THE UNIVERSITY
of EDINBURGH

Determining the weak couplings $|V_{ub}|/|V_{cb}|$ using semileptonic decays

Iwan Smith

Doctor of Philosophy
The University of Edinburgh
2019

Abstract

This thesis will present the first observation of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ and the determination of $|V_{ub}|/|V_{cb}|$. Using 2 fb^{-1} of data with a centre of mass energy of $\sqrt{s} = 8$ TeV provided by the Large Hadron Collider and collected using the LHCb experiment, a measurement of the ratio of branching fractions of the decays $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ is performed. This is the first observation of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ which is found to have the branching fraction,

$$\frac{\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} = (3.70 \pm 0.29 \pm 0.51) \times 10^{-3}, \tag{1}$$

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

A second set of branching fraction measurements are made, restricted to high and low regions of q^2. The experimental ratio of branching fractions is combined with form factor calculations allowing for measurements of $|V_{ub}|/|V_{cb}|$ to be performed. There is a long standing discrepancy of $\approx 3.5\sigma$ between exclusive and inclusive measurements of $|V_{ub}|$ and a new measurement of this parameter provides some clarity on this discrepancy. Form factors from lattice QCD in the high q^2 region give,

$$\frac{|V_{ub}|}{|V_{cb}|} = 0.0719 \pm 0.0056 \pm 0.0086, \tag{2}$$

and form factors from light-cone sum rules in the low q^2 region give

$$\frac{|V_{ub}|}{|V_{cb}|} = 0.0625 \pm 0.0090 \pm 0.0039, \tag{3}$$

where the first uncertainty is experimental and the second is theoretical. The two measurements are in agreement and differ by 1σ. This high precision measurement of $|V_{ub}|/|V_{cb}|$ provides an essential constraint for global fits to the CKM sector, and these results confirm the long-standing tension between inclusive and exclusive determinations of $|V_{ub}|$.

Lay Summary

The world around us is made of atoms, which in turn are made from fundamental particles known as quarks and leptons. The quarks combine to form protons and neutrons which form the nucleus of an atom orbited by electrons. There are six flavours of quarks grouped into three generations of matter and two types referred to as up and down. All physical matter around is made from the two lightest quarks. The up type quarks (up, charm and top) may transition to down type quarks (down, strange and bottom) and vice versa, with the relative coupling strengths described by a 3×3 unitary matrix, known as the CKM matrix. The strengths of the matrix elements may be determined by investigating decays of particles sensitive to different elements of the matrix. The element V_{ub} couples the up quark to the bottom quark and is the smallest of the elements with the largest relative uncertainty. Historically measurements of $|V_{ub}|$ have been performed using exclusive decays where a a specific decay is measured, and inclusive decays where many decays containing a $b \to u$ transition are measured simultaneously. There is a discrepancy of approximately 3.5σ between the inclusive and exclusive determinations of $|V_{ub}|$.

The LHCb experiment forms part of the Large Hadron Collider at CERN and was built and designed to detect the decays of b-hadrons. This thesis presents a first observation of the semileptonic decay $B_s^0 \to K^- \mu^+ \nu_\mu$, a tree level decay dependent on $|V_{ub}|$. The decay $B_s^0 \to D_s^- \mu^+ \nu_\mu$ is dependent on $|V_{cb}|$ and the ratio of branching fractions, $\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)/\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)$ is measured. Semileptonic decays containing a light hadron in the final state are beneficial to theoretical physicists as the hadronic and leptonic components of the decay rate can be factorised out. The ratio, $|V_{ub}|/|V_{cb}|$, is obtained by restricting the branching fraction measurement to specific regions in phase space and combining the branching fraction with theoretical predictions calculated using Lattice QCD and light-cone sum rules.

Declaration

The data presented in this thesis was collected by the LHCb experiment at CERN, and I played a major role in the analysis of the data containing $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ decays. All of the analysis presented in the thesis is my own work, apart from the regression model to select a q^2 solution, the choice of stripping selections and preselections, vetoes for the decay $B_s^0 \to K^- \mu^+ \nu_\mu$, and the development of the isolation tool to reject charged backgrounds, detailed in Sections 5.1.3, 5.6.2, 5.6.3 and 5.6.7.

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise, and that this work has not been submitted for any other degree or professional qualification.

(*Iwan Smith, 2019*)

To the memory of my brother, Izaak.

Acknowledgements

Firstly I would like to thank my supervisor Franz Muheim for his guidance, feedback and suggestions over the course of my PhD. His expertise and experience were highly valued during the course of this analysis. Thanks must go to the proponents of the $B_s^0 \to K^-\mu^+\nu_\mu$ analysis, Adlene, Bassem, Marta, Michael, Michel, Mika, and Svende, who's knowledge, and enthusiasm made this a very fulfilling analysis. I am also grateful to all members of the semileptonic working group for their input and feedback.

Theoretical predictions were an essential part of this work, and I am incredibly grateful to Alexander Khodjamirian, Aleksey V. Rusov, and all members of the Fermilab Lattice, MILC, RBC, UKQCD and HPQCD collaboration that worked tirelessly to produce the theoretical form factors needed for this analysis. Additional thanks go to Oliver Witzel and Chris Bouchard for their helpful discussions regarding the $B_s^0 \to K^-\mu^+\nu_\mu$ form factors.

I would like to thank CERN, the LHCb Collaboration, the University of Edinburgh, SUPA and the STFC for providing the infrastructure, training, resources and funding to make my doctoral research possible.

The four years I spent working towards this PhD. were an incredible experience and would not have been the same were it not for the friends I made along the way, both inside and outside the world of physics. I would like to extend a heartfelt thanks to all my friends and family for their support, energy and patience.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The standard model of particle physics is the name given in the 1970s to the theory describing all fundamental particles and the forces governing their interactions. It incorporates all that we know about subatomic particles and has predicted the existence of new particles, most famously the Higgs boson which was discovered in 2012 by the ATLAS and CMS experiments. The 17 particles in the standard model are divided into six quarks, six leptons, four gauge bosons and one scalar boson. The six quarks can be divided into three up-down pairs and the six leptons can be divided into three pairs containing a charged lepton and a neutrino. Quark and lepton pairs are known as flavours. Different quark and lepton pairs behave in exactly the same way and the masses of the charged leptons and quarks originate from their coupling to the Higgs field, with masses varying by five orders of magnitude in the quark sector and three orders of magnitude in the charged lepton sector. It remains unknown why the masses vary to such an extent and why there are exactly three flavours of quarks and leptons.

The standard model allows quarks to change flavour via the charged weak interaction mediated by the $W^{\pm}$ boson, a process that was first observed in 1896 via the radioactive decay of a neutron to proton via the emission of a $W^{\pm}$.

$$n \rightarrow pe^{-}\overline{\nu}_{e}, \tag{1.1}$$

in which a neutron, uud, decays to a proton, udd, The weak force only couples leptons of the same generation and for quarks cross-generational couplings are allowed. The strength of the couplings between quarks are proportional to the

elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$V^{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.2}$$

and the couplings between charged leptons and neutrinos are universal across the generations. The structure of the CKM matrix is nearly diagonal and is illustrated in Figure 1.1, with $|V_{ub}|$ being the smallest and least well known of the elements. It is worth noting that the CKM matrix is unitary,

$$\sum_{i}^{3} V_{ij} V_{ik}^{*} = \delta_{jk} \tag{1.3}$$

which provides an essential test of the Standard Model. The vanishing combinations of Equation 1.3 can be represented as triangles in the complex plane known as CKM unitary triangles.

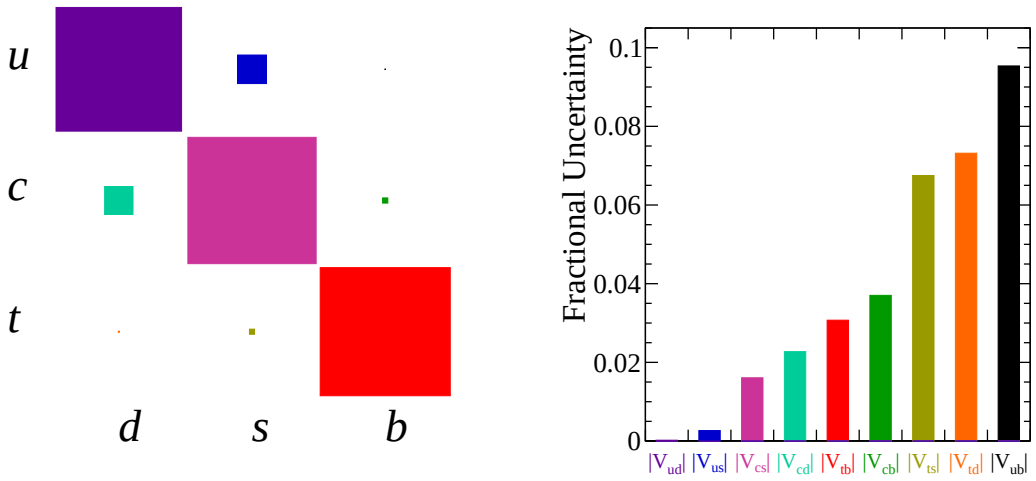


Figure 1.1 *The magnitudes of the CKM matrix elements are illustrated (left) with the almost diagonal structure clearly visible. The fractional uncertainties of the CKM matrix elements are plotted on the right and it can be seen that $|V_{ub}|$ is the smallest element with the largest relative uncertainty.*

The CKM matrix is parametrised by three mixing angles and a complex phase. The complex phase is responsible for all CP violation in the standard model. The CP operator is a product of the charge conjugation operator, $Ce^{-} \rightarrow e^{+}$, and the parity transformation operator, $Px_i \rightarrow -x_i$. CP violation is responsible for the difference in behaviour between matter and antimatter and is required to explain the matter-antimatter asymmetry we observe in the universe. However

the amount of CP violation required to explain the matter-antimatter asymmetry we see today is nine orders of magnitude larger than seen in the quark sector.

In order to test the unitarity of the CKM matrix and precisely measure the amount of CP violation in the quark sector, the parameters of the CKM matrix must be constrained. The CKM parameters can be constrained by performing measurements of observables sensitive to the magnitudes of the CKM matrix elements. Since $|V_{ub}|$ is the least well known of the CKM matrix elements it is the dominant limiting factor when drawing CKM unitary triangles. An improved uncertainty on $|V_{ub}|$ will improve the global precision of fits to the CKM unitary triangles and test the unitarity of the CKM matrix. Non unitarity of the CKM matrix would be indicative of new physics beyond the standard model.

The CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ can be determined from inclusive and exclusive semileptonic decays of a B hadron. When performing an exclusive measurement all visible[1] decay products of the B are reconstructed, and an inclusive decay, $B \to \ell^- \overline{\nu} X$, contains additional unreconstructed final state particles. Inclusive determinations of $|V_{cb}|$ combine measurements of the semileptonic $b \to c\mu^- \overline{\nu}_\mu X$ decay rate with the leptonic energy, the hadronic invariant mass spectra and theoretical calculations. Inclusive measurements of $|V_{cb}|$ were initially performed by the ARGUS and CLEO collaborations. Later came the B factories operating at the $\Upsilon(4S)$ production energy and LEP using B mesons produced from the decays of the Z boson. The B factories had the benefit of higher statistics and produced more precise determinations than LEP while the boosted B mesons from the Z allowed measurements to be made in a larger phase space.

An inclusive measurement of $|V_{ub}|$ is complicated due to the enormous backgrounds originating from $B \to X_c \ell^- \overline{\nu}$ decays. A kinematic approach is usually taken and inclusive measurements are performed in the region where charm backgrounds are kinematically forbidden although statistics can be increased by extending the phase space into the $B \to X_c e^- \overline{\nu}$ region. CLEO, Belle and BaBar have quoted partial rates of $B \to X_c \ell^- \overline{\nu}$ for $|\vec{p_e}| \geq 2.0$ GeV and $|\vec{p_e}| \geq 1.9$ GeV which is well below the charm kinematic endpoint.

