCb detector.

The effect seen during late 2011 and 2012 data-taking, in which central bands of high occupancy were seen in RICH HPD images, has been extensively studied. The aforementioned studies indicated certain hardware settings relating to the detection of the Cherekov photons on the silicon chips was sub-optimal. After modification of the settings, performance returned to that seen during 2011 operation.

The most accurate measurements of T violation and polarisation amplitudes in the $B_s^0 \to \phi\phi$ decay have been presented [60]. Results of studies of the polarisation amplitudes and strong phase difference were found to be

$$\begin{aligned} |A_0|^2 =& \ \ 0.365 \pm 0.022\,(\text{stat}) \pm 0.012\,(\text{syst})\,, \\ |A_\perp|^2 =& \ \ 0.291 \pm 0.024\,(\text{stat}) \pm 0.010\,(\text{syst})\,, \\ \cos\delta_\parallel =& -0.844 \pm 0.068\,(\text{stat}) \pm 0.029\,(\text{syst})\,, \end{aligned}$$

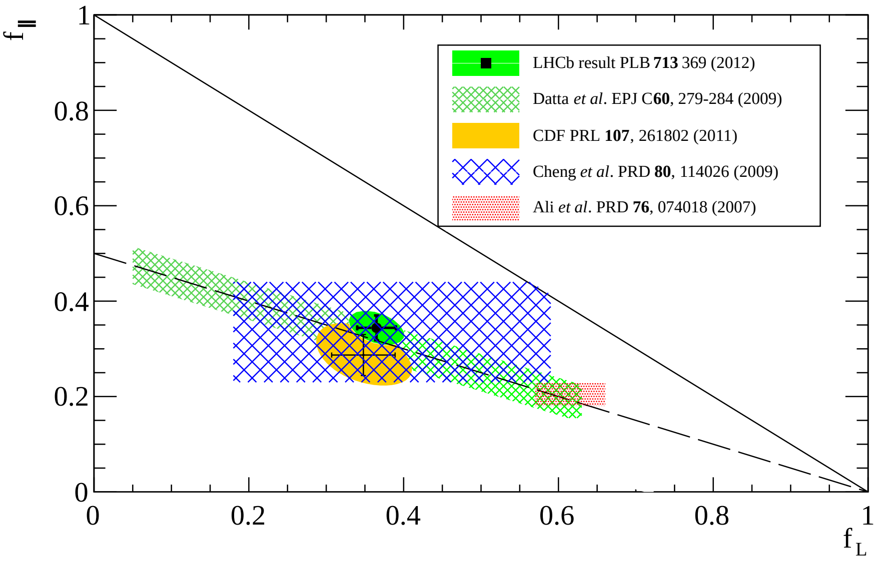


Figure 8.1.: Comparison of the LHCb polarisation fractions (solid green) [60] with the CDF measurement (solid yellow) [27], the perturbative QCD prediction of Ali *et al.* (medium hatched red) [21], the QCDF prediction of Datta *et al.* (small hatched green) [29], and the QCDF prediction of Cheng *et al.* (large hatched blue) [30].

which show good agreement with those measured by the CDF collaboration [27], though statistical uncertainties were reduced to almost half of the uncertainties reported by the CDF collaboration. Results were also found to be consistent with the predictions of QCD factorisation, though the uncertainties on the predictions of QCD factorisation are considerably larger than the experimental uncertainties. Figure 8.1 shows the polarisation fractions measured in the reported research compared to the values measured by the CDF collaboration, along with theoretical predictions.

The T-odd triple product asymmetries were measured to be

$$\begin{aligned} A_U &= -0.055 \pm 0.036\,(\text{stat}) \pm 0.018\,(\text{syst})\,, \\ A_V &= 0.010 \pm 0.036\,(\text{stat}) \pm 0.018\,(\text{syst})\,. \end{aligned}$$

Results are found to be consistent with those reported by the CDF collaboration, with significantly reduced uncertainties. No significant evidence of T violation was seen.

The first direct measurement of the CP-violating phase, $\phi^{s\bar{s}s\bar{s}}$, in $B_s^0 \rightarrow \phi\phi$ decays was performed and yielded a 68 % CL of

$$\phi^{s\bar{s}s\bar{s}}(68\,\%\ \text{CL}) = [-2.46, -0.76]. \tag{8.1}$$

The central value was therefore found to be far from the naive SM expectation, but with large uncertainties such that no evidence of *CP* violation can be claimed. The *p*-value of the SM hypothesis was found to be 16 %. The result was published in Physical Review Letters [65].

The measurement of the *CP*-violating phase was performed using $1.0\,\mathrm{fb}^{-1}$ of proton-proton collisions recorded with the LHCb experiment in 2011. An update of the analysis is underway, incorporating data taken in 2012, which will reduce the statistical uncertainties significantly.

The upgraded LHCb experiment can be expected to outperform the current detector, in almost every measure. Decay time resolution is expected to be significantly improved, along with reconstruction efficiencies and also trigger efficiencies. This is remarkable given the increasingly difficult environment provided by the upgrade luminosities.

The $50\,\mathrm{fb}^{-1}$ of data collected at the upgrade will provide the largest sample of $B_s^0 \to \phi\phi$ decays in existence and will allow $\phi^{s s\bar{s}}$ to be measured to a precision of below 0.02 rad, which is comparable to the theoretical uncertainties.

“After climbing a great hill, one only finds that there are many more hills to climb.”
— Nelson Mandela

Appendix A.

Detailed Angular Formulae Derivation

A.1. Form of the Amplitude

The general matrix element for a pseudo-scalar to vector vector ($P \to VV$) decay may be written as

$$\mathcal{M}_{fi} \sim \sum_{\lambda_1,\lambda_2} \langle \Omega_1, \phi_1, \lambda_1 | \langle \Omega_2, \phi_2, \lambda_2 | M_B | B_s^0 \rangle \langle \Omega_1^k | M_{\phi_1} | \phi_1, \lambda_1 \rangle \langle \Omega_2^k | M_{\phi_2} | \phi_2, \lambda_2 \rangle, \tag{A.1}$$

where λ_i and Ω_i describe the spins and phase space of the i^{th} ϕ meson, respectively, and $|\lambda_i| \leq 1$. The decay angles of the K^+K^- pairs resulting from the decay of the i^{th} ϕ meson are denoted Ω_i^k. From conservation of spin $\lambda_1 = \lambda_2$. Since the $\phi \to K^+K^-$ decay is a spin $1 \to 00$ transition, the polarisations of the two ϕ mesons will be correlated. The matrix element can then be written as

$$\mathcal{M}_{fi} \sim \sum_{i \in \{-1,0,1\}} \alpha_i^{eff} D^{*1}_{i,0}(\Omega_1) D^{*1}_{i,0}(\Omega_2), \tag{A.2}$$

where $\alpha_i^{eff} = \alpha^0_{i\,i}(B_s^0)\alpha^1_{0\,0}(\phi_1)\alpha^1_{0\,0}(\phi_2)$. Since the individual $\alpha^{\;j}_{\lambda_1\,\lambda_2}$ terms may not be experimentally distinguishable, it is the effective terms that can be related to measurable observables. The α parameters are defined through

$$\alpha^{\;j}_{\lambda_1\,\lambda_2} = \sqrt{\frac{4\pi(2J+1)}{\rho_s}} \langle J, M, \lambda_1, \lambda_2 | \hat{M} | J, M \rangle, \tag{A.3}$$

$$\Rightarrow \alpha^{\;j}_{\lambda_1\,\lambda_2} \propto \sqrt{2J+1},$$

where J and M are quantum numbers describing the total angular momentum and the projection onto the z-axis, respectively. The D functions describe angular distributions, while the α terms contain the decay time dependence and physics involved in the decay. The D functions may be related to spherical harmonics, $Y_l^m(\theta,\Phi)$, through

$$D^{*\,l}_{m\,0} = \sqrt{\frac{4\pi}{(2l+1)}}Y_l^m(\theta,\Phi). \tag{A.4}$$

The D functions that appear in the matrix element are then evaluated as

$$D^{*\,1}_{-1\,0}(\Omega_1)D^{*\,1}_{-1\,0}(\Omega_2) = \frac{1}{2}\sin\theta_1\sin\theta_2 e^{-i\Phi}, \tag{A.5}$$

$$D^{*\,1}_{0\,0}(\Omega_1)D^{*\,1}_{-1\,0}(\Omega_2) = \cos\theta_1\cos\theta_2, \tag{A.6}$$

$$D^{*\,1}_{+1\,0}(\Omega_1)D^{*\,1}_{-1\,0}(\Omega_2) = \frac{1}{2}\sin\theta_1\sin\theta_2, e^{i\Phi}, \tag{A.7}$$

where θ_1, θ_2 and Φ are the helicity angles shown in Figure 1.7 and the coordinate system has been chosen such that $\Phi_1 = 0 \Rightarrow \Phi_1 + \Phi_2 = \Delta\Phi \equiv \Phi$. Considering the real and imaginary parts of the D functions separately, we can write the amplitude for the P-wave terms ($A^{P-\text{wave}}$)as

$$\begin{aligned} A^{P-\text{wave}} = {} & 3\frac{(F_+^{eff}+F_-^{eff})}{\sqrt{2}}\sin\theta_1\sin\theta_2\cos\Phi \\ & +i3\frac{(F_+^{eff}-F_-^{eff})}{\sqrt{2}}\sin\theta_1\sin\theta_2\sin\Phi \\ & +3F_0^{eff}\cos\theta_1\cos\theta_2, \end{aligned} \tag{A.8}$$

where $F^{eff}_{(-,0,+)}$ are helicity amplitudes. In this formalism, we can add S-wave terms through observation of the spin structure of the decay and consider the cases individually, i.e P-wave: $0 \to 1(\to 00) + 1(\to 00)$;
S-wave: $0 \to 1(\to 00) + 0(\to 00)$ and $0 \to 0(\to 00) + 0(\to 00)$. In the rest of this Chapter, f_0 will be used to denote spin-0 K^+K^- pairs. For the case of $B_s^0 \to \phi f_0$, spin

conservation forces $\lambda_1 = \lambda_2 = 0$. The matrix element may therefore be written as

$$\begin{aligned}\mathcal{M}_{fi}^{S-\text{wave}(\phi f_0)} \sim\; & \langle\Omega_1, \phi_1, (\lambda_1 = 0)|\langle\Omega_2, f_0|M_B|B_s^0\rangle\langle\Omega_1^k|M_{\phi_1}|\phi_1, 0\rangle\langle\Omega_2^k|M_{f_0}|f_0, 0\rangle \\ & +\langle\Omega_1, f_0|\langle\Omega_2, \phi_2, (\lambda_2 = 0)|M_B|B_s^0\rangle\langle\Omega_1^k|M_{f_0}|f_0, 0\rangle\langle\Omega_2^k|M_{\phi_2}|\phi_2, 0\rangle,\end{aligned} \tag{A.9}$$

$$= \alpha^0_{0\,0}(B_s^0)\alpha^1_{0\,0}(\phi_1)\alpha^0_{0\,0}(f_0)D^{*\,1}_{0,0}(\Omega_1)D^{*\,0}_{0,0}(\Omega_2) + \alpha^0_{0\,0}(B_s^0)\alpha^0_{0\,0}(f_0)\alpha^1_{0\,0}(\phi_2)D^{*\,0}_{0,0}(\Omega_1)D^{*\,1}_{0,0}(\Omega_2). \tag{A.10}$$

Similarly, the equivalent expression for $B_s^0 \to f_0 f_0$ may be written as

$$\mathcal{M}_{fi}^{S-\text{wave}(f_0 f_0)} \sim \alpha^0_{0\,0}(B_s^0)\alpha^0_{0\,0}(f_0)\alpha^0_{0\,0}(f_0)D^{*\,0}_{0,0}(\Omega_1)D^{*\,0}_{0,0}(\Omega_2). \tag{A.11}$$