Exclusive determinations of $|V_{cb}|$ are based on semileptonic $B \to D^{(*)} \ell^- \overline{\nu}$ decays in the limit $m_{b,c,} \gg \Lambda_{\rm QCD}$. Exclusive measurements of $|V_{ub}|$ are made by combining the exclusive decay rate of B hadrons combined with form factor

[1] A visible particle is reconstructible by the detector. The neutrino is considered invisible.

predictions. The decays $B^0 \to \pi^- \mu^+ \nu_\mu$, $B_s^0 \to K^- \mu^+ \nu_\mu$ and $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$ which contain a ground state hadron in the final state are "golden modes" for lattice QCD predictions and have the lowest theoretical uncertainties.

Form factors provided by lattice QCD are most accurate in the kinematic region with high momentum transfer.

The averaged $|V_{cb}|$ measurements are

$$|V_{cb}|_{Incl} = (42.2 \pm 0.8) \times 10^{-3}, \qquad |V_{cb}|_{Excl} = (41.9 \pm 2.0) \times 10^{-3}, \tag{1.4}$$

and the averaged $|V_{ub}|$ measurements are

$$|V_{ub}|_{Incl} = (4.49 \pm 0.28) \times 10^{-3}, \qquad |V_{ub}|_{Excl} = (3.70 \pm 0.16) \times 10^{-3}. \tag{1.5}$$

The difference between inclusive and exclusive measurements of $|V_{cb}|$ and $|V_{ub}|$ of approximately 3σ has been a long-standing puzzle in particle physics.

The LHC experiment provides an abundance of B hadrons which are detected by the LHCb experiment making exclusive determinations of $|V_{ub}|$ possible with the decays $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$ and $B_s^0 \to K^- \mu^+ \nu_\mu$. This thesis presents a first observation of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ with a measurement of the ratio of branching fractions $\frac{\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)}$ and a ratio of the CKM matrix elements $|V_{ub}|/|V_{cb}|$. This measurement uses data collected from pp collision events collected by the LHCb experiment in the year 2012. The measured ratio of branching fractions is combined with theoretical inputs from Lattice QCD and Light-Cone Sum Rules allowing $|V_{ub}|/|V_{cb}|$ to be determined. This ratio provides an important constraint when performing global fits testing the unitarity of the CKM matrix.

The thesis is structured as follows. Chapter 2 presents an overview of the theoretical framework required for this measurement, including a discussion of the standard model of particle physics and the CKM sector. The theory of semileptonic decays is presented alongside the theory of lattice QCD and the latest form factor predictions for the decays $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ are presented. The LHC and LHCb experiments are introduced in Chapter 3 and the conditions for taking data are discussed. Chapter 4 briefly discusses the analysis strategy for the measurement of the ratio of branching fractions and CKM matrix elements. The main analysis work is presented in Chapters 5 and 6 and discusses the methods used to reconstruct $B_s^0 \to K^- \mu^+ \nu_\mu$ candidates

and separate a signal yield from the many backgrounds present at the LHCb experiment. Chapter 5 details the reconstruction of several non-trivial kinematic distributions essential for this analysis and goes on to detail the modelling of data and the selections used to reject backgrounds. Chapter 6 goes on to detail the fits used to extract the $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yields in data followed by a calculation of the selection efficiencies and systematics, and culminates with the results of $\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)/\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)$ and $|V_{ub}|/|V_{cb}|$. The implications of this measurement on the particle physics landscape is discussed in Chapter 7 which leads to the conclusion of this thesis in Chapter 8.

Chapter 2

Theory

This chapter provides a summary of the standard model of particle physics, and goes on to explain the CKM matrix and its parametrisation. The theory of lattice QCD is presented and the current theoretical predictions for the differential decay rates of $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ are presented.

The standard model is introduced in Section 2.1 and the CKM sector in Section 2.2. The theory behind semileptonic decays and a summary of the form factors used in this analysis are discussed in Section 2.3.

2.1 The Standard Model

The standard model (SM) of particle physics is a single theory describing all the fundamental forces, with the exception of gravity, and their interactions. The theory may be described as an $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge theory where the special unitary group, $\mathrm{SU}(n)$, is a subgroup of the unary group, $\mathrm{U}(n)$. The theories of quantum electrodynamics, QED, and hypercharge are both represented by the unary group $\mathrm{U}(1)_Y$, the electroweak sector and quantum chromodynamics, QCD, are represented by the special unary groups $\mathrm{SU}(2)_L$ and $\mathrm{SU}(3)_c$ respectively [1–7].

2.1.1 Quantum Electrodynamics

The theory of quantum electrodynamics (QED) describes the interactions of charged particles via the exchange of a photon. It is the quantum equivalent of classical electromagnetism and completely models the interactions between light and matter. The Dirac equation,

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{2.1}$$

where ψ is the Dirac spinor, a relativistic spin-$\frac{1}{2}$ field, is a relativistic wave equation describing all massive spin-$\frac{1}{2}$ particles and was the first prediction of antimatter. The QED Lagrangian, $\mathcal{L}_{QED}$ may be defined by taking the Dirac Lagrangian density,

$$\mathcal{L}_{Dirac} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi, \tag{2.2}$$

and demanding local gauge invariance under the transformation

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}\psi(x), \tag{2.3}$$

where α is an arbitrary phase independent of the space time position, x. The derivative, ∂_{μ}, is replaced by a covariant derivative which transforms in exactly the same way as $\psi(x)$,

$$D_{\mu}\psi(x) \rightarrow D'_{\mu}\psi'(x) = e^{i\alpha(x)}D_{\mu}\psi(x), \tag{2.4}$$

and is defined with the introduction of a gauge field, A_{μ},

$$D_{\mu} \equiv \partial_{\mu} + ieA_{\mu}, \tag{2.5}$$

with the introduction of a covariant derivative and addition of a kinetic energy term, $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, where $F_{\mu\nu}$ is the field strength tensor. The QED Lagrangian may be obtained

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi}\gamma^{\mu}D_{\mu}\psi - m\overline{\psi}\psi. \tag{2.6}$$

In the case of Abelian QED the classical result for the electromagnetic field strength is found

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}. \tag{2.7}$$

Local gauge transformations of the Dirac spinors are denoted

$$\psi(x) \to \psi'(x) = U(x)\psi(x), \qquad \overline{\psi}(x) \to \overline{\psi'}(x) = \overline{\psi}(x)U^{-1}(x), \tag{2.8}$$

where, for QED, $U(x)$ is an element of the non-Abelian Lie group U(1) and has the form

$$U(x) = \exp\left(i \sum_{j=1}^{N^2-1} \alpha_j(x) T_j\right), \tag{2.9}$$

where the sum is over the $N^2 - 1$ generators, T, of the group which satisfy the Lie algebra

$$[T_i, T_j] = i c_{ijk} T_k, \tag{2.10}$$

where c_{ijk} are the structure constants of the group. For Abelian groups the generators are commutative resulting in $c_{ijk} = 0$ for the U(1) of QED. The generators for the SU(2) and SU(3) groups involve the three Pauli matrices, $T_i = \sigma_i/2$, and eight Gell-Mann matrices, $T_i = \lambda_i/2$, respectively.

The covariant derivative is defined

$$D^\mu = (\partial_\mu - igA_\mu), \tag{2.11}$$

where g is the gauge coupling. Gauge invariance requires that

$$D^\mu \psi(x) \to D'^\mu \psi'(x) = U(x) \left[D^\mu \psi(x)\right], \tag{2.12}$$

and the transformation of A_μ follows

$$A\mu \to A'_\mu = U(x) A_\mu U^{-1}(x) + \frac{i}{g} U(x) \left[\partial_\mu U^{-1}(x)\right]. \tag{2.13}$$

The locally gauge invariant Lagrangian is obtained from the free Dirac Lagrangian by replacing ∂_μ with D_μ,

$$\mathcal{L} = i\overline{\psi}\gamma^\mu D_\mu \psi - m\overline{\psi}\psi, \tag{2.14}$$

and the non Abelian definition for $F_{\mu\nu}$ follows

$$[D_\mu, D_\nu] = -igF_{\mu\nu}\psi(x), \tag{2.15}$$

yielding the locally gauge invariant kinetic energy term.

2.1.2 Quantum Chromodynamics

Quantum chromodynamics, QCD, is the theory of the strong interaction and models the interactions of quarks via gluon exchange. QCD is a non-Abelian gauge theory with symmetry group SU(N_c) where $N_c = 3$ and contains 8, $N_c^2 - 1$, gluons. The QCD Lagrangian is

$$\mathcal{L}_{QCD} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi + g_s(\overline{\psi}\gamma^\mu T_a \psi)G_\mu^a - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}, \tag{2.16}$$

where $a = 1, 2, 3, \ldots, 8$, the SU(3) generators are $T_a = \lambda_a/2$, the Gell-Mann λ-matrices are λ_a, and $G_{\mu\nu}^a$ is the field strength tensor.

The quark fields carry a QCD analogue of electric charge referred to as colour, R, G, B,

$$\psi(x) = \begin{pmatrix} \psi_R(x) \\ \psi_G(x) \\ \psi_B(x) \end{pmatrix}, \tag{2.17}$$

and transform as a triplet under a local SU(3) gauge transformations

$$\psi(x) \to U(x)\psi(x) = e^{iT^a \alpha_a(x)}\psi(x), \tag{2.18}$$

under which $\mathcal{L}_{QCD}$ is invariant.

2.1.3 The Weak Force and $\mathbf{SU(2)_L \times U(1)_Y}$

The Glashow model couples the SU(2) representation of the weak sector with the U(1) representation of the hypercharge sector where the generators of the $\mathrm{U}(1)_Y$ commute with those of $\mathrm{SU}(2)_L$.

The weak isospin doublet containing a left handed electron and neutrino is defined with it's adjoint

$$\chi_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \equiv \begin{pmatrix} \nu \\ e \end{pmatrix}_L, \qquad \overline{\chi}_L = \begin{pmatrix} \overline{\nu}_L & \overline{e}_L \end{pmatrix} \tag{2.19}$$

with generators satisfying the SU(2) Lie Algebra

$$\left[\frac{1}{2}\alpha^i, \frac{1}{2}\alpha^j\right] = i\epsilon_{ijk}\frac{1}{2}\alpha^k, \tag{2.20}$$

where α^i are the Pauli matrices. The doublet has an isospin quantum number, $T = \frac{1}{2}$, and the upper and lower components of the doublet have $T^3 = +\frac{1}{2}, -\frac{1}{2}$ respectively. The isospin triplet of weak currents, J^1_μ and J^2_μ, couple the electron to the neutrino and the current J^3_μ couples the electron or neutrino to itself,

$$J^i_\mu = \overline{\chi}_L \gamma_\mu \frac{1}{2} \alpha^i \chi_L \quad (i = 1, 2, 3), \tag{2.21}$$

The electromagnetic current

$$J^{em}_\mu = Q(\overline{e}_L \gamma_\mu e_L + \overline{e}_R \gamma_\mu e_R) \tag{2.22}$$

where Q is the charge of the particle may be expressed in terms of the weak current, J^3_μ, and an additional current J^Y_μ which includes a coupling to the right handed electron

$$J^{em}_\mu = J^3_\mu + \frac{1}{2} J^Y_\mu, \tag{2.23}$$

yielding

$$J^Y_\mu = -\overline{\nu}_L \gamma_\mu \nu_L - \overline{e}_L \gamma_\mu e_L - 2\overline{e}_R \gamma_\mu e_R. \tag{2.24}$$

The identity between J^{em}_μ, J^3_μ and J^Y_μ given in Equation 2.24 yields the Gell-Mann Nishijima relation corresponding to electric charge, Q, the third component of isospin, T^3 and hypercharge, Y,

$$Q = T^3 + \frac{1}{2} Y. \tag{2.25}$$

The three generations of leptons all consist of the same weak isospin doublet with the same quantum numbers,

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \tag{2.26}$$

and their charges are given in Table 2.1.

Vector fields coupling the currents detailed above must be included to ensure the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge theory is invariant under local gauge transformations. An isotriplet of gauge bosons, W^i_μ, $(i = 1, 2, 3)$, is introduced to gauge the $\mathrm{SU}(2)_L$ symmetry with coupling strength, g, and a vector boson B_μ is introduced to gauge the $\mathrm{U}(1)_Y$ symmetry with coupling strength, $g'/2$. The lepton-gauge boson portion of the Lagrangian, $\mathcal{L}(l)$, couples vector boson fields to the weak isospin doublet and the right hand lepton to the vector boson B_μ. The full Lagrangian

Particle	T	T^3	Q	Y
$\nu_e,\ \nu_\mu,\ \nu_\tau$	$\frac{1}{2}$	$\frac{1}{2}$	0	-1
$e_L^-,\ \mu_L^-,\ \tau_L^-$	$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1
$e_R^-,\ \mu_R^-,\ \tau_R^-$	0	0	-1	-2
$u_L,\ c_L,\ t_L$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{3}$
$d_L,\ s_L,\ b_L$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{3}$	$\frac{1}{3}$
$u_R,\ c_R,\ t_R$	0	0	$\frac{2}{3}$	$\frac{4}{3}$
$d_R,\ s_R,\ b_R$	0	0	$-\frac{1}{3}$	$-\frac{2}{3}$

Table 2.1 *The fermion charge assignments for weak isospin, T, it's third component, T^3, electric charge, Q, and hypercharge Y.*

contains the sum over the three generations of lepton, $\sum_{l=e\mu\tau}\mathcal{L}(l)$.

The interaction part of $\mathcal{L}(l)$

$$\mathcal{L}_I = \overline{\chi}_L\gamma^\mu\left[-g\frac{1}{2}\vec{\tau}\cdot\vec{W}_\mu + \frac{1}{2}g'B_\mu\right]\chi_L + \overline{e}_R\gamma^\mu g'B_\mu e_R, \tag{2.27}$$

may be decomposed into a charged and neutral current corresponding to the physical $W^\pm$ and Z bosons respectively,

$$\mathcal{L}_I = \mathcal{L}_{CC} + \mathcal{L}_{NC}. \tag{2.28}$$

The charged vector fields are defined as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \tag{2.29}$$

and the neutral vector fields Z_μ and A_μ are an orthogonal mixture of W_μ^3 and B_μ with weak mixing angle θ_w,

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix}\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix}. \tag{2.30}$$

The interaction part of the lagrangian may now be written in terms of the full fermion fields and physical gauge bosons

$$\begin{aligned}\mathcal{L}_I = &-\frac{1}{\sqrt{2}}\left[\overline{\nu}\gamma_\mu\frac{1}{2}(1-\gamma_5)eW^{+\mu} + \overline{e}\gamma_\mu\frac{1}{2}(1-\gamma_5)\nu W^{-\mu}\right] + e(\overline{e}\gamma_\mu e A^\mu) \\ &-\frac{g}{2\cos\theta_\omega}\left[\overline{\nu}\gamma_\mu\frac{1}{2}(1-\gamma_5\nu) - \overline{e}\gamma_\mu\frac{1}{2}(1-\gamma_5)e + 2\sin^2\theta_\omega\overline{e}\gamma_\mu e\right]Z^\mu,\end{aligned} \tag{2.31}$$

where the coefficients of the $\bar{l}lV$ ($l = e, \nu,\ \ V = A, W^{\pm}, Z$) components gives the fermion-gauge boson vertex factors. The complete Glashow model Lagrangian requires the kinetic energy terms for the W^i_μ and B_μ fields

$$\begin{aligned} \mathcal{L}_W &= -\frac{1}{4}\vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu}, \\ \mathcal{L}_B &= -\frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \end{aligned} \tag{2.32}$$

which may be expressed in terms of the physical fields defined in Equations 2.29 and 2.30. The full Lagrangian for the Glashow model Lagrangian may be expressed,

$$\mathcal{L} = \sum_{l=e,\mu,\tau} \mathcal{L}(l) + \mathcal{L}_W + \mathcal{L}_B, \tag{2.33}$$

which contains no mass terms.

2.1.4 Electroweak Symmetry Breaking

In order to introduce mass terms for the $W^{\pm}$ and Z fields a Higgs doublet,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \qquad \Phi^\dagger = \begin{pmatrix} \phi^- & \overline{\phi^0} \end{pmatrix}, \tag{2.34}$$

must be included. The addition of the Yukawa couplings between the Higgs and the fermions provides the mechanism for generating fermion masses and the observed flavour structure of the CKM sector of the standard model. The covariant derivative for the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry is defined

$$D_\mu = \partial_\mu + \frac{i}{2}g\vec{\tau} \cdot \vec{W}_\mu + g'\frac{i}{2}B_\mu. \tag{2.35}$$

The Higgs Lagrangian,

$$\mathcal{L}_\Phi = -(D_\mu\Phi)^\dagger D^\mu\Phi - V(\Phi), \tag{2.36}$$

is added to the Glashow model of the Lagrangian given in Equation 2.33 with the scalar Higgs potential defined as,

$$V(\Phi) = \mu^2(\Phi^\dagger\Phi) - \lambda(\Phi^\dagger\Phi)^2, \tag{2.37}$$

which has a minima specified by

$$\frac{dV}{d(\Phi^\dagger\Phi)} = 0 \quad \Rightarrow \quad \mu^2 - 2\lambda(\Phi^\dagger\Phi) = 0 \quad \Rightarrow \quad \Phi^\dagger\Phi\big|_{min} = \frac{\mu^2}{2\lambda}. \tag{2.38}$$

The $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ may be spontaneously broken by choosing an arbitrary vacuum from the set of minima of the potential V. Without any loss of generality, this may be chosen as

$$\langle\Phi\rangle = \begin{pmatrix} 0 \\ \frac{\nu}{\sqrt{2}} \end{pmatrix}, \tag{2.39}$$

where ν is the vacuum expectation value of the Higgs Field and was found experimentally to be, $\nu = 246$ GeV [8, 9]. The unitary gauge is defined when the field Φ is expanded around this chosen vacuum,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H + \nu \end{pmatrix}, \tag{2.40}$$

where H is the neutral scalar Higgs field. In the unitary gauge "Goldstone" fields with zero vacuum expectation values are eliminated.

Evaluating the Higgs Lagrangian in the unitary gauge, one finds

$$\begin{aligned} \mathcal{L}_\Phi &= (D_\mu\Phi)^\dagger D_\mu\Phi + \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2 \\ &= \frac{1}{2}\partial_\mu H\partial^\mu H + \frac{1}{4}g^2(H^2 + 2\nu H + \nu^2)W\mu^+W^{-\mu} \\ &+ \frac{1}{8}(g^2 + g'^2)(H^2 + 2\nu H + \nu^2)Z_\mu Z^\mu \\ &+ \mu^2H^2 + \frac{\lambda}{4}(H^4 + 4\nu H^3), \end{aligned} \tag{2.41}$$

where the masses of the $W^\pm$ and Z may be read off by identifying the coefficients of the $W^+_\mu W^{-\mu}$ and $Z_\mu Z^\mu$ terms. One finds,

$$\begin{aligned} M_W &= \frac{1}{2}g\nu \\ M_Z &= \frac{1}{2}(g^2 + g'^2)^{1/2}\nu = \frac{1}{2}\frac{g\nu}{2\cos\theta_\omega}, \end{aligned} \tag{2.42}$$

yielding the famous relation between the masses of the vector bosons and the weak mixing angle,

$$\frac{M_W}{M_Z} = \cos\theta_\omega. \tag{2.43}$$

2.1.5 Yukawa Coupling and Leptons

Fermion masses are provided by the Yukawa coupling, which couples the fermion fields to the Higgs field. Take the electron,

$$\mathcal{L}_Y = -G_e \left[\overline{\chi}_L \Phi e_R + \overline{e}_R \Phi^\dagger \chi_L \right], \tag{2.44}$$

where the Higgs field may be substituted in using the unitary gauge given in Equation 2.40,

$$\begin{aligned} \mathcal{L}_Y(e) &= -\frac{G_e}{\sqrt{2}}(\nu + H)(\overline{e}_L e_R + \overline{e}_R e_L) \\ &= -\frac{G_e}{\sqrt{2}}(\nu + H)\overline{e}e = -\frac{G_e \nu}{\sqrt{2}}(\overline{e}e) - \frac{G_e}{\sqrt{2}}(\overline{e}eH), \end{aligned} \tag{2.45}$$

from which one can read off the electron's mass, $m_e = G_e\nu/2$, and the lepton Higgs coupling, $g(H\overline{e}e) = m_e/\nu = gm_e/(2M_W)$. It should be noted that the coupling between the leptons and the Higgs is proportional to the lepton mass.

2.1.6 Yukawa Coupling and Quarks

An $\mathrm{SU}(2)_L$ Isospin doublet analogous to the lepton case is created containing an up type quark and an admixture of the down type quarks

$$\chi_L^f = \begin{pmatrix} U_f \\ D'_f \end{pmatrix}, \quad f = 1, 2, 3 \tag{2.46}$$

where $U_1 = u$, $U_2 = c$, $U_3 = t$ and $D_1 = d$, $D_2 = s$, $D_1 = b$ and D'_f is the eigenstate of the weak interaction which is a rotated mixture of the flavour eigenstates

$$D'_f = \sum_{f'=1,2,3} V_{ff'} D_{f'}. \tag{2.47}$$

V is the 3×3 unitary Cabibbo-Kobyashi-Maskawa (CKM) matrix [10] and describes the coupling strengths of the quarks. The charged $W^\pm$ interactions couple to the physical u_{Lj} and d_{Lk} quarks as

$$\frac{-g}{\sqrt{2}} = \left(\overline{u_L}, \overline{c_L}, \overline{t_L}\right) \gamma^\mu W_\mu^+ V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + h.c., \tag{2.48}$$

where

$$V_{CKM} \equiv V_L^u V_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.49}$$

Following a similar process to that outlined in Section 2.1.4, the electromagnetic and hypercharge currents may be defined

$$\begin{aligned} J_\mu^{f(em)} &= \left(\frac{2}{3}\right) \overline{U}_{fR}\gamma_\mu U_{fR} + \left(\frac{2}{3}\right) \overline{U}_{fL}\gamma_\mu U_{fL} + \left(\frac{-1}{3}\right) \overline{D}_{fR}\gamma_\mu D_{fR} + \left(\frac{-1}{3}\right) \overline{D}_{fL}\gamma_\mu D_{fL} \\ J_\mu^{fY} &= \left(\frac{1}{3}\right) (\overline{U}_{fL}\gamma_\mu U_{fL} + \overline{D}_{fL}\gamma_\mu D_{fL}) + \left(\frac{4}{3}\right) \overline{U}_{fR}\gamma_\mu U_{fR} + \left(\frac{-2}{3}\right) \overline{D}_{fR}\gamma_\mu D_{fR}, \end{aligned} \tag{2.50}$$

where the numbers in brackets denote the charges and hypercharges of the quarks.