The two extra D functions evaluate to

$$D^{*\,0}_{0\,0}(\Omega_1)D^{*\,0}_{0\,0}(\Omega_2) = \cos\theta_1 + \cos\theta_2, \tag{A.12}$$

$$D^{*\,0}_{0\,0}(\Omega_1)D^{*\,1}_{0\,0}(\Omega_2) = 1, \tag{A.13}$$

for the case of the $B_s^0 \to \phi f_0$ (denoted with a subscript S) and $B_s^0 \to f_0 f_0$ (denoted with a subscript SS) terms respectively. This leads to S-wave terms in the amplitude ($A^{S-\text{wave}}$) of the form

$$A^{S-\text{wave}} = \sqrt{3}F_S^{eff}(\cos\theta_1 + \cos\theta_2) + F_{SS}^{eff}. \tag{A.14}$$

The well known transformation from the helicity basis to the linear amplitude basis is given by

$$A_0(t) = \alpha_0^{eff} = 3F_0^{eff}, \tag{A.15}$$

$$A_\parallel(t) = \frac{\alpha_+^{eff} + \alpha_-^{eff}}{\sqrt{2}} = 3\frac{F_+^{eff} + F_-^{eff}}{\sqrt{2}}, \tag{A.16}$$

$$A_\perp(t) = \frac{\alpha_+^{eff} - \alpha_-^{eff}}{\sqrt{2}} = 3\frac{F_+^{eff} - F_-^{eff}}{\sqrt{2}}. \tag{A.17}$$

In order to keep consistent normalization the S-wave polarization amplitudes (A_S & A_{SS}) are defined as

$$A_S(t) = 3F_S^{eff}, \tag{A.18}$$

$$A_{SS}(t) = 3F_{SS}^{eff}. \tag{A.19}$$

This then allows the final amplitude ($\mathcal{A}$) to be written as

$$\mathcal{A}(t,\theta_1,\theta_2,\Phi) = A_0(t)\cos\theta_1\cos\theta_2 + \frac{A_\parallel(t)}{\sqrt{2}}\sin\theta_1\sin\theta_2\cos\Phi \\ + i\frac{A_\perp(t)}{\sqrt{2}}\sin\theta_1\sin\theta_2\sin\Phi + \frac{A_S(t)}{\sqrt{3}}(\cos\theta_1+\cos\theta_2) + \frac{A_{SS}(t)}{3}. \quad \text{(A.20)}$$

The differential decay rate may be found through the square of the total amplitude leading to the fifteen terms

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t\mathrm{d}\cos\theta_1\mathrm{d}\cos\theta_2\mathrm{d}\Phi} \propto 4|\mathcal{A}(t,\theta_1,\theta_2,\Phi)|^2 = \sum_{i=1}^{15} K_i(t)f_i(\theta_1,\theta_2,\Phi), \quad \text{(A.21)}$$

where the $K_i(t)$ & $f_i(\theta_1,\theta_2,\Phi)$[1] are shown in Table A.1.

i	K_i	f_i
1	$\lvert A_0(t)\rvert^2$	$4\cos^2\theta_1\cos^2\theta_2$
2	$\lvert A_\parallel(t)\rvert^2$	$\sin^2\theta_1\sin^2\theta_2(1+\cos 2\Phi)$
3	$\lvert A_\perp(t)\rvert^2$	$\sin^2\theta_1\sin^2\theta_2(1-\cos 2\Phi)$
4	$\Im(A_\parallel^*(t)A_\perp(t))$	$-2\sin^2\theta_1\sin^2\theta_2\sin 2\Phi$
5	$\Re(A_\parallel^*(t)A_0(t))$	$\sqrt{2}\sin 2\theta_1\sin 2\theta_2\cos\Phi$
6	$\Im(A_0^*(t)A_\perp(t))$	$-\sqrt{2}\sin 2\theta_1\sin 2\theta_2\sin\Phi$
7	$\lvert A_{SS}(t)\rvert^2$	$\frac{4}{9}$
8	$\lvert A_S(t)\rvert^2$	$\frac{4}{3}(\cos\theta_1+\cos\theta_2)^2$
9	$\Re(A_S^*(t)A_{SS}(t))$	$\frac{8}{3\sqrt{3}}(\cos\theta_1+\cos\theta_2)$
10	$\Re(A_0(t)A_{SS}^*(t))$	$\frac{8}{3}\cos\theta_1\cos\theta_2$
11	$\Re(A_\parallel(t)A_{SS}^*(t))$	$\frac{4\sqrt{2}}{3}\sin\theta_1\sin\theta_2\cos\Phi$
12	$\Im(A_\perp(t)A_{SS}^*(t))$	$-\frac{4\sqrt{2}}{3}\sin\theta_1\sin\theta_2\sin\Phi$
13	$\Re(A_0(t)A_S^*(t))$	$\frac{8}{\sqrt{3}}\cos\theta_1\cos\theta_2(\cos\theta_1+\cos\theta_2)$
14	$\Re(A_\parallel(t)A_S^*(t))$	$\frac{4\sqrt{2}}{\sqrt{3}}\sin\theta_1\sin\theta_2(\cos\theta_1+\cos\theta_2)\cos\Phi$
15	$\Im(A_\perp(t)A_S^*(t))$	$-\frac{4\sqrt{2}}{\sqrt{3}}\sin\theta_1\sin\theta_2(\cos\theta_1+\cos\theta_2)\sin\Phi$

Table A.1.: Terms found in the total differential decay rate.

---

[1] The factor of 4 in equation A.21 is included as a matter of convenience and is absorbed by the f_i terms. This makes no difference to fitting as the f_i terms are normalised to 1 on integration over the helicity angles.

A.2. Time Evolution

The $A_0(t)$, $A_\parallel(t)$ and $A_\perp(t)$ terms encode all of the physics of B^0_s mixing. The time evolution of B^0_s mesons is described by

$$|B^0_s(t)\rangle = g_+(t)|B^0_s(0)\rangle + \frac{q}{p}g_-(t)|\overline{B}^0_s(0)\rangle, \tag{A.22}$$

$$|\overline{B}^0_s(t)\rangle = \frac{p}{q}g_-(t)|B^0_s(0)\rangle + g_+(t)|\overline{B}^0_s(0)\rangle, \tag{A.23}$$

where

$$g_+(t) = \frac{1}{2}\left(e^{-(im_L+\Gamma_L/2)t} + e^{-(im_H+\Gamma_H/2)t}\right), \tag{A.24}$$

$$g_-(t) = \frac{1}{2}\left(e^{-(im_L+\Gamma_L/2)t} - e^{-(im_H+\Gamma_H/2)t}\right). \tag{A.25}$$

The masses and decay rates of the B^0_s mass eigenstates have been denoted by $m_{L(H)}$ and $\Gamma_{L(H)}$, respectively, with a subscript L for the lighter eigenstate and H for the heavier eigenstate. Therefore, the time dependence of each polarization amplitude is given by

$$A_h(t) \equiv \langle f|B^0_s(t)\rangle_h = [g_+(t)A_h + \eta_h\frac{q}{p}g_-(t)\overline{A}_h], \tag{A.26}$$

$$\overline{A}_h(t) \equiv \langle \overline{f}|\overline{B}^0_s(t)\rangle_h = [\frac{p}{q}g_+(t)A_h + \eta_h\frac{q}{p}g_+(t)\overline{A}_h], \tag{A.27}$$

where $h \in \{0, \parallel, \perp, S, SS\}$, $A_h \equiv A_h(0) = \langle f|B^0_s\rangle_h$ and η_h is the *CP* eigenvalue of the final state. The explicit form of the $K_i(t)$ functions may be derived noting two useful formulae

$$|g_\pm(t)|^2 = \frac{1}{2}e^{-\Gamma_s t}(\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) \pm \cos\Delta m_s t), \tag{A.28}$$

$$g^*_+(t)g_-(t) = \frac{1}{2}e^{-\Gamma_s t}(-\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + i\sin\Delta m_s t), \tag{A.29}$$

where $\Delta m_s = m_H - m_L$, $\Gamma_s = (\Gamma_\mathrm{L} + \Gamma_\mathrm{H})/2$ and $\Delta\Gamma_s = \Gamma_L - \Gamma_H$ are the mass and lifetime difference of the B^0_s mass eigenstates respectively. The individual $K_i(t)$ terms may be found in a slightly less tedious method, through the observation that the time evolution of each term can be categorised and derivations recycled, i.e. the forms of single amplitude type terms are similar as are those of the interference terms. A generalised derivation yields for single amplitude type terms

$$\begin{aligned}
|A_h(t)|^2 &= [g_+(t)A_h + \eta_h \frac{q}{p} g_-(t)\overline{A}_h][g_+^*(t)A_h^* + \eta_h (\frac{q}{p})^* g_-^*(t)\overline{A}_h^*] \\
&= \frac{|A_h|^2}{2} e^{\Gamma_s t}(\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + \cos\Delta m_s t) + \eta_h^2 |\frac{q}{p}|^2 \frac{|\overline{A}_h|^2}{2} e^{\Gamma_s t}(\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - \cos\Delta m_s t) \\
&+ \eta_h \left(\frac{q}{p}\right)^* \frac{A_0\overline{A}_0^*}{2} e^{-\Gamma_s t}(-\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - i\sin\Delta m_s t) \\
&+ \eta_h \left(\frac{q}{p}\right) \frac{A_0^*\overline{A}_0}{2} e^{-\Gamma_s t}(-\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + i\sin\Delta m_s t).
\end{aligned} \tag{A.30}$$

On using the assumptions $|\frac{q}{p}| = 1$ and $\frac{q}{p} = e^{-i\phi_s}$, equation A.30 simplifies to

$$\begin{aligned}
|A_h(t)|^2 = \frac{e^{-\Gamma_s t}}{2}\{&(|A_h|^2 + \eta_h^2|\overline{A}_h|^2)\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + (|A_h|^2 - \eta_h^2|\overline{A}_h|^2)\cos(\Delta m_s t) \\
&- 2\eta_h \Re(A_0^*\overline{A}_0)(\cos\phi_s \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - \sin\phi_s \sin(\Delta m_s t)) \\
&- 2\eta_h \Im(A_0^*\overline{A}_0)(\cos\phi_s \sin(\Delta m_s t) + \sin\phi_s \sinh\left(\frac{\Delta\Gamma_s t}{2}\right))\}.
\end{aligned} \tag{A.31}$$

Two other possibilities remain. These are the real and imaginary components of the interference between two different types of amplitudes.

$$\begin{aligned}
A_h(t)A_k^*(t) &= [g_+(t)A_h + \eta_h \frac{q}{p} g_-(t)\overline{A}_h][g_+^*(t)A_k^* + \eta_k (\frac{q}{p})^* g_-^*(t)\overline{A}_k^*] \\
&= \frac{A_hA_k^*}{2} e^{\Gamma_s t}(\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + \cos\Delta m_s t) + \eta_h\eta_k |\frac{q}{p}|^2 \frac{\overline{A}_h\overline{A}_k^*}{2} e^{\Gamma_s t}(\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - \cos\Delta m_s t) \\
&+ \eta_k \frac{A_h\overline{A}_k^*}{2}(\cos\phi_s + i\sin\phi_s) e^{-\Gamma_s t}(-\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - i\sin\Delta m_s t) \\
&+ \eta_h \frac{A_k^*\overline{A}_h}{2}(\cos\phi_s - i\sin\phi_s) e^{-\Gamma_s t}(-\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + i\sin\Delta m_s t).
\end{aligned} \tag{A.32}$$

Taking the real part of equation A.32 yields

$$\begin{aligned}\Re(A_h(t)A_k^*(t)) = \frac{e^{-\Gamma_s t}}{2}\{&(\Re(A_h A_k^*) + \eta_h\eta_k\Re(\overline{A}_h\overline{A}_k^*))\cosh\left(\frac{\Delta\Gamma_s t}{2}\right)\\ &+ (\Re(A_h A_k^*) - \eta_h\eta_k\Re(\overline{A}_h\overline{A}_k^*))\cos(\Delta m_s t)\\ &+ (\eta_k\Re(A_h\overline{A}_k^*) + \eta_h\Re(\overline{A}_h A_k^*))(\sin\phi_s\sin(\Delta m_s t) - \cos\phi_s\sinh\left(\frac{\Delta\Gamma_s t}{2}\right))\\ &+ (\eta_k\Im(A_h\overline{A}_k^*) - \eta_h\Im(\overline{A}_h A_k^*))(\cos\phi_s\sin(\Delta m_s t) + \sin\phi_s\sinh\left(\frac{\Delta\Gamma_s t}{2}\right))\}.\end{aligned} \tag{A.33}$$