The quark electroweak Lagrangian is defined as

$$\begin{aligned} \mathcal{L}(q) = \sum_{f=1,2,3} \Bigg(& \overline{\chi}_L^f \gamma_\mu \left[i\partial_\mu - \frac{1}{2}\vec{\tau}\cdot\vec{W}_\mu - \left(\frac{1}{3}\right) B_\mu \right] \chi_L^f \\ & + \overline{U}_{fR}\gamma^\mu \left[i\partial_\mu - \frac{g'}{2}\left(\frac{4}{3}\right) B_\mu \right] U_{fR} + \overline{D}_{fR}\gamma_\mu \left[i\partial_\mu - \frac{g'}{2}\left(\frac{-2}{3}\right) B_\mu \right] D_{fR} \Bigg), \end{aligned} \tag{2.51}$$

with masses originating from the quark Yukawa term

$$\mathcal{L}_Y(q) = \sum_{f=1,2,3} - \left[\overline{\chi}_L^f G_{ff'}^D \Phi D_{f'R} + \overline{\chi}_L^f G_{ff'}^U \Phi^c U_{f'R} + \text{h.c.} \right], \tag{2.52}$$

where $G_{ff'}^U$ and $G_{ff'}^D$ are matrices of the couplings between the quark and Yukawa fields. The conjugate Higgs scalar field, Φ^c after spontaneous symmetry breaking is given in the unitary gauge by

$$\Phi^c = \begin{pmatrix} \overline{\phi}^0 \\ -\phi^- \end{pmatrix} = \begin{pmatrix} H + \nu \\ 0 \end{pmatrix}. \tag{2.53}$$

The quarks in Equation 2.52 yield mass terms when ϕ acquires a vacuum expectation value, $\langle\phi\rangle = (0, \nu/\sqrt{2})$

2.2 The CKM Sector

2.2.1 The CKM Matrix

The CKM matrix can be parametrised by three mixing angles and a complex phase, with the standard convention being

$$\begin{aligned} V_{CKM} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \end{aligned} \tag{2.54}$$

where $s_{ij} = \sin(\theta_{ij})$, $c_{ij} = \cos(\theta_{ij})$ and δ is the phase responsible for all CP-violation in flavour changing phenomena in the standard model. Using this formalism V_{ub} and V_{cb} are defined as,

$$V_{ub} = s_{13}e^{-i\delta}, \quad V_{cb} = s_{23}c_{13}. \tag{2.55}$$

The exact formalism given in Formula 2.54 is a little unwieldy so an approximation is made which better captures the essential physics of the CKM matrix. The first approximation was made by Wolfenstein after he noticed that the orders of magnitude of the CKM matrix visualised in Figure 1.1 follow a pattern:

$$|V_{CKM}| \sim \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}, \tag{2.56}$$

where $\lambda \approx 0.2$. This was refined by the addition of three different real parameters, A, ρ, η, all $\mathcal{O}(1)$. The Wolfenstein parameters can be defined in terms of the standard parameters.

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \tag{2.57}$$

$$c_{13} = A\lambda^2 = \lambda \left| \frac{V_{cb}}{V_{us}} \right|, \tag{2.58}$$

$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = V_{ub}^* = \frac{A\lambda^3(\overline{\rho} + i\overline{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4(\overline{\rho} + i\overline{\eta})]}, \tag{2.59}$$

where $\overline{\rho} + i\overline{\eta} = -V_{ud}V_{ub}^*/V_{cd}V_{cb}^*$ and does not depend on ones choice of definition for the CKM phase. ρ and η are non-exact expansions of $\overline{\rho}$ and $\overline{\eta}$, e.g. $\overline{\rho} = \rho(1 - \lambda^2/2 + ...)$.

Using the Wolfenstein parametrisation the CKM matrix can be expressed as

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.60}$$

Attention should be drawn to the 2×2 matrix at the top left of the Wolfenstein parametrisation which is a first order expansion of the 2D rotation matrix. This is the Cabibbo mixing matrix [11, 12] and its inclusion informs us that to first order in λ the first two generations of quarks do not know about the third. It should be noted that complex numbers only appear in the $3-1$ matrix elements, which has the curious feature of removing CP violation from the Kaon system. Curious as the first direct of observation of CP violation was in the decays of neutral kaons. Finally it should be noted that this parametrisation of the CKM matrix is not unitary! Both of the above quirks can be resolved by extending the parametrisation to higher powers in λ.

The unitarity of the CKM matrix,

$$\sum_k V_{ik}V_{jk}^* = \delta_{ij}, \quad \sum_k V_{ki}V_{kj}^* = \delta_{ij}, \tag{2.61}$$

provides an essential test of the standard model. The six vanishing relations given in Equation 2.61 can be plotted, forming triangles in the complex plane which are called unitary triangles. The most interesting unitary triangle

$$\sum_i V_{id}V_{ib}^* = V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.62}$$

has sides with lengths $\mathcal{O}(\lambda^3)$ and is the most *triangular looking*. When plotting the unitary triangle it is customary to normalise the sides,

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0, \tag{2.63}$$

the geometrical interpretation of which is plotted in Figure 2.1.

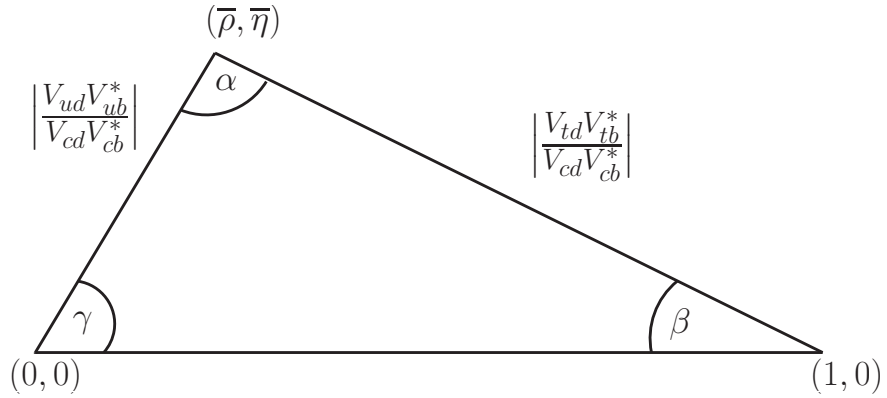


Figure 2.1 *The unitary triangle given in Equation 2.63 is plotted in the complex plane.*

All of the unitary triangles have the same area, $J/2$. The Jarlskog invariant, J, is a basis independent way of quantifying the amount of CP violation in the quark sector, and is given by

$$\Im[V_{ij}V_{kl}V_{il}^*V_{kj}^*] = J\sum_{mn}\epsilon_{ikm}\epsilon_{jln}, \tag{2.64}$$

which can be expressed using the mixing angles and Wolfenstein parameters

$$\begin{aligned} J &= c_{12}c_{23}c_{13}^2 s_{12}s_{23}s_{13}\sin(\delta) \\ &\approx \lambda^6 A^2 \eta. \end{aligned} \tag{2.65}$$

2.2.2 Constraining the CKM sector

The parameters of the CKM matrix can be overconstrained by making measurements of key observables which are sensitive to combinations of the magnitudes and the phases of the matrix elements. This serves to improve the determination of the CKM elements and could reveal the effects of physics beyond the standard model. The magnitudes of the matrix elements are a determining factor in the rates of semileptonic and leptonic decays and the phases of the CKM elements can be determined by measuring processes susceptible to the effects of oscillation and CP violation.

The limiting factor when performing global fits to the CKM matrix originates from the uncertainty on the magnitude of V_{ub}. The length of the side of the unitary triangle opposite the angle β is proportional to $|V_{ub}|/|V_{cb}|$ and an improved

measurement of this ratio could significantly improve CKM fits. The side of the triangle opposite γ is dependent on the magnitudes of V_{tb} and V_{td} which have large uncertainties. The side of the triangle opposite γ is more precisely constrained by measuring the mass difference between the B^0 and $\overline{B}^0$ mesons. The $B^0\overline{B}^0$ oscillation frequency is driven by the mass difference, Δm_d, which is related to the combination of CKM elements, $\Delta m_d \propto |V_{td}V_{tb}^*|^2$. And similarly for $B_s^0\overline{B}_s^0$ mixing, $\Delta m_s \propto |V_{ts}V_{tb}^*|^2$. Frequently, the ratio of the mass differences, $\Delta m_d/\Delta m_s \propto |V_{td}V_{tb}^*|^2/|V_{ts}V_{tb}^*|^2$ is used as a constraint as the theoretical uncertainties cancel in the ratio producing a parameter with a significantly improved uncertainty. The leading source of uncertainty when determining the magnitudes of the CKM elements and their combinations come from the theoretical uncertainty on the form factors which encompass the nature of QCD.

Consider the decay of a neutral B^0 meson to a final state f, the decay can proceed as $B^0 \to f$ or $\overline{B}^0 \to B^0 \to f$. If f is a CP eigenstate and the decay amplitudes from one CKM phase dominate the decay, the time dependent CP asymmetry can be written

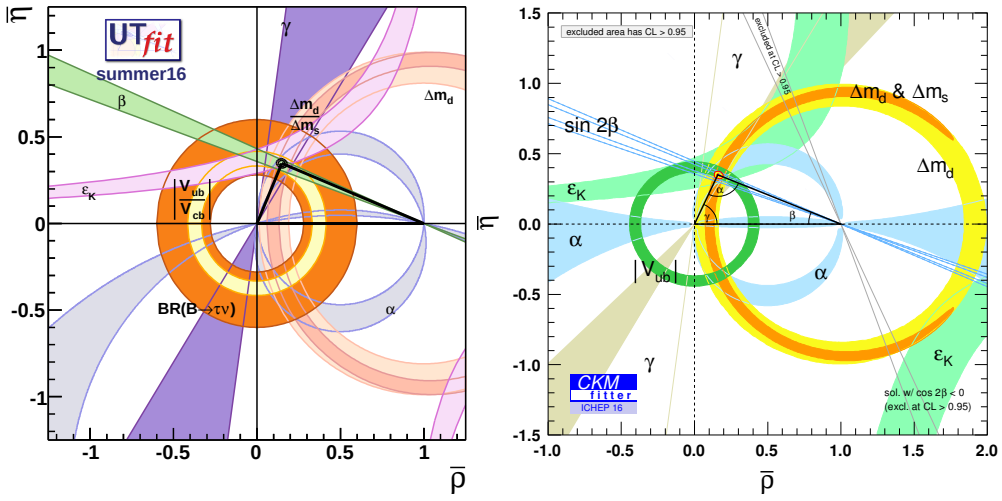
$$\mathcal{A}_f = \frac{\Gamma(\overline{B}^0(t) \to f) - \Gamma(B^0(t) \to f)}{\Gamma(\overline{B}^0(t) \to f) + \Gamma(B^0(t) \to f)} = \eta_f \sin(2\beta)\sin(\Delta m_d t), \tag{2.66}$$

where η_f is the CP eigenvalue of f. A measurement of $\sin(2\beta)$ can be performed by measuring the time dependent decay rates of $B^0 \to f$ and $\overline{B}^0 \to f$ using the transitions $b \to c\bar{c}s$, $b \to c\bar{c}d$ and $b \to c\bar{u}d$ with CP eigenstates to the same final state. Measurements have been performed using the decays, $B^0 \to J/\psi\, K^0{}_{S/L}$ and $B^0 \to J/\psi\, \pi^0$. There is a factor four ambiguity in β from $\sin(2\beta)$ which can be removed by performing a global fit to the unitary triangle.

The angle α is the phase between $V_{ub}^*V_{td}$ and $V_{ub}^*V_{ud}$ and can only be measured from time dependent CP asymmetries of $b \to u\bar{u}d$ decays. Penguin contributions from $b \to d$ decays are the same order in λ as the tree level decay and are a sizeable contribution of the decay rate. α has been measured in the decays, $B \to \pi\pi$, $B \to \rho\pi$ and $B \to \rho\rho$. The angle γ, unlike α and β does not depend on CKM elements coupling to the top quark. Consequently it can be measured from tree level decays of the B and is unlikely to be affected by new physics beyond the standard model. The angle γ may be determined by measuring the interference in the decays $B^- \to \overline{D}^0K^-$ and $B^- \to D^0K^-$ with the D^0 and $\overline{D}^0$ decaying to the same final state [13–16].

The most precise determinations of the CKM matrix elements come from global

fits to all available measurements and by imposing the constraints of the standard model. There are several approaches used to combine the data, the two best known come from the UTfit [17–19] and CKMfitter [20, 21] collaborations which use Bayesian and frequentist statistics respectively. The results from both collaborations are compatible and the fits to the CKM parameters are plotted in Figure 2.2.