The imaginary part in a similar way yields

$$\begin{aligned}\Im(A_h(t)A_k^*(t)) = \frac{e^{-\Gamma_s t}}{2}\{&(\Im(A_h A_k^*) + \eta_h\eta_k\Im(\overline{A}_h\overline{A}_k^*))\cosh\left(\frac{\Delta\Gamma_s t}{2}\right)\\ &+ (\Im(A_h A_k^*) - \eta_h\eta_k\Im(\overline{A}_h\overline{A}_k^*))\cos(\Delta m_s t)\\ &+ (\eta_k\Im(A_h\overline{A}_k^*) + \eta_h\Im(\overline{A}_h A_k^*))(\sin\phi_s\sin(\Delta m_s t) - \cos\phi_s\sinh\left(\frac{\Delta\Gamma_s t}{2}\right))\\ &+ (\eta_h\Re(\overline{A}_h A_k^*) - \eta_k\Re(A_h\overline{A}_k^*))(\cos\phi_s\sin(\Delta m_s t) + \sin\phi_s\sinh\left(\frac{\Delta\Gamma_s t}{2}\right))\}.\end{aligned} \tag{A.34}$$

It can be assumed that the complex amplitudes take the form $A_h = |A_h|e^{i\delta_h}e^{i\phi_h}$ where δ_h is a *CP*-conserving phase and ϕ_h is a *CP*-violating phase. Equations A.31, A.33 and A.34 are the starting points for the evaluation of all time-dependent functions $K_i(t)$. A simplified expression for each of the time dependent functions may be obtained from the assumption that no *CP* violation is found in the decay, i.e. $A_h = |A_h|e^{i\delta_h}$. On noting the two useful formulae

$$e^{-\Gamma_s t}\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) = \frac{1}{2}(e^{-\Gamma_\mathrm{H} t} + e^{-\Gamma_\mathrm{L} t}), \tag{A.35}$$

$$e^{-\Gamma_s t}\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) = \frac{1}{2}(e^{-\Gamma_\mathrm{H} t} - e^{-\Gamma_\mathrm{L} t}), \tag{A.36}$$

$$\tag{A.37}$$

the explicit form of each $K_i(t)$ function can be found. For $|A_0(t)|^2$

$$|A_0(t)|^2 = \frac{|A_0|^2}{2}[(1+\cos\phi_s)e^{-\Gamma_{\rm L}t} + (1-\cos\phi_s)e^{-\Gamma_{\rm H}t} \\ \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s], \tag{A.38}$$

where the upper (lower) of the ± or ∓ signs is for the case of the B_s^0 ($\overline{B}_s^0$) decay[2]. Similarly for $|A_\parallel(t)|^2$

$$|A_\parallel(t)|^2 = \frac{|A_\parallel|^2}{2}[(1+\cos\phi_s)e^{-\Gamma_{\rm L}t} + (1-\cos\phi_s)e^{-\Gamma_{\rm H}t} \\ \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s]. \tag{A.39}$$

The *CP*-odd $|A_\perp(t)|^2$ term evaluates to

$$|A_\perp(t)|^2 = \frac{|A_\parallel|^2}{2}[(1-\cos\phi_s)e^{-\Gamma_{\rm L}t} + (1+\cos\phi_s)e^{-\Gamma_{\rm H}t} \\ \mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s]. \tag{A.40}$$

The interference term $\Im(A_\parallel(t)^* A_\perp(t))$ is found through equation A.34 to be

$$\Im(A_\parallel(t)^* A_\perp(t)) = |A_\parallel||A_\perp|\{\pm e^{-\Gamma_s t}[\sin\delta_1\cos(\Delta m_s t) - \cos\delta_1\sin(\Delta m_s t)\cos\phi_s] \\ -\frac{1}{2}\cos\delta_1(e^{-\Gamma_{\rm H}t} - e^{-\Gamma_{\rm L}t})\sin\phi_s\}, \tag{A.41}$$

where $\delta_1 = \delta_\perp - \delta_\parallel$. The form of $\Re(A_\parallel(t)^* A_0(t))$ is the same as that of equation A.2, with only the prefactor changing, i.e.

$$\Re(A_\parallel(t)^* A_0(t)) = \frac{1}{2}|A_\parallel||A_0|\cos(\delta_2-\delta_1)[(1+\cos\phi_s)e^{-\Gamma_{\rm L}t} + (1-\cos\phi_s)e^{-\Gamma_{\rm H}t} \\ \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s], \tag{A.42}$$

where $\delta_2 = \delta_\perp - \delta_0$. The interference term $\Im(A_0(t)^* A_\perp(t))$ is of the same form as that of equation A.41, with prefactor and strong phase changes, i.e.

$$\Im(A_0(t)^* A_\perp(t)) = |A_0||A_\perp|\{\pm e^{-\Gamma_s t}[\sin\delta_2\cos(\Delta m_s t) - \cos\delta_2\sin(\Delta m_s t)\cos\phi_s] \\ -\frac{1}{2}\cos\delta_2(e^{-\Gamma_{\rm H}t} - e^{-\Gamma_{\rm L}t})\sin\phi_s\}. \tag{A.43}$$

---

[2] The expression for $\overline{B}_s^0$ decays may be found by applying the transformations $\phi_s \to -\phi_s$ and $A_h \leftrightarrow \eta_h \overline{A}_h$ to equations A.31, A.33 and A.34.

The *CP*-odd S-wave polarization $|A_S(t)|^2$ term is evaluated as

$$\begin{aligned}|A_S(t)|^2 = \frac{|A_S|^2}{2}[&(1-\cos\phi_s)e^{-\Gamma_{\mathrm{L}}t} + (1+\cos\phi_s)e^{-\Gamma_{\mathrm{H}}t} \\ &\mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s],\end{aligned} \tag{A.44}$$

and the *CP*-even S-wave polarization $|A_{SS}(t)|^2$ term is evaluated as

$$\begin{aligned}|A_{SS}(t)|^2 = \frac{|A_{SS}|^2}{2}[&(1+\cos\phi_s)e^{-\Gamma_{\mathrm{L}}t} + (1-\cos\phi_s)e^{-\Gamma_{\mathrm{H}}t} \\ &\pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s].\end{aligned} \tag{A.45}$$

The interference between the two types of S-wave term yields a time-dependent term

$$\begin{aligned}\Re(A_S(t)^* A_{SS}(t)) = |A_S||A_{SS}|\{&\pm e^{-\Gamma_s t}[\cos(\delta_{SS}-\delta_S)\cos(\Delta m_s t) - \sin(\delta_{SS}-\delta_S)\sin(\Delta m_s t)\cos\phi_s] \\ &-\frac{1}{2}\sin(\delta_{SS}-\delta_S)(e^{-\Gamma_{\mathrm{H}}t} - e^{-\Gamma_{\mathrm{L}}t})\sin\phi_s\}.\end{aligned} \tag{A.46}$$

The interference between the *CP*-even S-wave and the P-wave takes the same form as $K_5(t)$ with only modifications of the prefactor required, i.e.

$$\begin{aligned}\Re(A_0(t)A_{SS}(t)^*) &= \frac{1}{2}|A_0||A_{SS}|\cos(\delta_{SS})[(1+\cos\phi_s)e^{-\Gamma_{\mathrm{L}}t} + (1-\cos\phi_s)e^{-\Gamma_{\mathrm{H}}t} \\ &\quad \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s], \qquad \text{(A.47)} \\ \Re(A_\parallel(t)A_{SS}(t)^*) &= \frac{1}{2}|A_\parallel||A_{SS}|\cos(\delta_2-\delta_1-\delta_{SS})[(1+\cos\phi_s)e^{-\Gamma_{\mathrm{L}}t} + (1-\cos\phi_s)e^{-\Gamma_{\mathrm{H}}t} \\ &\quad \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s]. \qquad \text{(A.48)}\end{aligned}$$

The interference between the *CP*-even S-wave and the *CP*-odd P-wave yields a term of the same form as equation A.41, i.e.

$$\begin{aligned}\Im(A_\perp(t)A_{SS}(t)^*) = |A_\perp||A_{SS}|\{&\pm e^{-\Gamma_s t}[\sin(\delta_2-\delta_{SS})\cos(\Delta m_s t) - \cos(\delta_2-\delta_{SS})\sin(\Delta m_s t)\cos\phi_s] \\ &-\frac{1}{2}\cos(\delta_2-\delta_{SS})(e^{-\Gamma_{\mathrm{H}}t} - e^{-\Gamma_{\mathrm{L}}t})\sin\phi_s\}.\end{aligned} \tag{A.49}$$

The interference between the *CP*-odd *S*-wave and the *CP*-even *P*-wave terms is of the same form as found in equation A.46. These equations can then be written as

$$\begin{aligned}\Re(A_0(t)A_S(t)^*) = |A_S||A_0|\{&\pm e^{-\Gamma_s t}[\cos\delta_S \cos(\Delta m_s t) + \sin\delta_S \sin(\Delta m_s t)\cos\phi_s] \\ &+ \frac{1}{2}\sin\delta_S(e^{-\Gamma_\mathrm{H} t} - e^{-\Gamma_\mathrm{L} t})\sin\phi_s\},\end{aligned} \tag{A.50}$$

$$\begin{aligned}\Re(A_\parallel(t)A_S(t)^*) = |A_S||A_\parallel|\{&\pm e^{-\Gamma_s t}[\cos(\delta_2-\delta_1-\delta_S)\cos(\Delta m_s t) - \sin(\delta_2-\delta_1-\delta_S)\sin(\Delta m_s t)\cos\phi_s] \\ &- \frac{1}{2}\sin(\delta_2-\delta_1-\delta_S)(e^{-\Gamma_\mathrm{H} t} - e^{-\Gamma_\mathrm{L} t})\sin\phi_s\}.\end{aligned} \tag{A.51}$$

The final term is that of the interference between the *CP*-odd *S*-wave and the *CP*-odd *P*-wave. This is evaluated with the assistance of equation A.34 to be

$$\begin{aligned}\Im(A_\perp(t)A_S(t)^*) = \frac{1}{2}|A_\perp||A_S|\sin(\delta_2-\delta_S)[&(1-\cos\phi_s)e^{-\Gamma_\mathrm{L} t} + (1+\cos\phi_s)e^{-\Gamma_\mathrm{H} t} \\ &\mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s].\end{aligned} \tag{A.52}$$

Appendix B.

T-odd Triple Products in $P \to VV$ Decays

Scalar triple products of three momentum or spin vectors are odd under time reversal. Observation of non-zero triple product asymmetries can either be due to the presence of *CP*-violating phases or a *CP*-conserving phase in association with final-state interactions. T-odd triple products in $B_s^0 \to \phi\phi$ decays are defined through [28]

$$\sin\Phi = (\hat{n}_1 \times \hat{n}_2) \cdot \hat{p}_1, \tag{B.1}$$

$$\sin(2\Phi)/2 = (\hat{n}_1 \cdot \hat{n}_2)(\hat{n}_1 \times \hat{n}_2) \cdot \hat{p}_1, \tag{B.2}$$

where $\hat{n}_i$ $(i = 1, 2)$ is a unit vector perpendicular to the ϕ_i decay plane, $\hat{p}_1$ is a unit vector in the direction of the ϕ_1 momentum in the B_s^0 rest frame, and Φ is the angle between the two ϕ decay planes in the B_s^0 rest frame, shown in Figure 1.7.

The U and V observables introduced in Section 1.4.2.2, and used to measure the triple product asymmetries in Section 5 correspond to the T-odd triple products defined in equations B.1 and B.2.

From the full decay time dependent terms, shown in Table A.1, it can be seen that after integrating over B_s^0 decay time and the $\cos\theta_{1,2}$ helicity angles, the resulting differential decay distribution is of the form [28]

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}\Phi} \propto |A_0^{\mathrm{TI}}|^2 + 2|A_\perp^{\mathrm{TI}}|^2 \sin^2\Phi + 2|A_\parallel^{\mathrm{TI}}|^2 \cos^2\Phi - 2\Im(A_\perp A_\parallel^*)^{\mathrm{TI}} \sin 2\Phi, \tag{B.3}$$

where a superscript TI denotes a decay time integrated amplitude. If an asymmetry, $A_U(B_s^0)$ is constructed based on the $\sin 2\Phi$ triple product of equation B.2, this then corresponds to

$$A_U(B_s^0) = \frac{\int_0^{\pi/2}(\mathrm{d}\Gamma/\mathrm{d}\Phi)\mathrm{d}\Phi + \int_{\pi}^{3\pi/2}(\mathrm{d}\Gamma/\mathrm{d}\Phi)\mathrm{d}\Phi - \int_{3\pi/2}^{2\pi}(\mathrm{d}\Gamma/\mathrm{d}\Phi)\mathrm{d}\Phi - \int_{\pi/2}^{\pi}(\mathrm{d}\Gamma/\mathrm{d}\Phi)\mathrm{d}\Phi}{\int_0^{2\pi}(\mathrm{d}\Gamma/\mathrm{d}\Phi)\mathrm{d}\Phi}. \tag{B.4}$$

On the insertion of equation B.3, this evaluates to

$$A_U(B_s^0) = -\frac{4}{\pi}\frac{\Im(A_\perp A_\parallel^*)^{\mathrm{TI}}}{|A_0^{\mathrm{TI}}|^2 + |A_\parallel^{\mathrm{TI}}|^2 + |A_\perp^{\mathrm{TI}}|^2}. \tag{B.5}$$

It can be seen from Table A.1 that another interference term is present, namely $\Im(A_\perp A_0^*)^{\mathrm{TI}}$. An asymmetry, $A_V(B_s^0)$, may be constructed based on the $\sin\Phi$ triple product of equation B.1, assigning a factor of -1 for regions in which $\cos\theta_1\cos\theta_2 < 0$. This is allowed as $\cos\theta_1\cos\theta_2$ is a T-even quantity, therefore the resulting triple product is still T-odd. The asymmetry is then found to be

$$A_V(B_s^0) = -\frac{2\sqrt{2}}{\pi}\frac{\Im(A_\perp A_0^*)^{\mathrm{TI}}}{|A_0^{\mathrm{TI}}|^2 + |A_\parallel^{\mathrm{TI}}|^2 + |A_\perp^{\mathrm{TI}}|^2}. \tag{B.6}$$

In the measurement of triple product asymmetries in B_s^0 decays, the initial flavour of the B_s^0 meson is not taken in to account. This then means that the overall asymmetry is the average of the contributions from B_s^0 and $\overline{B}_s^0$ decays, which is given by

$$\frac{1}{2}(A_U(B_s^0) + A_U(\overline{B}_s^0)) = -\frac{2}{\pi}\left(\frac{\Im(A_\perp A_\parallel^*)^{\mathrm{TI}}}{|A_0^{\mathrm{TI}}|^2 + |A_\parallel^{\mathrm{TI}}|^2 + |A_\perp^{\mathrm{TI}}|^2} + \frac{\Im(\overline{A}_\perp \overline{A}_\parallel^*)}{|\overline{A}_0^{\mathrm{TI}}|^2 + |\overline{A}_\parallel^{\mathrm{TI}}|^2 + |\overline{A}_\perp^{\mathrm{TI}}|^2}\right), \tag{B.7}$$

$$\frac{1}{2}(A_V(B_s^0) + A_V(\overline{B}_s^0)) = -\frac{\sqrt{2}}{\pi}\left(\frac{\Im(A_\perp A_0^*)^{\mathrm{TI}}}{|A_0^{\mathrm{TI}}|^2 + |A_\parallel^{\mathrm{TI}}|^2 + |A_\perp^{\mathrm{TI}}|^2} + \frac{\Im(\overline{A}_\perp \overline{A}_0^*)^{\mathrm{TI}}}{|\overline{A}_0^{\mathrm{TI}}|^2 + |\overline{A}_\parallel^{\mathrm{TI}}|^2 + |\overline{A}_\perp^{\mathrm{TI}}|^2}\right). \tag{B.8}$$

If the overall decay rate of $B_s^0 \to \phi\phi$ is the same as $\overline{B}_s^0 \to \phi\phi$, then the asymmetries, A_U and A_V, are simply proportional to $\Im(A_\perp A_\parallel^* + \overline{A}_\perp \overline{A}_\parallel^*)^{\mathrm{TI}}$ and $\Im(A_\perp A_0^* + \overline{A}_\perp \overline{A}_0^*)^{\mathrm{TI}}$, respectively.

In the derivation presented in Appendix A, it has been assumed that the *CP*-violating weak phase is the same for all polarisations. If this assumption is not enforced, the asymmetries may be found to be of the form [28]

$$\begin{aligned}&\int_0^\infty \Im(A_\perp(t)A_\parallel^*(t)+\overline{A}_\perp(t)\overline{A}_\parallel^*(t))\mathrm{d}t = 2|A_\perp||A_\parallel|\cos(\delta_\perp-\delta_\parallel)\\&\big(\sin(\phi_\perp-\phi_\parallel)-\sin(\phi_\perp+\phi_\parallel)(\Delta\Gamma_s/2\Gamma_s)\big)+\mathcal{O}[(\Delta\Gamma_s/2\Gamma_s)^2],\end{aligned}\tag{B.9}$$

$$\begin{aligned}&\int_0^\infty \Im(A_\perp(t)A_0^*(t)+\overline{A}_\perp(t)\overline{A}_0^*(t))\mathrm{d}t = 2|A_\perp||A_0|\cos(\delta_\perp-\delta_0)\\&(\sin(\phi_\perp-\phi_0)-\sin(\phi_\perp+\phi_0)(\Delta\Gamma_s/2\Gamma_s))+\mathcal{O}[(\Delta\Gamma_s/2\Gamma_s)^2],\end{aligned}\tag{B.10}$$

where $\phi_i\, i\in\{0,\perp,\parallel\}$ denotes polarisation-dependent *CP*-violating weak phases. It can be seen from equations B.9 and B.10 that the asymmetries are only theoretically non-zero in the presence of *CP* violation, though *CP* violation can be hidden with certain values of the strong phases. If there is no difference in the *CP*-violating phase obtained from different polarisations, then the triple product asymmetries are suppressed by factors of $\Delta\Gamma_s/2\Gamma_s$.

It has been pointed-out in the work of Datta *et al.* [82] that fake triple product asymmetries can be complementary to the true triple product asymmetries in the identification of New Physics. Fake triple product asymmetries are constructed from the difference in the interference terms between $B_s^0\to\phi\phi$ and $\overline{B}_s^0\to\phi\phi$ decays, and so are proportional to $\Im(A_\perp A_i^*-\overline{A}_\perp\overline{A}_i^*)^{\mathrm{TI}}$, for $i\in\{0,\parallel\}$. As the difference is required rather than the sum, tagging of the initial flavour of the B_s^0 meson is necessary, which will limit the power of such measurements experimentally. Phenomenologically, the fake triple products may be evaluated as [28]

$$\begin{aligned}&\int_0^\infty \Im(A_\perp(t)A_\parallel^*(t)-\overline{A}_\perp(t)\overline{A}_\parallel^*(t))\mathrm{d}t =\\&2|A_\perp||A_\parallel|(\sin(\delta_\perp-\delta_\parallel)\cos(\phi_\perp-\phi_\parallel)(\Gamma_s/\Delta m_s)^2\\&-\cos(\delta_\perp-\delta_\parallel)\cos(\phi_\perp+\phi_\parallel)(\Gamma_s/\Delta m_s)+\mathcal{O}[(\Gamma_s/\Delta m_s)^3],\end{aligned}\tag{B.11}$$

$$\begin{aligned}&\int_0^\infty \Im(A_\perp(t)A_0^*(t)-\overline{A}_\perp(t)\overline{A}_0^*(t))\mathrm{d}t =\\&2|A_\perp||A_0|(\sin(\delta_\perp-\delta_0)\cos(\phi_\perp-\phi_0)(\Gamma_s/\Delta m_s)^2\\&-\cos(\delta_\perp-\delta_0)\cos(\phi_\perp+\phi_0)(\Gamma_s/\Delta m_s)+\mathcal{O}[(\Gamma_s/\Delta m_s)^3],\end{aligned}\tag{B.12}$$

where it can be seen that the fast B_s^0-$\overline{B}_s^0$ oscillations will suppress fake triple product asymmetries. The fake triple product asymmetries, while suppressed in B_s^0 decays and experimentally more challenging to measure, contain a different dependence on the

CP-conserving strong phases, and hence provide more experimental information on parameters that can be affected by the presence of NP.

Appendix C.

Calculation of the B_s^0 Mixing Box Diagram

The theoretical prediction of the B_s^0 mixing box diagram makes extensive use of effective field theory to separate long distance and short distance QCD effects through the three scales relevant for B_s^0 mixing, consisting of the masses of the top quark (m_t), the mass of the $W^\pm$ bosons, and the mass of the bottom quark (m_b). The sizes of the scales relative to the QCD scale (Λ_{QCD}) obey $m_t \sim m_W >> m_b >> \Lambda_{QCD}$. The defining relation of the effective field theory is [23]

$$\langle f|\mathcal{T}\exp\left(-i\int \mathrm{d}^4x\mathcal{H}_{int}^{SM}(x)\right)|i\rangle = \langle f|\mathcal{T}\exp\left(-i\int \mathrm{d}^4x\mathcal{H}_{int}^{eff}(x)\right)|i\rangle[1+\mathcal{O}\left(\frac{m_{light}}{m_{heavy}}\right)], \tag{C.1}$$

where f is the final state, i the initial state and $\mathcal{H}_{int}^{eff}(x)$ refers to the effective Hamiltonian. In the case of B_s^0-$\overline{B}_s^0$ mixing, we take the final state to be B_s^0 and the initial state to be $\overline{B}_s^0$, the light mass scale $m_{light} \sim m_b$ and the heavy mass scale $m_{heavy} = m_W \sim m_t$. The effective Hamiltonian at leading order relevant for B_s^0 mixing is an effective 4-quark coupling

$$\mathcal{H}^{eff} = H^{|\Delta B|=2} = \frac{G_F^2}{4\pi^2}(V_{tb}V_{ts}^*)^2 C^{|\Delta B|=2}(m_t, m_W, \mu)Q(\mu) + h.c., \tag{C.2}$$

where $C^{|\Delta B|=2}(m_t, m_W, \mu)$ is the Wilson coefficient and $Q(\mu) = \overline{s}\gamma_\nu b s\gamma^\nu\overline{b}$ at scale μ. The Wilson coefficients are perturbative in QCD, i.e. $C^{|\Delta B|=2} = C^{|\Delta B|=2(0)} + \frac{\alpha_s}{4\pi}C^{|\Delta B|=2(1)}$, and are found through matching the effective theory to the SM box diagram. At next-to-leading order (NLO), the effective Hamiltonian must be supplemented by an additional term (H^{QCD}), corresponding to the gluons that must dress the box diagram. In this case, the SM amplitude becomes

$$\begin{aligned} -\mathcal{M} &= -\mathcal{M}^{(0)} - \mathcal{M}^{(1)} \\ &= \frac{G_F^2}{4\pi^2}(V_{tb}V_{ts}^*)\left(C^{|\Delta B|=2(0)} + \frac{\alpha_s}{4\pi}C^{|\Delta B|=2(1)}\right)\cdot\left(< Q >^{(0)} + \frac{\alpha_s}{4\pi} < Q >^{(1)}\right)\cdot[1+\mathcal{O}\left(\frac{m_b^2}{m_W^2}\right)], \end{aligned} \tag{C.3}$$

where $< Q >^{(n)} \equiv \langle B_s^0|Q(\mu)|\overline{B}_s^0\rangle^{(n)}$. Loop calculations in quantum field theories such as the Standard Model contain divergences that must be handled in order to make reasonable predictions. Divergences in quantum field theories can be separated into short distance effects, known as ultra-violet (UV) divergences, long distance effects, known as infra-red (IR) divergences and colinear divergences. UV divergences are handled within the framework of a renormalisation, such as the $\overline{\text{MS}}$ scheme [83]. IR divergences can only be handled through cancellation between real and virtual Feynman diagrams. In such cases, IR safe observables must be measured in order to have sensible theoretical predictions. The renormalised operator, Q, is related to the bare operator Q_{bare}, through $Q = Z_Q(\mu)^{-1}Q_{\text{bare}}$, where $Z_Q^{-1}(\mu)$ serves the same purpose as a counterterm in the renormalisation of the coupling constant or the fermion mass.