$\overline{\rho} = 0.153 \pm 0.013$, $\overline{\eta} = 0.343 \pm 0.011$. $\overline{\rho} = 0.1598^{+0.0076}_{-0.0072}$, $\overline{\eta} = 0.3499^{+0.0063}_{-0.0061}$.

Figure 2.2 *Constraints on the $\overline{\rho}, \overline{\eta}$ plane from UTfit (left) and CKMfitter (right). Images taken from [18, 21]*

2.3 Semileptonic B meson Decays

In order to extract the electroweak parameters $|V_{ub}|$ and $|V_{cb}|$ from the physically observable decay rates hadronic form factors are required. This section will present an overview of the current form factor calculations for $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$. In a scattering interaction the form factor modifies the point-like model of the interaction to consider the spatial extent and shape of the interacting particles.

The amplitude of the semileptonic decay $B_s^0 \to K^- \mu^+ \nu_\mu$ can be written as a term proportional to the product of a leptonic current L^μ and a hadronic current H_μ [22]. When $q^2 \ll m_{W^\pm}^2$ the matrix element, $\mathcal{M}$, of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$

may be written

$$\begin{aligned}\mathcal{M} &= -i\frac{G_F}{\sqrt{2}}V_{ub}L^{\mu}H_{\mu} \\ &= -i\left\langle K^{-}(p')\mu^{+}(k')\nu_{\mu}(k)|\mathcal{H}_{\text{eff}}|B_s^0\right\rangle,\end{aligned} \tag{2.67}$$

where $\mathcal{H}_{eff}$ is the effective Hamiltonian, and G_F is the Fermi coupling constant. The leptonic current is

$$L^{\mu} = \overline{\mu}\gamma^{\mu}(1-\gamma_5)\nu_{\mu}, \tag{2.68}$$

and the hadronic current is

$$H_{\mu} = \left\langle K^{-}(p')|\overline{u}\gamma^{\mu}b|B_s^0(p)\right\rangle - \left\langle K^{-}(p')|\overline{u}\gamma^{\mu}\gamma_5 b|B_s^0(p)\right\rangle, \tag{2.69}$$

which leads to an effective Hamiltonian of

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}}V_{ub}^{L}\left[\overline{u}\gamma^{\mu}b - \overline{u}\gamma^{\mu}\gamma_5 b\right]\mu^{+}\gamma_{\mu}(1-\gamma_5)\nu_{\mu}, \tag{2.70}$$

where $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ which separates ψ into left and right handed currents, $\psi_L = \frac{1-\gamma_5}{2}\psi$ and $\psi_R = \frac{1+\gamma_5}{2}\psi$. Since $B_s^0 \to K^{-}\mu^{+}\nu_{\mu}$ is a pseudoscalar meson transition, $B_s^0(J^P = 0^-) \to K^-(J^P = 0^-)$, the axial-vector component of H_{μ} is zero due to constraints on the spin of the outgoing u quark. The vector component of H^{μ} is parametrised by the vector and scalar form factors, f_+ and f_0, and may be written as:

$$\left\langle K^{-}(p\prime)|\overline{u}\gamma^{\mu}b|B_s^0(p)\right\rangle = f_+(q^2)\left(p^{\mu} + p'^{\mu} - \frac{m_{B_s^0}^2 - m_{K^-}^2}{q^2}q^{\mu}\right) + f_0(q^2)\frac{m_{B_s^0}^2 - m_{K^-}^2}{q^2}q^{\mu}, \tag{2.71}$$

where $q^{\mu} = p_{B_s^0}^{\mu} - p_{K^-}^{\mu}$ is the momentum transfer. The determination of the vector and scalar form factors, f_+, f_0 are given in Section 2.3.2. The vector form factor parametrises transitions mediated by a vector boson, such as the $W^{\pm}$, and the scalar form factor parametrises transitions mediated by a scalar boson. As the decays $B_s^0 \to K^{-}\mu^{+}\nu_{\mu}$ and $B_s^0 \to D_s^{-}\mu^{+}\nu_{\mu}$ are mediated by the $W^{\pm}$ boson, the scalar form factor is heavily suppressed and its contributions are negligible. Decays coupling to the τ and new physics models with scalar states couple have an increased dependence on the scalar form factor.

2.3.1 Lattice QCD

Lattice QCD is a non perturbative method for solving the QCD action

$$S_{QCD} = \int d^4x \mathcal{L}_{QCD}, \tag{2.72}$$

numerically via the discretisation of space and time [23–25]. Consider a particle traversing the quantum mechanical path, $x(t)$ in time, t, between $x(0)$ and $x(t_f)$. Quantum mechanically the particle can be seen as traversing all possible paths with the probability of a given path proportional to $\exp(-\int dt\mathcal{L})$. The expectation value of an operator combination is known as a correlation function and is calculated according to

$$\langle \mathcal{O}(x(t_1)x(t_2)\rangle = \frac{\int \mathcal{D}x(t)\mathcal{O}(x(t_1)x(t_2))e^{-\int dt\mathcal{L}}}{\int \mathcal{D}x(t)e^{-\int dt\mathcal{L}}}, \tag{2.73}$$

where $\int \mathcal{D}x(t)$ is used to denote an integral over all possible paths $x(t)$. The expectation value may be solved numerically using a one-dimensional lattice in time with spacing a. Hybrid Monte Carlo methods [26] are used to generate large combinations of N_{conf} lattice configurations. Each configuration corresponds to a different path along the lattice where the probability of finding a given configuration is proportional to $\exp(-\int dt\mathcal{L})$. The calculation of the correlation function using LQCD is the discretised sum over all configurations

$$\langle \mathcal{O}(x(t_1)x(t_2)\rangle = \frac{1}{N_{\text{conf}}} \sum_{n=1}^{N_{\text{conf}}} \mathcal{O}(x(t_1), x(t_2)). \tag{2.74}$$

The corresponding statistical uncertainty of the expectation of the correlation functions is proportional to $1/\sqrt{N_{\text{conf}}}$.

In addition to the statistical uncertainty there are several sources of systematic errors which must be quantified:

- Extrapolation to the continuum limit: The results of calculations must be extrapolated to a lattice spacing of zero, $a \to 0$, using knowledge of the functional form of discretisation errors.
- Extrapolation to infinite volume: Lattice QCD calculations cover finite volume of space while the true quantum mechanical treatment integrates over an infinite volume of space time resulting in a shift away from the true

value.

- Chiral extrapolation: The mass of the pion varies between lattice configurations requiring an extrapolation to the true value.
- Operator matching: Operators defined in lattice calculations must be matched to those from the quantum mechanical integral using renormalisations requiring non-perturbative techniques which come with systematic uncertainties.
- Quark mass extrapolation: LQCD simulations use quark masses above the true masses requiring an extrapolation to the true value.
- B_s^0 mass fits: During the calculation of form factors the ground state B_s^0 mass is determined by fitting the 2-point correlation function, which may be different to the experimentally measured B_s^0 mass.

2.3.2 $B_s^0 \to K^-\mu^+\nu_\mu$ Form Factors

The current non perturbative methods for the calculations of form factors for $B_s^0 \to K^-\mu^+\nu_\mu$ include lattice QCD [24, 25] and light-cone sum rules [27]. The two calculation methods provide predictions which are complimentary in phase space, calculations from lattice QCD are most precise at high values of q^2 and calculations from light-cone sum rules are most precise at low values of q^2. Lattice QCD and light-cone sum rules calculations are performed using Monte Carlo simulations and the *cost* of generating Monte Carlo data at low q^2 is too high to be useful for LQCD, and vice versa for LCSR. For LQCD calculations there is typically no Monte Carlo data below $q^2 = 13\ \text{GeV}^2/c^4$ and for LCSR calculations there is typically no Monte Carlo data above $q^2 = 13\ \text{GeV}^2/c^4$. Despite the lack of data, requirements on unitarity and analyticity can be used to extrapolate form factor results into the regions with no Monte Carlo data. The decay $B_s^0 \to K^-\mu^+\nu_\mu$ is normalised to the decay $B_s^0 \to D_s^-\mu^+\nu_\mu$ for which form factor calculations from LQCD are available. Due to tighter kinematic theoretical constraints at low q^2, the form factor calculations for $B_s^0 \to D_s^-\mu^+\nu_\mu$ need not be restricted to high q^2 momentum transfer and the full phase space in q^2 is used.

The differential decay rate for $B_s^0 \to K^-\mu^+\nu_\mu$ in the B_s^0 rest frame is given by

$$\frac{d\Gamma(B_s^0 \to K^-\mu^+\nu_\mu)}{dq^2} = \frac{G_F^2|V_{ub}|^2}{24\pi^3}\frac{(q^2-m_\mu^2)^2\sqrt{E_K^2-m_K^2}}{q^4 m_{B_s^0}^2}\times$$
$$\left[\left(1+\frac{m_\mu^2}{2q^2}\right)m_{B_s^0}^2(E_K^2-m_K^2)|f_+(q^2)|^2 + \frac{3m_\mu^2}{8q^2}(m_{B_s^0}^2-m_K^2)^2|f_0(q^2)|^2\right], \quad (2.75)$$

which in the limit, $m_\mu^2 \ll q^2$, becomes

$$\frac{d\Gamma(B_s^0 \to K^-\mu^+\nu_\mu)}{dq^2} = \frac{G_F^2|V_{ub}|^2}{24\pi^3}(E_K^2-m_K^2)^{3/2}|f_+(q^2)|^2, \quad (2.76)$$

where G_F is the Fermi coupling constant, q is the momentum transfer or the invariant mass of the muon and neutrino, m_{μ,K,B_s^0}, are the masses of the muon, kaon and B_s^0 respectively. $|f_+|$ and $|f_0|$ are the vector and scalar form factors which parametrise the hadronic contributions to the electroweak decay and are calculated nonperturbatively using either lattice QCD or Light-Cone Sum Rules.

The form factors are parametrised using the BCL parametrisation detailed in reference [28] and formalised in Equation 2.79. The BCL parametrisation has K degrees of freedom where $K = 2, 3$, and is parametrised to the variable, z,

$$z = (q^2, t_0) = \frac{\sqrt{1-q^2/t_+}-\sqrt{1-t_0/t_+}}{\sqrt{1-q^2/t_+}+\sqrt{1-t_0/t_+}}, \quad (2.77)$$

where,

$$\begin{aligned} t_0 &= (m_{B_s^0}+m_{K^-})\cdot(\sqrt{m_{B_s^0}}-\sqrt{m_{K^-}})^2, \\ t_\pm &= (m_{B_s^0}\pm m_{K^-})^2. \end{aligned} \quad (2.78)$$

The $K = 3$ BCL parametrisation [29] for the vector and scalar form factors are

$$\begin{aligned} f_+(q^2) &= \frac{1}{1-q^2/m_{B^*}^2}\sum_{k=1}^{K-1} b_+^{(k)}\left[z^k - (-1)^{k-K}\frac{k}{K}z^k\right], \\ f_0(q^2) &= \frac{1}{1-q^2/m_{B^*}^2}\sum_{k=1}^{K-1} b_0^{(k)} z^k, \end{aligned} \quad (2.79)$$

where a pole is included at the theoretically predicted $m_{B^*} = 5.63$ GeV [30]. The $K = 2$ BCL parametrisation for the vector and scalar form factors are

$$f_{+,0}(q^2) = \frac{f_{+,0}(0)}{1-q^2/m_{B^*}^2}\left\{1 + b_{+,0}^{(1)}\left[z(q^2)-z(0)+\frac{1}{2}\left(z(q^2)^2-z(0)^2\right)\right]\right\}. \quad (2.80)$$

At low q^2 the vector and scalar form factors may be described by a single independent form factor

$$f_0(q^2) = \frac{m_{B_s^0}^2 - q^2}{m_{B_s^0}^2} f_+(q^2). \tag{2.81}$$

The coefficients, $b_{+,0}^{(k)}, f_{+,0}(0)$, for all models discussed in this section are given in Appendix A.

Three form factor calculations for $B_s^0 \to K^- \mu^+ \nu_\mu$ are used in the determination of $|V_{ub}|/|V_{cb}|$. Predictions from lattice QCD provide a precise determination of the form factors at high q^2 and are provided by Bouchard et al. [31] and Flynn et al. [30]. Calculations from light-cone sum rules are most precise at low q^2 and are provided by Khodjiamirian and Rusov (K&R) [32].

The predicted form factors are plotted in Figure 2.3 and the predicted decay rates are plotted in Figure 2.4. The results of the form factor calculations are given at the end of this section.

Attention should be drawn to the discrepancies at low q^2, the two lattice QCD calculations differ significantly, and the consensus within the theoretical community is that the systematic uncertainties are underestimated. Additionally the light cone sum rules calculations differ significantly from the lattice QCD calculations at $q^2 = 13$ GeV$^2/c^4$, the region at which predictions from LQCD and LCSR are both valid. There are two possible reasons for the LQCD discrepancy at low q^2; the form factor predictions provided by Bouchard et. al. perform a simultaneous lattice, quark mass and kinematic extrapolation while two extrapolations are performed in the prediction provided by Flynn et.al. Another possibility for the discrepancy is the assessment of the perturbative matching error, matching the lattice results to their continuum counterparts. The matching is carried out assuming zero kaon momentum whereas it varies with kaon momentum, although this effect is likely very small [33].

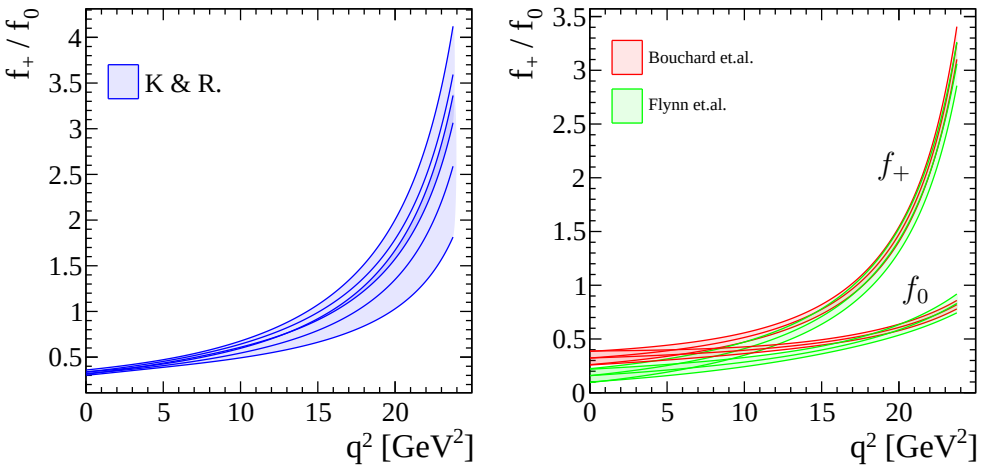


Figure 2.3 *The form factor predictions f_+ and f_0 for $B_s^0 \to K^-\mu^+\nu_\mu$ calculated using QCD sum rules (left) and lattice QCD (right) from references [30–32]*

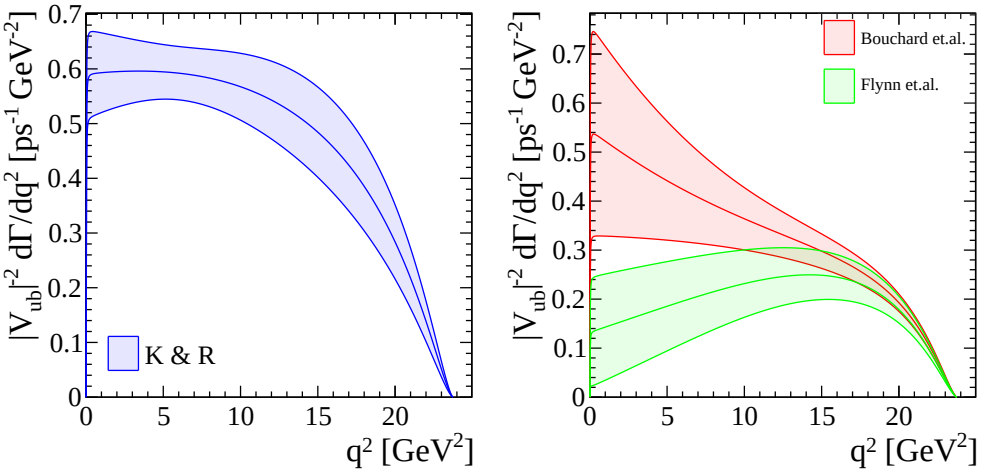


Figure 2.4 *The predicted differential decay rates for $B_s^0 \to K^-\mu^+\nu_\mu$ calculated using QCD sum rules (left) and lattice QCD (right) from references [30–32].*

2.3.3 $B_s^0 \to D_s^- \mu^+ \nu_\mu$ Form Factors

The $B_s^0 \to D_s^- \mu^+ \nu_\mu$ differential decay rate is given by

$$\frac{d\Gamma(B_s^0 \to D_s^- \mu^+ \nu_\mu)}{d\omega} = \frac{G_F^2 |V_{cb}|^2}{48\pi^3} m_{D_s^+}^3 (m_{B_s^0} + m_{D_s^+})^2 (\omega^2 - 1)^{3/2} |\mathcal{G}(\omega)|^2, \quad (2.82)$$

where $\mathcal{G}(\omega)$ is conventionally introduced as

$$\mathcal{G}(\omega) = \frac{2\sqrt{r}}{1+r} f_+(\omega), \quad (2.83)$$

with

$$\omega(q^2) = 1 + \frac{{q^2}_{\max} - q^2}{2 m_{B_s^0} m_{D_s^+}} \quad \text{and} \quad r = \frac{m_{D_s^+}}{m_{B_s^0}}. \quad (2.84)$$

The form factor f_+ is parametrised using a modification of the BCL parametrisation [34] with J=3

$$f_+(q^2) = \frac{1}{P_+} \sum_{j=0}^{J-1} a_j^{(+)} \left[z^j - (-1)^{j-J} \frac{j}{J} z^j \right], \quad (2.85)$$

where P_+ is called the Blaschke factor. $P_+(q^2)$ and z are given by

$$P_+(q^2) = \left(1 - \frac{q^2}{m_+^2}\right) \quad \text{and} \quad z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \quad (2.86)$$

where $m_+ = m_{B_c^*} = 6.3309$ GeV, $t_+ = (m_{B_s^0} + m_{D_s^+})^2$ and $t_0 = (m_{B_s^0} - m_{D_s^+})^2$. The coefficients, $a_{+,0}^{(k)}$, for both models discussed in this section are given in Appendix A.

Two sets of lattice QCD form factor calculations for $B_s^0 \to D_s^- \mu^+ \nu_\mu$ are used in the determination of $|V_{ub}|/|V_{cb}|$ from Bailey et al. [35] and Monahan et al. [36]. The calculated form factors and differential decay rates are plotted in Figure 2.5.

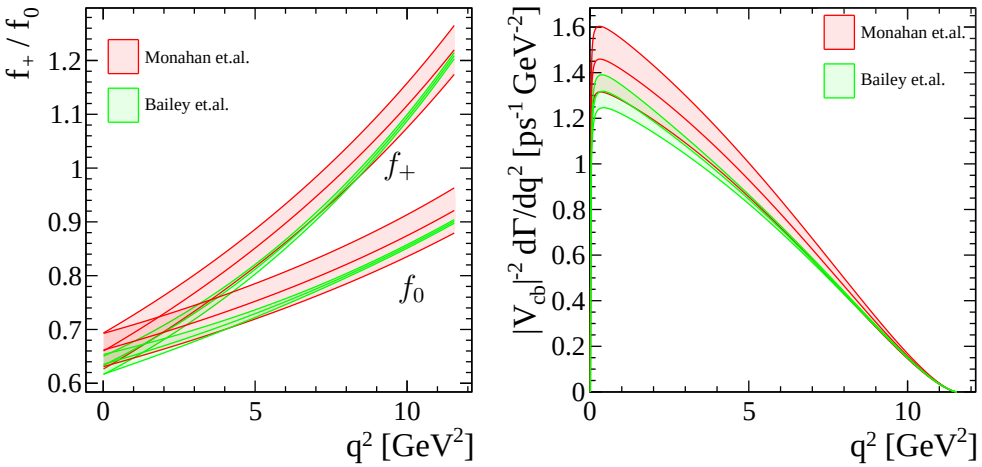


Figure 2.5 *The form factor predictions (left) and differential decay rates (right) for $B_s^0 \to D_s^- \mu^+ \nu_\mu$ calculated using lattice QCD from references [35, 36].*

	$\Gamma\|V_{ub}\|^{-2}[\mathrm{ps}^{-1}]$	$\Gamma\|V_{ub}\|^{-2}[\mathrm{ps}^{-1}]$ $q^2 > 7\ \mathrm{GeV}^2/c^4$	$\Gamma\|V_{ub}\|^{-2}[\mathrm{ps}^{-1}]$ $q^2 < 7\ \mathrm{GeV}^2/c^4$	$\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)$ $\times 10^{-4}$
Flynn et al.	4.54 ± 1.35	3.37 ± 0.70	1.18 ± 0.67	0.93 ± 0.27
Bouchard et al.	7.75 ± 1.57	4.47 ± 0.61	3.29 ± 0.99	1.59 ± 0.32
K & R	11.07 ± 1.13	6.94 ± 1.02	4.14 ± 0.40	2.29 ± 0.23

Table 2.2 *The predicted decay widths and branching fractions of $B_s^0 \to K^- \mu^+ \nu_\mu$ are presented for the form factor predictions given in References [30–32] for the full q^2 region and the high and low bins. The exclusive average of $|V_{ub}|$ and $|V_{cb}|$ as determined by the PDG are used in the calculation of branching fractions [37]*

	$\Gamma\|V_{cb}\|^{-2}[\mathrm{ps}^{-1}]$	$\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)$
Bailey et al.	8.17 ± 0.24	0.0215 ± 0.0006
Monahan et al.	8.98 ± 0.73	0.0238 ± 0.0020

Table 2.3 *The predicted decay widths and branching fractions of $B_s^0 \to D_s^- \mu^+ \nu_\mu$ are presented for the form factor predictions given in References [35, 36].*

2.3.4 Form factor Results

The form factor results for $B_s^0 \to K^- \mu^+ \nu_\mu$ are

$$\Gamma|V_{ub}|^{-2}\Big|_{q^2<7\ \ \mathrm{GeV}^2} = 4.14 \pm 0.40\ \mathrm{ps}^{-1}, \qquad \Gamma|V_{ub}|^{-2}\Big|_{q^2>7\ \ \mathrm{GeV}^2} = 3.92 \pm 0.88\ \mathrm{ps}^{-1}, \tag{2.87}$$

where the value for high q^2 is the weighted average of two LQCD results under the assumption that the uncertainties between the two calculations are linearly correlated. A visualisation of the averaging procedure,which by construction includes the extrapolation uncertainty to low q^2, is given in Figure 2.6.

The weighted average of the form factor results for $B_s^0 \to D_s^- \mu^+ \nu_\mu$ is

$$\Gamma|V_{cb}|^{-2} = 8.57 \pm 0.69\ \mathrm{ps}^{-1}, \tag{2.88}$$

where the uncertainties between the two models are assumed to be completely correlated. A visualisation of the averaging procedure is plotted in Figure 2.6.

The full set of results from the form factor calculations including decay widths in different regions of phase space and predicted branching fractions using global averages of the exclusive values of $|V_{ub}|$ and $|V_{cb}|$ are given in Tables 2.2 and 2.3 for $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ respectively.