The box diagram describing B_s^0-$\overline{B}_s^0$ mixing is known to contain both IR and UV divergences. The effective theory diagram contains the same IR divergences and dependence on the light mass scales. At one-loop, the IR sensitive diagrams are those in which there is a virtual correction to a vertex. As other diagrams are finite, light masses may be neglected. The IR structure of $\mathcal{M}$ is contained in $\langle Q\rangle$, therefore perturbative methods are meaningful for $C^{|\Delta B|=2}$, but not for $\langle Q\rangle$. This means that $\langle Q\rangle$ must be calculated using non-perturbative methods. The factorisation into short distance coefficients and long distance operators is known as the Operator Product Expansion (OPE).

The amplitude for the box diagram (at leading order), including all three quark flavours, may be written as

$$\mathcal{M}^{(0)} = \sum_{j,k=u,c,t} V_{jb}^* V_{js} V_{kb}^* V_{ks} M_{jk}^{(0)} \langle Q\rangle^{(0)}, \tag{C.4}$$

where V_{ij} are elements of the CKM matrix and $M_{jk}^{(0)} = -\frac{\mathrm{G}_F}{4\pi^2} m_W^2 \widetilde{S}(x_j, x_k)$ with $x_j \equiv \frac{m_j^2}{m_W^2}$. CKM unitarity gives the relation $V_{ub}^* V_{us} = -V_{tb}^* V_{ts} - V_{cb}^* V_{cs}$. This allows the amplitude to be rewritten as (setting $x_u = 0$)

$$\begin{aligned}-\mathcal{M}^{(0)} &= \frac{\mathrm{G}_F^2}{4\pi^2}\langle Q\rangle^{(0)}\Big((V_{tb}^*V_{ts})^2(\widetilde{S}(x_t,x_t)-\widetilde{S}(0,x_t)-\widetilde{S}(x_t,0)+\widetilde{S}(0,0))\\ &+(V_{cb}^*V_{cs})^2(\widetilde{S}(x_c,x_c)-\widetilde{S}(0,x_c)-\widetilde{S}(x_c,0)+\widetilde{S}(0,0))\\ &+2(V_{tb}^*V_{ts})(V_{cb}^*V_{cs})(\widetilde{S}(x_t,x_c)-\widetilde{S}(0,x_t)-\widetilde{S}(x_c,0)+\widetilde{S}(0,0))\Big)\\ &= \frac{\mathrm{G}_F^2}{4\pi^2}\langle Q\rangle^{(0)}\left((V_{tb}^*V_{ts})^2S(x_t)+(V_{cb}^*V_{cs})^2S(x_c)+2(V_{tb}^*V_{ts})(V_{cb}^*V_{cs})S(x_t,x_c)\right),\end{aligned} \tag{C.5}$$

where the symmetric functions $\widetilde{S}$ are related to the Inami-Lim [84] functions, S, through

$$S(x_j,x_k)=\widetilde{S}(x_j,x_k)-\widetilde{S}(x_j,0)-\widetilde{S}(0,x_k)+\widetilde{S}(0,0). \tag{C.6}$$

The last two terms in equation C.5 may be neglected as $x_c \sim 10^{-4}$, $S(x_c) \sim \mathcal{O}(x_c)$ and $S(x_c) \sim \mathcal{O}(x_c, x_t)$. This is known as Glashow-Iliopoulos-Maiani (GIM) supression. Therefore, the leading order Wilson coefficient is matched as

$$C^{|\Delta B|=2,(0)}(m_t,m_W)=m_W^2S(x_t). \tag{C.7}$$

The renormalisation scale, μ, is present in equation C.7 as there is an implicit dependence on the renormalisation scale in $m_t \equiv m_t(\mu)$, even though there is no explicit dependence. As long as $\mu \sim \mathcal{O}(m_W, m_t)$, there are no large differences in scales in the Wilson coefficients and therefore no large logarithms to spoil convergence. At next-to-leading order (NLO), logarithms of the form $\ln\left(\frac{\mu}{m_W}\right)$ are present in $C^{|\Delta B|=2,(1)}$. The large logarithm $\ln\left(\frac{m_b}{m_W}\right)$ present in $\mathcal{M}^{(1)}$ is split between the Wilson coefficient and the matrix element operator through

$$\ln\left(\frac{m_b}{m_W}\right)=\ln\left(\frac{\mu}{m_W}\right)+\ln\left(\frac{m_b}{\mu}\right). \tag{C.8}$$

The scale at which equation C.3 is evaluated is known as the matching scale. Clearly, the scale has to be the same for both the Wilson coefficients and the matrix element operators. Therefore, a common choice is to use the renormalisation group (RG) equation to relate $C^{|\Delta B|=2}(m_t, m_W, \mu_{tW})$ to $C^{|\Delta B|=2}(m_t, m_W, \mu_b)$ such that the large logarithms, $\alpha_s^n \ln\left(\frac{\mu_{tW}}{\mu_b}\right)$ are summed to all orders in perturbation theory. This RG improvement then promotes the LO result to a leading log (LL) result. As a consequence that physical quantities should not depend on the renormalisation scale, the following relation must hold

$$\frac{\mathrm{d}\ln C}{\mathrm{d}\ln \mu} = -\frac{\mathrm{d}\ln Q}{\mathrm{d}\ln \mu}. \tag{C.9}$$

Renormalisation group evolution leads to the relation between the Wilson coefficients at the two scales, μ_{tW} and μ_b, of

$$C^{|\Delta B|=2,(0)}(m_t, m_W, \mu_b) = \left(\frac{\alpha_s(\mu_{tW})}{\alpha_s(\mu_b)}\right)^{\frac{\gamma_1}{\beta_1}} C^{|\Delta B|=2,(0)}(m_t, m_W, \mu_{tW}), \tag{C.10}$$

where $\gamma_1 = 4$ is first term of the expansion of the anomalous dimension, γ_Q, corresponding to Q, defined by

$$\gamma_Q = \frac{\mathrm{d}\ln Q}{\mathrm{d}\ln \mu} = \gamma_1 \alpha_s + \gamma_2 \alpha_s^2 + \mathcal{O}(\alpha_s^2), \tag{C.11}$$

and $\beta_0 = 2(11 - 2N_f/3)$ is the first term in the expansion of the QCD beta function, defined by

$$\beta = \frac{1}{\alpha_s}\frac{\mathrm{d}\alpha_s}{\mathrm{d}\ln \mu} = \beta_0 \alpha_s^2 + \beta_1 \alpha_s^3 + \mathcal{O}(\alpha_s^4). \tag{C.12}$$

Equation C.10 is usually written to separate the scales[1] as

$$C^{|\Delta B|=2,(0)}(m_t, m_W, \mu_b) = \eta_B b_B(\mu_b) C^{|\Delta B|=2,(0)}(m_t, m_W, \mu_{tW}), \tag{C.13}$$

where η_B and $b_B(\mu_b)$ can be evaluated in the context of naive dimensional regularisation in the $\overline{\mathrm{MS}}$ scheme as 0.55 and 1.5, respectively [23]. Equation C.2 then becomes

$$H^{|\Delta B|=2} = \frac{\mathrm{G}_F^2}{4\pi^2} m_W^2 (V_{tb} V_{ts}^*)^2 \eta_B S(x_t) b_B(\mu_b) Q(\mu_b) + h.c. \tag{C.14}$$

The hadronic matrix element is usually parametrised as

$$\langle B_s^0 | Q(\mu_b) | \overline{B}_s^0 \rangle = \frac{2}{3} M_{B_s^0}^2 f_{B_s^0}^2 \frac{\hat{B}_{B_s^0}}{b_B(\mu_b)}, \tag{C.15}$$

where $\hat{B}_{B_s^0}$ is a bag factor and $f_{B_s^0}$ is the decay constant of the B_s^0 meson. This parametrisation is chosen to not only make the scale cancellation explicit, but also to ensure

---

[1] This is possible to all orders in α_s.

$\frac{\hat{B}_{B_s^0}}{b_B(\mu_b)}$ is close to unity. The factor $f_{B_s^0}\sqrt{\hat{B}_{B_s^0}}$ is obtained from lattice gauge theory to be 270 ± 45 MeV [23].

Putting everything together results in

$$M_{12} = \frac{\langle B_s^0 | H^{|\Delta B|=2} | \overline{B}_s^0 \rangle}{2M_{B_s^0}} = \frac{\mathrm{G}_F^2}{12\pi^2} \eta_B M_{B_s^0} \hat{B}_{B_s^0} f_{B_s^0}^2 m_W^2 S\left(\frac{m_t^2}{m_W^2}\right) (V_{ts}^* V_{tb})^2. \tag{C.16}$$

Appendix D.

Theory of the $\overline{B}{}^0_s \to \phi_L\phi_L$ Amplitude

The methods of factorisation have been introduced in Appendix C in order to compute the SM box diagram contributing to B^0_s-$\overline{B}{}^0_s$ mixing. The same framework can be used to make predictions of the SM contributions to the $B^0_s \to \phi\phi$ physics parameters from penguin diagrams. Note that factorisation can refer to the separation of scales in the operator product expansion and also to the factorisation of hadronic matrix elements, known as QCD factorisation (QCDF). In analogy with equation C.2, the effective Hamiltonian describing the $B^0_s \to \phi\phi$ decay can be written as

$$\mathcal{H}^{eff} = \frac{G_F}{\sqrt{2}} \sum_{p=u,c} \lambda_p \sum_{i=3,\ldots,10} C_i Q_i + h.c., \tag{D.1}$$

where C_i are the Wilson coefficients[1], elements of the CKM matrix, V_{kl}, enter through $\lambda_p = V_{pb}V^*_{ps}$, G_F is the Fermi constant and the operators, Q_i are given by [22]

$$\begin{aligned}
Q_3 &= (\bar{s}b)_{V-A}(\bar{s}s)_{V-A}, & Q_4 &= (\bar{s}_i b_j)_{V-A}(\bar{s}_j s_i)_{V-A}, \\
Q_5 &= (\bar{s}b)_{V-A}(\bar{s}s)_{V+A}, & Q_6 &= (\bar{s}_i b_j)_{V-A}(\bar{s}_j s_i)_{V+A}, \\
Q_7 &= (\bar{s}b)_{V-A}\frac{3}{2}e_s(\bar{s}s)_{V+A}, & Q_8 &= (\bar{s}_i b_j)_{V-A}\frac{3}{2}e_s(\bar{s}_j s_i)_{V+A}, \\
Q_9 &= (\bar{s}b)_{V-A}\frac{3}{2}e_s(\bar{s}s)_{V-A}, & Q_{10} &= (\bar{s}_i b_j)_{V-A}\frac{3}{2}e_s(\bar{s}_j s_i)_{V-A},
\end{aligned} \tag{D.2}$$

where i, j are colour indices, e_s represents the strange quark charge, and $(\bar{p}q)_{V-A} = \bar{p}\gamma_\mu(1 \pm \gamma_5)q$ for quark fields, p and q, and Dirac matrices, γ_μ and γ_5. Note that other contributions to the effective Hamiltonian are present for different final states of B decays.

---

[1] The numbering scheme of the Wilson coefficients and operators is chosen to be consistent with the numbering scheme of the general B decay effective Hamiltonian in the literature.