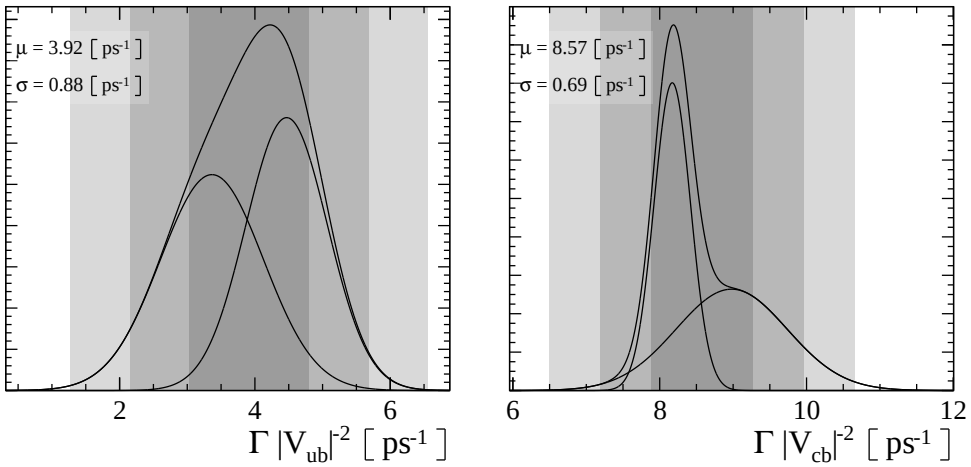


Figure 2.6 *Plots demonstrating the averaging of the form factor predictions for $B_s^0 \to K^- \mu^+ \nu_\mu$ (left) and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ (right) with* 1, 2 and, 3σ *error bars shaded in grey.*

Chapter 3

The LHCb experiment

The measurement presented in this thesis was performed using data collected by the LHCb experiment during the year 2012. The Large Hadron Collider (LHC) produced proton-proton collisions which were detected by LHCb.

This chapter provides an overview of the relevant aspects of the LHC and LHCb machines. The Large Hadron Collider and LHCb experiment are introduced in Sections 3.1 and 3.2 respectively. The reconstruction of semileptonic B_s^0 decays using the LHCb experiment is discussed in Section 3.3. Tracking and calorimetry are presented in Sections 3.4 and 3.5 respectively. Finally the trigger and simulation are discussed in Sections 3.6 and 3.7 respectively.

3.1 The Large Hadron Collider

The LHC is the world's largest and most powerful particle accelerator and collider with a circumference of 27 km. The LHC straddles the French-Swiss border near Geneva at the European Organization for Nuclear Research (CERN). The LHC accelerated protons to centre of mass energies of $\sqrt{s} = 7$ TeV, $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV during the years of 2011, 2012 and 2015-2018 respectively.

Protons for the LHC are sourced from a bottle of hydrogen. The hydrogen atoms are ionised, and the protons accelerated through a series of linear and circular accelerators prior to injection into the LHC at an energy of 450 GeV. Figure 3.1 shows the accelerator chain used to accelerate and inject protons and ions for the LHC. Protons inside the LHC are grouped into bunches with a maximum

design capacity of 2808 bunches per beam, and each bunch containing 1.2×10^{11} protons. Eight radiofrequency, RF, cavities per beam accelerate protons to the desired energies. Dipole magnets bend the beam around the ring while quadrupole, sextupole and octopole magnets focus the beam [38]. Bunches are spaced 25/50 ns apart and are focused at the interaction points by the LHC producing collisions at a rate of 40/20 MHz[1].

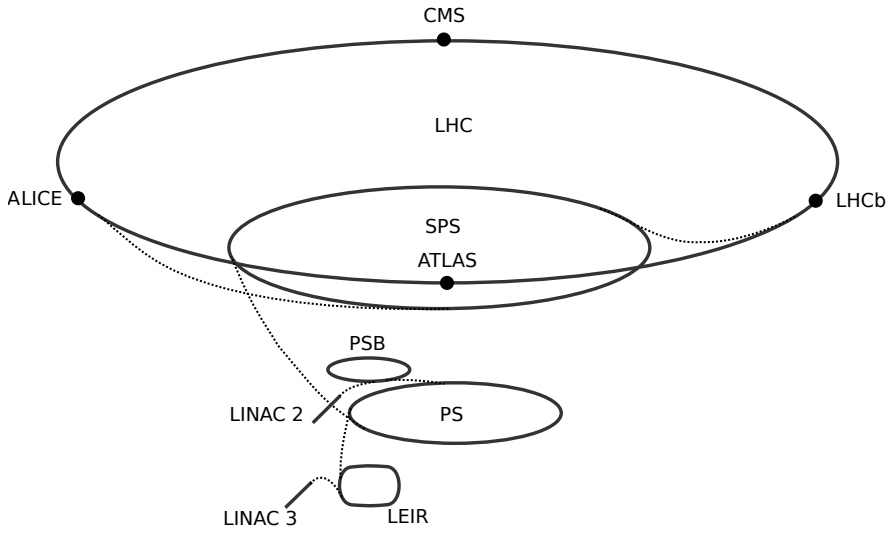


Figure 3.1 *The accelerator chain used to provdie protons and ions for the LHC is shown. Protons originate at LINAC 2 and ions originate at LINAC 3. Image taken from [39].*

3.2 The LHCb experiment

The LHCb experiment is dedicated to the study and precise measurement of b and c-physics. The experiment exploits the high production cross sections for $b\bar{b}$ and $c\bar{c}$ pairs, $\sigma(pp \to b\bar{b}X) = 72.0 \pm 0.3 \pm 6.8$ μb for $b\bar{b}$ within the acceptance of the LHCb experiment at $\sqrt{s} = 7$ TeV [40], with 10^{12} $b\bar{b}$ pairs produced during 2012. The $b\bar{b}$ production cross section at LHCb is five orders of magnitude larger than at Belle and BaBar, $\sigma(e^+e^- \to \Upsilon(4S) \to b\bar{b}) = 1.2, 1.1$ nb respectively, [41, 42] providing an ideal environment for high statistics measurements of standard model parameters.

The collisions at BaBar and Belle produce a very clean environment due to the nature of the annihilation type collision and the collision centre of mass energy

[1]The effective collision frequency during 2012 was closer to 11 MHz due to gaps in the beam and bunch crossings with no visible collisions.

being tuned to the $\Upsilon(4S)$ mass. A hadronic environment, as present at the LHC, produces an event with considerably more activity and in order to keep the detector occupancy at a manageable level beam optics limit the number of collisions per bunch crossing to approximately 1.5, equivalent to a modest instantaneous luminosity of $\mathcal{L} \approx 2 \times 10^{32}\text{cm}^{-2}\ \text{s}^{-1}$.

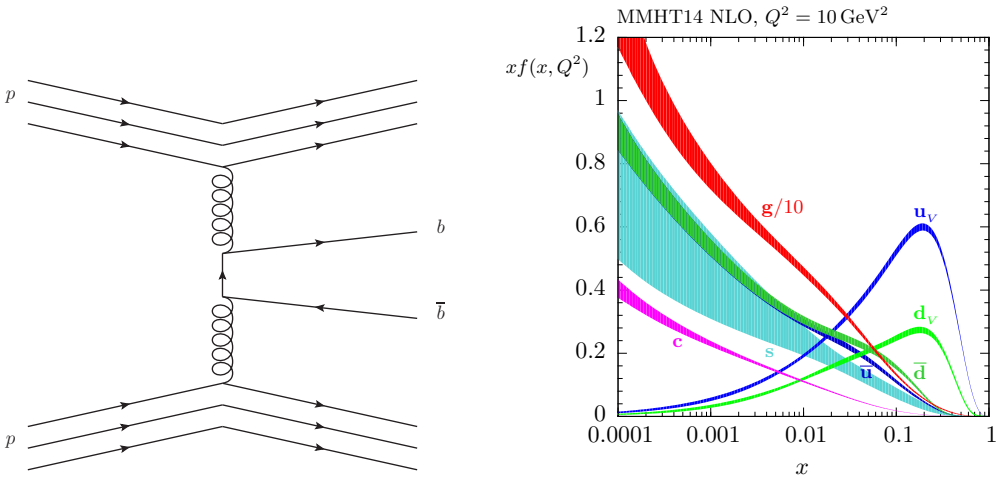


Figure 3.2 *A Feynman digram depicting $b\bar{b}$ production via gluon-gluon fusion (left) and NLO PDFs at $q^2 = 10\ \text{GeV}^2/c^4$ (right). The fractional momentum of the proton is dominated by the gluon. Image taken from [43].*

The dominant $b\bar{b}$ production method at LHCb is via gluon-gluon fusion where the incoming protons radiate gluons which fuse to produce a $b\bar{b}$ pair. Gluons carry a large fraction of the proton's momentum, the fractions of which are plotted in Figure 3.2. A difference in the momentum of the gluons is propagated to the $b\bar{b}$ pair boosting the interaction with respect to the centre of mass frame of the pp collision. Consequently the $b\bar{b}$ and $c\bar{c}$ pairs are produced in the forward and backward regions of the detector.

The LHCb detector [44, 45] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks, where

$$\eta = -\ln\left|\tan\frac{\theta}{2}\right|, \tag{3.1}$$

and θ is the angle between the particle's momentum and the beam line. Despite covering just 1.8% of the solid angle, 25% of $b\bar{b}$ pairs are produced within the detector acceptance. The full acceptance is $10 < \theta < 250[300]$ mrad in the horizontal bending [vertical non-bending] plane.

A schematic of the LHCb detector is shown in Figure 3.3. A right handed

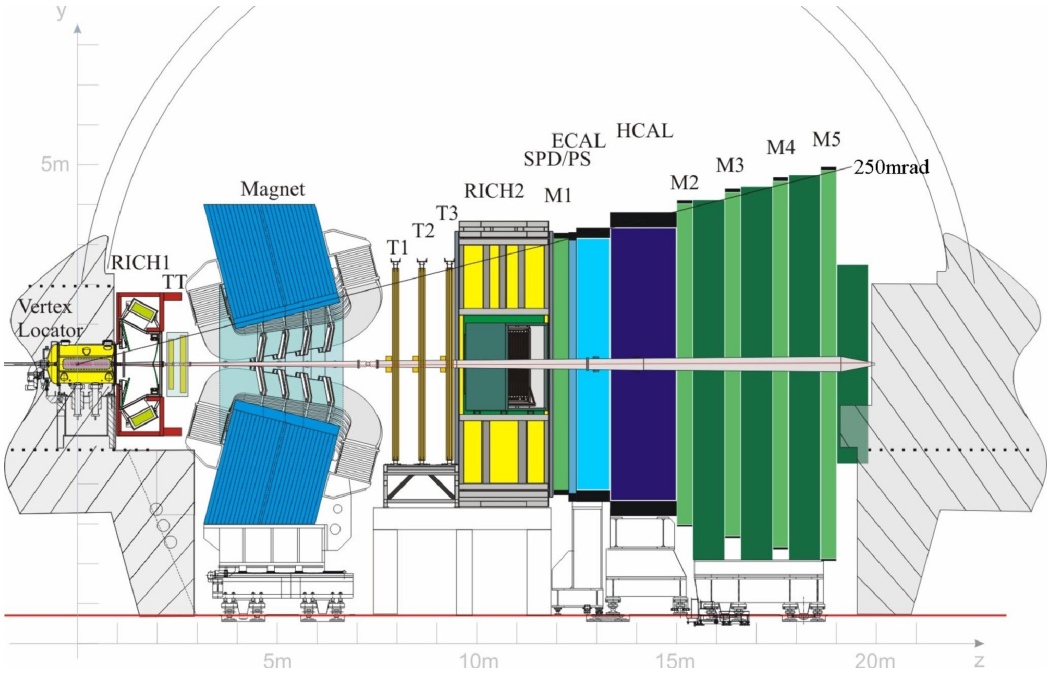


Figure 3.3 *Schematic of the LHCb experiment*

coordinate system is used with the origin centred at the interaction point; the positive z axis points 'downstream' towards the end of the detector, the positive x axis points towards the centre of the LHC and the positive y axis points straight up. The downstream direction is defined to point from the collision point towards the muon stations.