The operators Q_3, Q_4, Q_5 and Q_6 correspond to gluonic penguins diagrams, while the operators Q_7, Q_8, Q_9 and Q_{10} correspond to the electroweak penguin diagrams [85], which take the form shown in Figure 1.6. To calculate amplitudes for the $\overline{B}_s^0 \to \phi_L\phi_L$ decay from the Hamiltonian, the operators in equation D.2 can be computed from QCD factorisation [86]. This involves splitting the operator in to two parts $\langle\phi_1|J_\mu|B_s^0\rangle$ and $\langle\phi_2|\tilde{J}^\mu|0\rangle$, where $J_\mu = \bar{s}\gamma_\mu(1-\gamma_5)b$ and $\tilde{J}^\mu = \bar{s}\gamma^\mu s$, and ϕ_2 refers to the ϕ meson containing the spectator quark from the $\overline{B}_s^0$ decay. The matrix element $\langle\phi_1|J_\mu|B_s^0\rangle$ can be written in terms of form factors ($F_0^{B_s^0\to\phi}$, $F_1^{B_s^0\to\phi}$ and $F_2^{B_s^0\to\phi}$) defined through [22]

$$\begin{aligned}\langle\phi_1|\bar{s}\gamma^\mu\gamma_5 b|B_s^0\rangle = {}& 2m_\phi F_0(q^2)\frac{\varepsilon_1^\mu\cdot q}{q^2}q^\mu + (m_{B_s^0}+m_\phi)F_1(q^2)\left(\varepsilon_1 - \frac{\varepsilon_1\cdot q}{q^2}q^\mu\right)\\ & - F_2(q^2)\frac{\varepsilon_1\cdot q}{m_{B_s^0}+m_\phi}\left((p_{B_s^0}+p_{\phi 1})^\mu - \frac{m_{B_s^0}^2-m_\phi^2}{q^2}q^\mu\right),\end{aligned} \tag{D.3}$$

where $\varepsilon_{1(2)}$ denotes the polarisation of the $\phi_{1(2)}$ meson, $p_{B_s^0}$ and $p_{\phi 1(2)}$ denotes the 4-momenta of the B_s^0 and $\phi_{1(2)}$ mesons, respectively, $m_{B_s^0}$ and m_ϕ are the masses of the B_s^0 and ϕ mesons, and $q = p_{B_s^0} - p_{\phi 1}$ is the momentum transfer. The matrix element $\langle\phi_2|\tilde{J}^\mu|0\rangle$ can be written in terms of the decay constant (f_ϕ) defined through

$$\langle\phi_2|\bar{s}\gamma_\mu s|0\rangle = -if_\phi m_\phi \varepsilon_{2\mu}. \tag{D.4}$$

This allows the factorised matrix element to be written as

$$\langle\phi_L\phi_L|(\bar{s}b)_{V-A}(\bar{s}s)_V|\overline{B}_s^0\rangle = im_{B_s^0}^2 F_0^{B_s^0\to\phi}(m_\phi^2)f_\phi, \tag{D.5}$$

where the assumption that the ϕ meson is much lighter than the B_s^0 meson is used. The complete $\overline{B}_s^0 \to \phi_L\phi_L$ amplitude, $\mathcal{A}(\overline{B}_s^0 \to \phi_L\phi_L)$ takes the form [22]

$$\mathcal{A}(\overline{B}_s^0 \to \phi_L\phi_L) = \langle\phi_L\phi_L|\mathcal{H}^{eff}|\overline{B}_s^0\rangle = \frac{G_F}{\sqrt{2}}\sum_{p=u,c}\lambda_p\langle\phi_L\phi_L|\mathcal{T}_p|\overline{B}_s^0\rangle, \tag{D.6}$$

where

$$\begin{aligned}\mathcal{T}_p = &(a_3 + a_4^p + a_5)(\bar{s}b)_{V-A}\otimes(\bar{s}s)_V\\ &-\frac{1}{2}(a_7^p + a_9 + a_{10}^p)(\bar{s}b)_{V-A}\otimes(\bar{s}s)_V.\end{aligned} \tag{D.7}$$

In equation D.7, the $\otimes$ symbol indicates that the operator is calculated in factorised form. Bartsch *et al.* [22] have shown that the coefficients for the $B_s^0 \to \phi_L\phi_L$ amplitudes can be evaluated as

$$a_3 = C_3 + \frac{C_4}{N_c}\left(1 + \frac{C_F\alpha_s}{4\pi}V_\phi + \frac{C_F\pi\alpha_s}{N_c}H_{\phi\phi}\right), \tag{D.8}$$

$$a_4^p = C_4 + \frac{C_3}{N_c}\left(1 + \frac{C_F\alpha_s}{4\pi}V_\phi + \frac{C_F\pi\alpha_s}{N_c}H_{\phi\phi}\right) + \frac{C_F\alpha_s}{4\pi N_c}P^p_{\phi,2}, \tag{D.9}$$

$$a_5 = C_5 + \frac{C_6}{N_c}\left(1 + \frac{C_F\alpha_s}{4\pi}(-V'_\phi) + \frac{C_F\pi\alpha_s}{N_c}(-H'_{\phi\phi})\right), \tag{D.10}$$

$$a_7^p = C_7 + \frac{C_8}{N_c}\left(1 + \frac{C_F\alpha_s}{4\pi}(-V'_\phi) + \frac{C_F\pi\alpha_s}{N_c}(-H'_{\phi\phi})\right) + \frac{\alpha_{EW}}{9\pi}P^{p,EW}_{\phi,n}, \tag{D.11}$$

$$a_9 = C_9 + \frac{C_{10}}{N_c}\left(1 + \frac{C_F\alpha_s}{4\pi}V_\phi + \frac{C_F\pi\alpha_s}{N_c}H_{\phi\phi}\right) + \frac{\alpha_{EW}}{9\pi}P^{p,EW}_{\phi,n}, \tag{D.12}$$

$$a_{10}^p = C_{10} + \frac{C_9}{N_c}\left(1 + \frac{C_F\alpha_s}{4\pi}V_\phi + \frac{C_F\pi\alpha_s}{N_c}H_{\phi\phi}\right) + \frac{\alpha_{EW}}{9\pi N_c}P^{p,EW}_{\phi,2}, \tag{D.13}$$

where α_s and α_{EW} are the coupling constants of the QCD and EW interactions, respectively, and $C_F = 4/3$ is the colour factor from QCD for 3 colours (N_c). The coupling constants and Wilson coefficients are renormalised to scale, μ. Note that scalar penguin corrections to equations D.8 to D.13 are ignored for simplicity. The hadronic quantities, $V_\phi^{(\prime)}$, $H_{\phi\phi}^{(\prime)}$, $P^p_{\phi,2}$, $P^{p,EW}_{\phi,2}$, ϕ and $P^{p,EW}_{\phi,n}$, in equations D.8 to D.13 correspond to

- $V_\phi^{(\prime)}$: Gluonic vertex corrections between the quarks in the ϕ_1 meson and the external quark lines except for the spectator $\overline{s}$-quark.
- $H_{\phi\phi}^{(\prime)}$: Gluonic corrections between the spectator quark and the quarks of the ϕ_1 meson, known as hard spectator scattering.
- $P^p_{\phi,n}$: Gluonic penguin corrections to the effective vertex of twist n, where the twist is the difference between the dimension of the operator and the spin.
- $P^{p,EW}_{\phi,n}$: Electroweak penguin corrections to the effective vertex of twist n.

Equation D.6 then evaluates to

$$\begin{aligned}\mathcal{A}(\overline{B}{}^0_s \to \phi_L\phi_L) &= 2i\frac{G_F}{\sqrt{2}}\lambda_p m^2_{B^0_s}F_0^{B^0_s\to\phi}(m_\phi)f_\phi\left(a_3 + a_4^p + a_5 - \frac{1}{2}(a_7^p + a_9 + a_{10}^p)\right)\\ &\equiv i\frac{G_F}{\sqrt{2}}\lambda_p a^p.\end{aligned} \tag{D.14}$$

The decay time-dependent *CP* asymmetry is defined in analogy with equation 1.48 as

$$\frac{\Gamma(B_s^0 \to \phi_L\phi_L) - \Gamma(\overline{B}_s^0 \to \phi_L\phi_L)}{\Gamma(B_s^0 \to \phi_L\phi_L) + \Gamma(\overline{B}_s^0 \to \phi_L\phi_L)} = S_\phi \sin(\Delta m_s t), \tag{D.15}$$

where Δm_s is the B_s^0 oscillation frequency and no direct *CP* violation is assumed. The coefficient, S_ϕ, is given by

$$S_\phi = \frac{2\Im(\xi)}{1+|\xi|^2}, \tag{D.16}$$

where

$$\xi = -\frac{M_{12}^*}{|M_{12}|}\frac{\mathcal{A}(\overline{B}_s^0 \to \phi_L\phi_L)}{\mathcal{A}(B_s^0 \to \phi_L\phi_L)}, \tag{D.17}$$

with M_{12} denoting the B_s^0-$\overline{B}_s^0$ mixing amplitude. In terms of the coefficients of the $\overline{B}_s^0 \to \phi_L\phi_L$ amplitude in equation D.14, the *CP* violation parameter S_ϕ is given by [22]

$$S_\phi = 2\lambda^2\eta\Re\left(\frac{a^c - a^u}{a^c}\right), \tag{D.18}$$

where λ and η are parameters of the CKM matrix in the Wolfenstein parametrisation. The value of $|a^c - a^u|$ can be computed in QCDF as the leading contributions from diagrams in which the spectator $\overline{s}$-quark annihilates cancel[2]. Bartsch *et al.* therefore evaluate this as $|a^c - a^u| = 0.057^{+0.023}_{-0.029}$, where the largest contributions to the uncertainty arise from the renormalisation scale and charm quark mass.

As has been explained in Section 1.4.2.1, annihilation contributions play a significant role in the determination of $|a^c|$, which is then found from measurements of the $B_s^0 \to \phi_L\phi_L$ branching ratio, and can be evaluated as

$$|a^c| = 0.177\,\mathrm{GeV}^3\left(\frac{\mathcal{B}(B_s^0 \to \phi_L\phi_L)}{15\cdot 10^{-6}}\right)^{1/2}\left(\frac{1.53\ \mathrm{ps}}{\tau_{B_s^0}}\right)^{1/2}, \tag{D.19}$$

where $\tau_{B_s^0}$ is the B_s^0 lifetime. The SM upper limit can then be written as

$$S_\phi \leq 2\lambda^2\eta\frac{|a^c - a_u|}{|a^c|} \leq \lambda^2\eta\left(\frac{\mathcal{B}(B_s^0 \to \phi_L\phi_L)}{15\cdot 10^{-6}}\right)^{-1/2}, \tag{D.20}$$

---

[2] Such penguin annihilation diagrams are difficult to calculate numerically in the framework of QCDF.

where the value of S_ϕ is found to be positive in the SM. Bartsch *et al.* therefore evaluated the upper limit using the QCDF value of $\mathcal{B}(\overline{B}{}^0_s \to \phi_L\phi_L)^{QCDF} = 15.5 \cdot 10^{-6}$ [22], arriving at $S_\phi < 0.020$. However, this number can be updated using the value of the CDF measurement of the branching ratio, $\mathcal{B}(B^0_s \to \phi_L\phi_L)^{exp} = (2.32 \pm 0.84) \cdot 10^{-5}$ [27], in combination with the value of the longitudinal polarisation fraction presented in Chapter 5, $|A_0|^2 = 0.365 \pm 0.025$, to give an upper limit of $S_\phi < 0.016$.

Appendix E.

Comparisons Between Data and Monte Carlo

To check the agreement between data and simulation the ${}_s\mathcal{P}lot$ [54] technique is used to create background-unfolded distributions for the momentum, p_{T} and η of the reconstructed B_s^0 meson in addition to the minimum kaon p_{T} in each event. These are compared to the corresponding distributions from simulation in Figs E.1 - E.5 for the multi-variate selection used in the decay time-dependent analysis of the $B_s^0 \to \phi\phi$ decay.

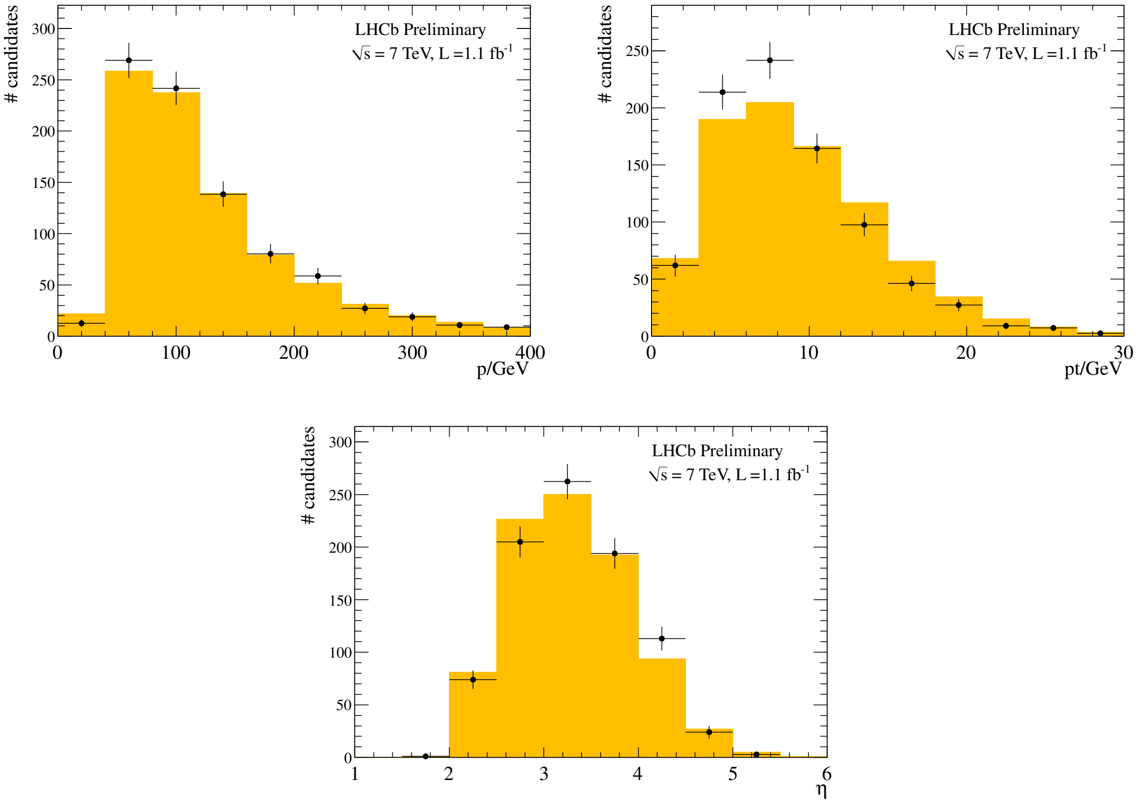


Figure E.1.: Comparison of the B_s^0 momentum (top left), B_s^0 p_T (top right) and B_s^0 η (bottom) distributions obtained using the ${}_s\mathcal{P}lot$ technique from the $B_s^0 \to \phi\phi$ data (black points) with the Monte Carlo simulation (solid yellow). The Monte Carlo histogram is normalised to the number of events seen in the data.

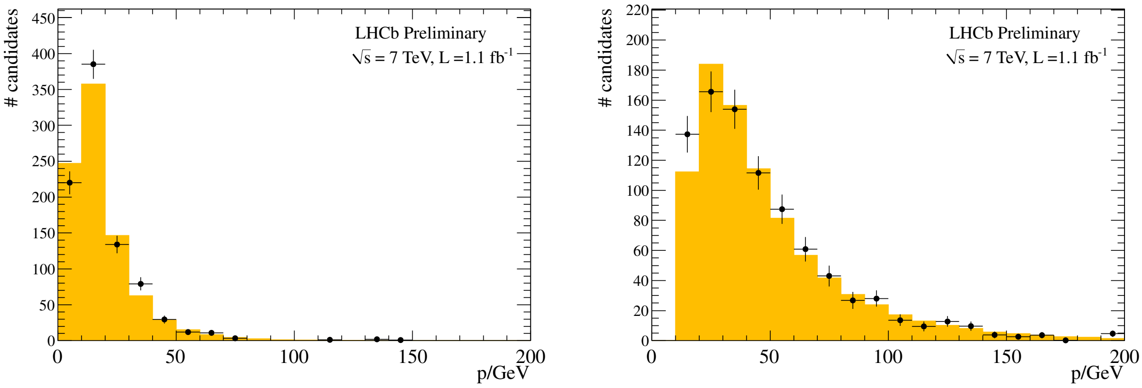


Figure E.2.: Comparison of the minimum (left) and maximum (right) kaon momentum obtained using the ${}_s\mathcal{P}lot$ technique from the $B_s \to \phi\phi$ data (black points) with the Monte Carlo simulation (solid yellow). The Monte Carlo histogram is normalised to the number of events seen in the data.

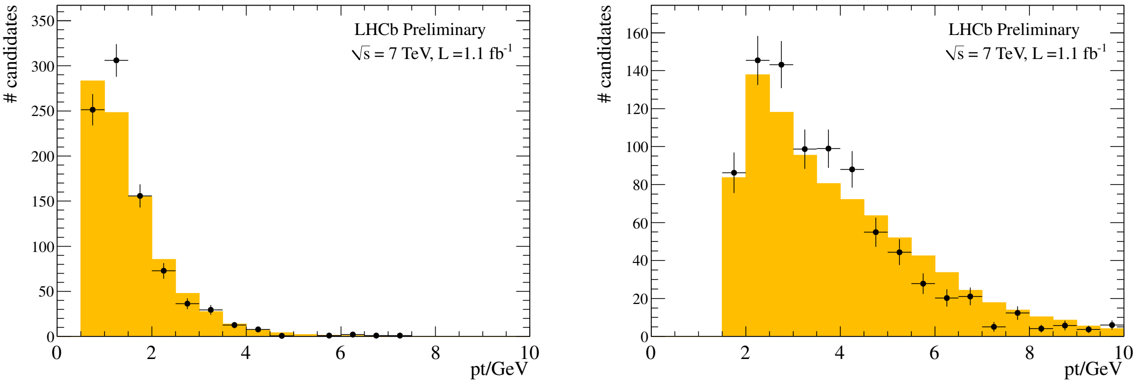


Figure E.3.: Comparison of the minimum (left) and maximum (right) kaon pt obtained using the ${}_s\mathcal{P}lot$ technique from the $B_s \to \phi\phi$ data (black points) with the Monte Carlo simulation (solid yellow). The Monte Carlo histogram is normalised to the number of events seen in the data.

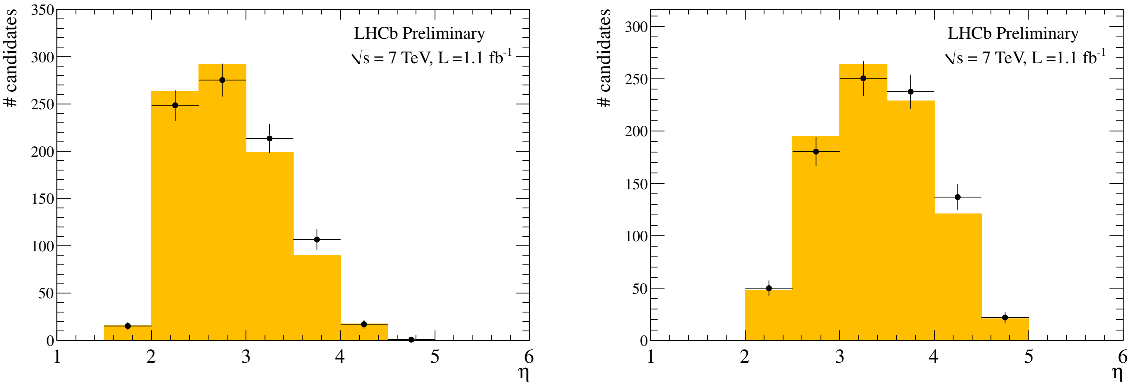


Figure E.4.: Comparison of the minimum (left) and maximum (right) kaon η obtained using the ${}_s\mathcal{P}lot$ technique from the $B_s \to \phi\phi$ data (black points) with the Monte Carlo simulation (solid yellow). The Monte Carlo histogram is normalised to the number of events seen in the data.

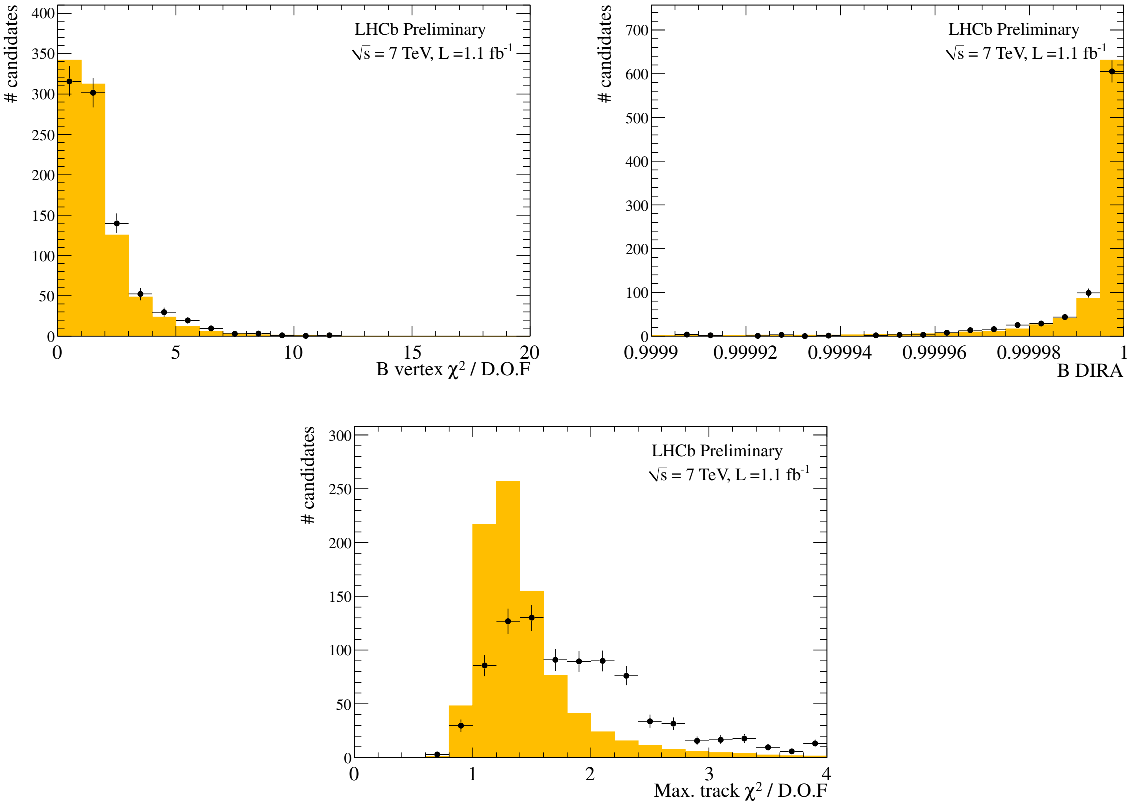


Figure E.5.: Comparison of the B_s^0 vertex χ^2 per NDF (top left), B_s^0 DIRA (top right) and maximum kaon track χ^2 per NDF (bottoms) obtained using the ${}_s\mathcal{P}lot$ technique from the $B_s \to \phi\phi$ data (black points) with the Monte Carlo simulation (solid yellow). The Monte Carlo histogram is normalised to the number of events seen in the data.

Bibliography

[1] E. Abers and B. Lee, *Gauge Theories*, Phys. Rept. **9** (1973) 1.

[2] P. W. Anderson, *Plasmons, gauge invariance, and mass*, Phys. Rev. **130** (1963) 439.

[3] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[4] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[5] G. S. Guralnik, C. R. Hagen, and T. W. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[6] J. Goldstone, A. Salam, and S. Weinberg, *Broken symmetries*, Phys. Rev. **127** (1962) 965.

[7] R. B. Mann, *An introduction to particle physics and the standard model.* CRC Press, 2010.

[8] L.-L. Chau and W.-Y. Keung, *Comments on the parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[9] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP violation*, Phys. Rev. Lett. **55** (1985) 1039.

[10] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[11] J. Charles *et al.*, *Predictions of selected flavour observables within the Standard Model*, Phys. Rev. **D84** (2011) 033005, `arXiv:1106.4041`.

[12] M. Gronau and D. Wyler, *On determining a weak phase from CP asymmetries in charged B decays*, Phys. Lett. **B265** (1991) 172.

[13] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing CP violation in $B^+ \to KD$ and measuring the CKM phase gamma*, Phys. Rev. **D63** (2001) 036005, `arXiv:hep-ph/0008090`.

[14] A. Lenz and U. Nierste, *Numerical updates of lifetimes and mixing parameters of B mesons*, Proceedings, 6th International Workshop, CKM 2010 (2011) `arXiv:1102.4274`.

[15] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays*, Phys. Rev. D 87, **112010** (2013) , `arXiv:1304.2600`.

[16] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`, Updated results and plots available at: ckmfitter.in2p3.fr.

[17] A. Lenz *et al.*, *Constraints on new physics in $B-\bar{B}$ mixing in the light of recent LHCb data*, Phys. Rev. **D86** (2012) 033008, `arXiv:1203.0238`.

[18] I. I. Bigi, *CP violation.* Cambridge monographs on particle physics, nuclear physics and cosmology, 2009.

[19] G. D'Ambrosio, G. Giudice, G. Isidori, and A. Strumia, *Minimal flavor violation: An effective field theory approach*, Nucl. Phys. **B645** (2002) 155, `arXiv:hep-ph/0207036`.

[20] B. Bhattacharya, A. Datta, M. Duraisamy, and D. London, *Searching for new physics with $\bar{b} \to \bar{s}$ $B_s^0 \to V_1V_2$ penguin decays*, `arXiv:1306.1911`.

[21] A. Ali *et al.*, *Charmless non-leptonic B_s decays to PP, PV and VV final states in the pQCD approach*, Phys. Rev. **D76** (2007) 074018, `arXiv:hep-ph/0703162`.

[22] M. Bartsch, G. Buchalla, and C. Kraus, *$B \to V_LV_L$ decays at next-to-leading order in QCD*, `arXiv:0810.0249`.

[23] U. Nierste, *Three lectures on meson mixing and CKM phenomenology*, Proceedings, Helmholtz International School (2009) `arXiv:0904.1869`.

[24] CDF Collaboration, D. Acosta *et al.*, *First evidence for $B_s^0 \to \phi\phi$ decay and measurements of branching ratio and A_{CP} for $B^+ \to \phi K^+$*, Phys. Rev. Lett. **95** (2005) 031801, `arXiv:hep-ex/0502044`.

[25] BaBar Collaboration, B. Aubert *et al.*, *Measurement of CP asymmetries in $B^0 \to \phi K^0$ and $B^0 \to K^+K^-K_S^0$ decays*, Phys. Rev. **D71** (2005) 091102, `arXiv:hep-ex/0502019`.

[26] Belle Collaboration, K. Abe *et al.*, *Measurement of time dependent CP violating asymmetries in $B^0 \to \phi K_S^0$, $K^+K^-K_S^0$, and $\eta' K_S^0$ decays*, Phys. Rev. Lett. **91** (2003) 261602, `arXiv:hep-ex/0308035`.

[27] CDF Collaboration, T. Aaltonen *et al.*, *Measurement of polarization and search for CP violation in $B_s^0 \to \phi\phi$ decays*, Phys. Rev. Lett. **107** (2011) 261802, `arXiv:1107.4999`.

[28] M. Gronau and J. L. Rosner, *Triple product asymmetries in K, $D_{(s)}$ and $B_{(s)}$ decays*, Phys. Rev. **D84** (2011) 096013, `arXiv:1107.1232`.

[29] A. Datta, D. London, J. Matias, M. Nagashima, and A. Szynkman, *Final-state*

polarization in B_s decays, Eur. Phys. J. **C60** (2009) 279, `arXiv:0802.0897`.

[30] H.-Y. Cheng and C.-K. Chua, *QCD factorization for charmless hadronic B_s decays revisited*, Phys. Rev. **D80** (2009) 114026, `arXiv:0910.5237`.

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

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

[33] L. Evans, *The LHC machine*, PoS **EPS-HEP2009** (2009) 004.

[34] J.-L. Caron, *Layout of the LEP tunnel including future LHC infrastructures*, AC Collection, Feb, 1997.

[35] R. Wunstorf *et al.*, *Results on radiation hardness of silicon detectors up to neutron fluences of* $10^{15} n/cm^2$, Nucl. Instrum. Meth. **A315** (1992) 149.

[36] A. Affolder *et al.*, *Radiation damage in the LHCb Vertex Locator*, Nucl. Instrum. Meth. **A718** (2013) 312, `arXiv:1302.5259`.

[37] D. Esperante *et al.*, *Performance results of the LHCb Silicon Tracker detector at the LHC*, PoS **RD11** (2011) 001.

[38] LHCb Outer Tracker group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, LHCb-DP-2013-003.

[39] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, LHCb-DP-2012-003.

[40] J. B. Birks, *The Theory and practice of scintillation counting.* Pergamon, 1964.

[41] LHCb Collaboration, I. Machikhiliyan, *Current status and performance of the LHCb electromagnetic and hadron calorimeters*, J. Phys. Conf. Ser. **293** (2011) 012052.

[42] P. Perret, *Overview of the LHCb Calorimeter detectors*, LHCb-TALK-2013-023.

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

[44] LHCb Collaboration, R. Aaij *et al.*, *Evidence for the decay $B^0 \to J/\psi\omega$ and measurement of the relative branching fractions of B_s^0 meson decays to $J/\psi\eta$ and $J/\psi\eta'$*, Nucl. Phys. **B867** (2013) 547, `arXiv:1210.2631`.

[45] J. Alves, A.A. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[46] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

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

[48] O. Callot and S. Hansmann-Menzemer, *The forward tracking: Algorithm and performance studies*, LHCb-2007-015.

[49] A. Powell, *Analogue Pilot Testing*, LHCb RICH L0 Production Readiness Review, June, 2005.

[50] R. J. Barlow, *Statistics: A guide to the use of statistical methods in the physical sciences.* Wiley Blackwell, 1989.

[51] G. J. Feldman and R. D. Cousins, *A unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, `arXiv:physics/9711021`.

[52] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*, Phys. Rev. Lett. **108** (2012) 101803, `arXiv:1112.3183`.

[53] T. du Pree, M. H. M. Merk, and H. G. Raven, *Search for a strange phase in beautiful oscillations.* PhD thesis, Amsterdam, Vrije Universiteit, 2010.

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

[55] C. Fitzpatrick, *simpletools: Handy command line tools for ntuple manipulation and analysis*, LHCb-INT-2009-029.

[56] B. P. Roe *et al.*, *Boosted decision trees, an alternative to artificial neural networks*, Nucl. Instrum. Meth. **A543** (2005) 577, `arXiv:physics/0408124`.

[57] D. J. Amit, K. Wong, and C. Campbell, *Perceptron learning with sign constrained weights*, J. Phys. **A22** (1989) 2039.

[58] J. Therhaag, *TMVA Toolkit for multivariate data analysis in ROOT*, PoS **ICHEP2010** (2010) 510.

[59] Y. Xie, *Principles to optimize event selections for measurements of CP asymmetries*, LHCb-INT-2009-012.

[60] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes and triple product asymmetries in the $B_s^0 \to \phi\phi$ decay*, Phys. Lett. **B713** (2012) 369, `arXiv:1204.2813`.

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

[62] C.-W. Chiang and L. Wolfenstein, *Observables in the decays of B to two vector mesons*, Phys. Rev. **D61** (2000) 074031, `arXiv:hep-ph/9911338`.

[63] C.-W. Chiang, *Angular distribution of charming $B \to VV$ decays and time evolution effects*, Phys. Rev. **D62** (2000) 014017, `arXiv:hep-ph/0002243`.

[64] M. Needham, *Momentum scale calibration using resonances*, LHCb-2008-037.

[65] LHCb collaboration, R. Aaij *et al.*, *First measurement of the CP-violating phase in $B_s^0 \to \phi\phi$ decays*, Phys. Rev. Lett. **110** (2013) 241802, `arXiv:1303.7125`.

[66] Y. Xie, *Some physics and technical issues in analysis of the decay $B_s^0 \to J/\Psi K^+K^-$*, LHCb-INT-2012-017.

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

[68] LHCb Collaboration, R. Aaij *et al.*, *Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data*, LHCb-CONF-2012-033.

[69] H. Moser and A. Roussarie, *Mathematical methods for B^0 anti-B^0 oscillation analyses*, Nucl. Instrum. Meth. **A384** (1997) 491.

[70] R. Aaij *et al.*, *Selections and lifetime measurements for exclusive $b \to J/\Psi X$ decays with $J/\Psi \to \mu^+\mu^-$ with 2010 data*, LHCb-ANA-2011-001.

[71] LHCb collaboration, *Measurement of Δm_s in the decay $B_s^0 \to D_s^-(K^+K^-\pi^-)\pi^+$ using opposite-side and same-side flavour tagging algorithms*, LHCb-CONF-2011-050.

[72] S. M. Flatte, *On the Nature of 0^+ Mesons*, Phys. Lett. **B63** (1976) 228.

[73] HERA-B collaboration, I. Abt *et al.*, K^{*0}*and ϕ meson production in proton-nucleus interactions at $\sqrt{s} = 41.6$ GeV*, Eur. Phys. J. **C50** (2007) 315, `arXiv:hep-ex/0606049`.

[74] LHCb collaboration, R. Aaij *et al.*, *Analysis of the resonant components in $\overline{B}^0 \to J/\psi\pi^+\pi^-$*, Phys. Rev. **D87** (2013) 052001, `arXiv:1301.5347`.

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

[76] LHCb collaboration, R. Aaij *et al.*, *Framework TDR for the LHCb Upgrade: Technical design report*, CERN-LHCC-2012-007. LHCb-TDR-12.

[77] LHCb Collaboration, *LHCb VELO upgrade technical design report*, CERN-LHCC-2013-021. LHCB-TDR-013.

[78] LHCb collaboration, R. Aaij *et al.*, *LHCb PID upgrade technical design report*, CERN-LHCC-2013-022. LHCB-TDR-014.

[79] LHCb collaboration, R. Aaij *et al.*, *Letter of intent for the LHCb upgrade*, CERN-LHCC-2011-001. LHCC-I-018.

[80] T. Gerson, *Updated sensitivity projections for the LHCb upgrade*, LHCb-PUB-2013-015. CERN-LHCb-PUB-2013-015.

[81] R. Jacobsson, *Plans and physics outlook for non-high luminosity experiments until and after LS3*, CERN-LHCb-PROC-2013-078.

[82] A. Datta, M. Duraisamy, and D. London, *Searching for new physics with B-decay fake triple products*, Phys. Lett. **B701** (2011) 357, `arXiv:1103.2442`.

[83] G. 't Hooft, *Dimensional regularization and the renormalization group*, Nucl. Phys. **B61** (1973) 455.

[84] T. Inami and C. Lim, *Effects of superheavy quarks and leptons in low-energy weak Processes* $K(L) \to \mu^+\mu^-$*,* $K^+ \to \pi^+\nu\overline{\nu}$ *and* $K^0 \to \overline{K}^0$, Prog. Theor. Phys. **65** (1981) 297.

[85] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, `arXiv:hep-ph/9512380`.

[86] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *QCD factorization for exclusive, nonleptonic B meson decays: General arguments and the case of heavy light final states*, Nucl. Phys. **B591** (2000) 313, `arXiv:hep-ph/0006124`.