The principal components of the detector can be divided into a few key categories, tracking, calorimetry and particle identification. A spectrometer dipole magnet bends charged particles in the horizontal plane allowing the charge and momentum of the particles to be determined from the direction and radius of the curvature. The vertex locator, VELO, envelopes the collision point and accurately measures the location of tracks produced close to the interaction point. Additional downstream tracking is provided by the tracker turicensis, TT, before the magnet and three tracking stations after the magnet consisting of inner and outer trackers, IT and OT respectively.

Charged hadronic particle identification is made possible by two ring imaging Cherenkov detectors, RICH1 before the magnet, and RICH2 after the magnet. the two RICH detectors contain gasses with different refractive indexes giving optimal performance at different momentum ranges.

Calorimetry and additional particle identification are provided by the electromag-

netic and hadronic calorimeters, ECAL and HCAL, a preshower and scintillating pad detector, PS and SPD, and five muon stations, M1-5.

3.3 Reconstructing Semileptonic Decays

Consider the decay $B_s^0 \to K^-\mu^+\nu_\mu$, two protons collide at the interaction point, close to the centre of the VELO producing a B_s^0 meson and many prompt tracks. Prompt tracks are defined as originating from the point of collision. On average the B_s^0 meson flies a distance of 14 mm before decaying into a charged kaon, muon and invisible neutrino. The charged particles leave hits in the VELO as they traverse the sub-detector, and a reconstruction algorithm reconstructs the trajectories of the charged tracks from the hits. The point of the collision and production of the B_s^0 is precisely determined by performing a vertex fit on the tracks and is referred to as the primary vertex, PV, and the decay location of the B_s^0, known as the secondary vertex, SV, is determined by performing a vertex fit on two oppositely charged particles with high kaon and muon likelihoods. As the PV is calculated using a larger number of tracks than the SV, the resolution on the PV is significantly better than the resolution on the SV.

Moving along the beam axis in a downstream direction, the particles traverse the first ring imaging Cherenkov, RICH, detector and emit Cherenkov radiation. The light radiated by the particles is focused onto and recorded by hybrid photon detectors, HPDs. A likelihood hypothesis for the particle types is calculated from the pattern of the radiation, thus allowing the kaon to be positively identified.

The particles then cross the first tracking station, the Tracker Turicensis, TT, followed by the dipole magnet and are bent in opposite directions. The particles then cross the inner tracker, IT, and outer tracker, OT. Hits left in the tracking station before and the three tracking stations after the magnet allow the curvature of the charged tracks, and hence their momenta to be measured.

If the particles are close to the beam pipe they will pass through the second RICH detector allowing a second measurement of the angle of the emitted photons. This will strengthen the particle identification likelihood hypothesis. Behind RICH2 the $K^-\mu^+$ enter the calorimeter. Hits are left in both the SPD and PS indicating that the particles are charged and vetoing any possibility that the particles are photons or neutral pions. Both particles leave hits in the ECAL. The muon

traverses the HCAL leaving a small signal while the kaon showers and is fully absorbed, positively identifying the kaon as a hadron. Finally the muon enters the muon station where it's location is accurately measured and is positively identified as muon. The large signals in the muon system are detected and passed on to the hardware trigger, L0, which flags the event as interesting. Two software triggers, HLT1 confirms the presence of a muon, and HLT2 performs a full event reconstruction allowing the event to be saved permanently for offline analysis.

3.4 Tracking

3.4.1 Magnet

A dipole magnet with an integrated field strength of $\int B\ dl = 4$ Tm bends the paths of charged tracks allowing their charge and momenta to be determined with a resolution of $\delta p/p = 4 \times 10^{-3}$. The bending force, $\vec{F}$, for a particle with charge, q, moving with velocity, v, in a magnetic field, B is:

$$\vec{F} = q\left(\vec{v} \times \vec{B}\right). \tag{3.2}$$

In order to effectively determine the momentum of charged particles, the magnetic field must be as high as possible, however the vertex locator and Hybrid Photodetectors (HPD) of the RICH1 detector are sensitive to magnetic fields, and the field strength must be minimised outside the region of the magnet. Figure 3.4 plots the field strength of the magnet against the z axis of the detector with the location of the trackers overlaid.

The magnet is composed of two saddle shaped coils in a window frame yoke with sloping poles. Each magnet coil consists of fifteen pancakes arranged in five triplets and are made from pure Al-99.7 with a central channel for water cooling. The nominal current passing through the coils during operation is 5.85 kA.

3.4.2 VELO

The LHCb vertex locator, VELO, is a silicon strip detector operating in a secondary vacuum around the primary LHC vacuum. The VELO measures the

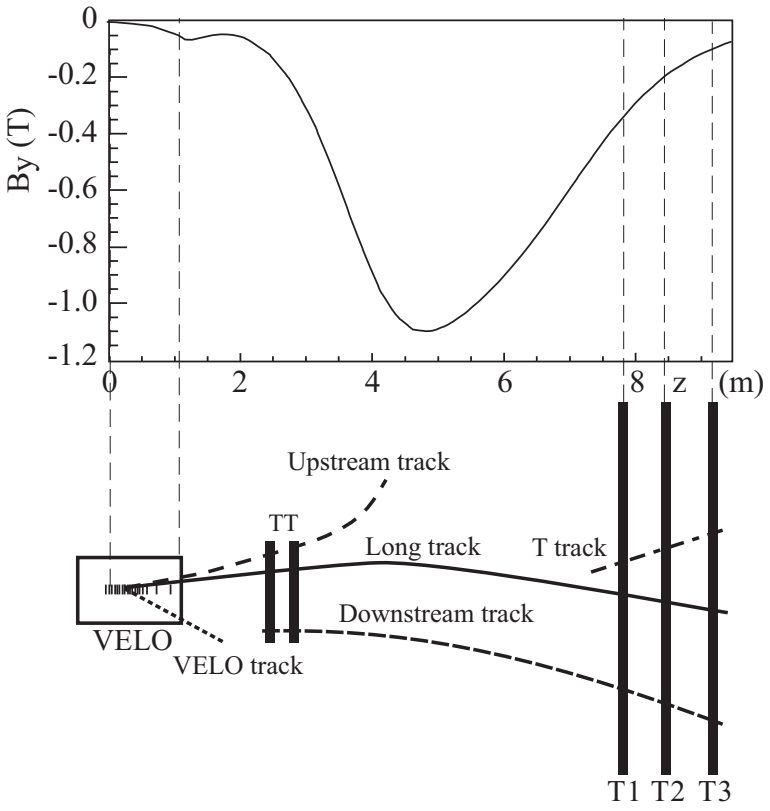


Figure 3.4 *A plot of the magnetic field profile is shown with a digram of the tracking systems and a characterisation of the different tracks types. Image taken from [45].*

location of the collision and decay vertex of beauty and charm hadrons from the hits left behind as charged particles traverse the detector. This information is used to accurately measure the decay times of hadrons and the impact parameters of long lived particles. Detached vertices make up a vital component of the trigger and are used to enrich the b-hadron content of events saved to trigger. The beam width during injection is larger than the inner radius of the VELO requiring that the VELO be retracted during injection.

The VELO is required to meet several performance, geometric, environmental and machine criteria. In order to accurately measure the location of the production and decay vertices, the signal to noise ratio of the VELO should be grater than 14 [46], corresponding to ≈ 200 noise hits per event. A spacial cluster resolution of 4 μm is required for tracks with an angle of 100 mrad from the beam line. The final consideration is the spillover probability, the fraction of the signal remaining after 25 ns which is required to be less than 0.3. From a geometric point of view, the VELO must cover the angular acceptance of the downstream detectors, $1.6 < \eta < 4.9$, and tracks emerging from primary vertices, $|z| < 10.6$ cm, must traverse three VELO stations. The minimum distance between the innermost VELO sensors and the beam is 8 mm while the outer radius is greater than 4.2 cm and modules are spaced by 3.5 cm. In order to cover the full azimuthal angle the two detector halves overlap slightly. This range is achieved by offsetting one half of the detector by 1.5 cm in z. The VELO is operated in an extreme radiation environment with the dose from one year of operation equivalent to a 1 MeV neutron flux of $1.3 \times 10^{14} n_{eq}/cm^2$ in the innermost region. The VELO must be capable of operating in these conditions for the duration of data taking. Given that the VELO must be positioned as close as possible to the beam and that the VELO functions optimally in a vacuum, the integration of the detector with the LHC introduces several design constraints. To protect the LHC from VELO out gassing, and the VELO from wakefield currents and beam halo effects, the VELO must be shielded from the LHC by a metallic foil, the RF foil. The section of the VELO closest to the beam is exposed to beam induced bombardment and is protected from beam induced effects such as synchrotron radiation and secondary electrons. Variations in the closed-orbit of the LHC and the thickness of the RF foil limit the minimum distance from the beam to 8 mm. During injection of the LHC the beam width is considerably larger than the 8 mm inner radius requiring that the VELO be retracted 29 mm into the shadow of the LHC beam.

The VELO is made up of a series of modules placed around the interaction

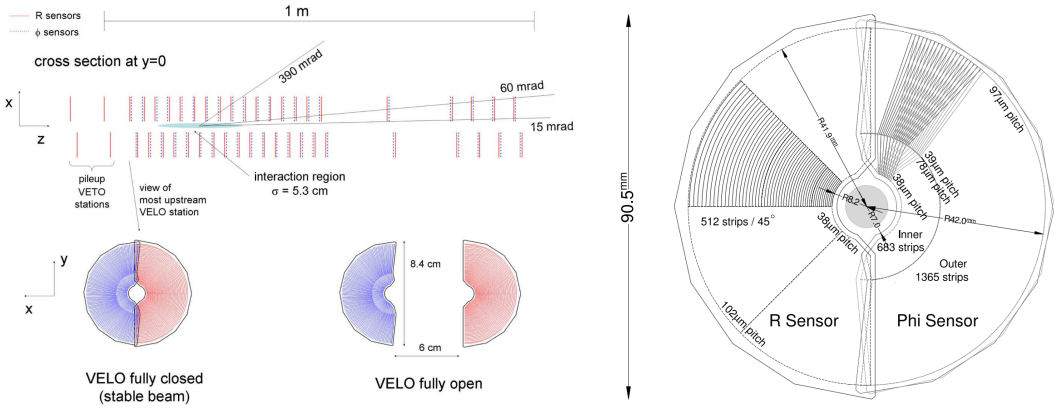


Figure 3.5 *The cross section in the x, z plane of the VELO sensors in the closed position, with the front face of the first modules shown open and closed (left). A sketch illustrating the $r - \phi$ geometry of the sensors is shown (right), showing only a few of the strips. The strips of two adjacent ϕ sensors are overlaid to demonstrate the stereo angle. Images taken from [44].*

point. The modules perform three functions, they hold the sensors in position, connect the electrical readouts to the to the sensors and enable cooling while in a vacuum. Each module holds two sensors, an R-sensor and a ϕ-sensor. The R-sensor measures the radius of a charged track while the ϕ-sensor measures the azimuthal angle. The silicon sensors use diode strip implants with a minimum strip separation (pitch) of 32 µm. The layout of the strips is illustrated in Figure 3.5. For the R-sensors the strips form semi-circles divided into 45° regions (in order to reduce occupancy and capacitance), centred on the LHC beam. The strip pitch increases from 38 µm at the point closest to the beam up to 101.6 µm at the point furthest from the beam. The ϕ-sensors are divided into an inner and outer region, the strips in the inner region are skewed by 20° to the radial and the outer region begins at a radius of 17.25 mm and are skewed by 10°. There are approximately twice as many strips in the outer region than in the inner region. Adjacent modules have the skew reversed. The material budget of the VELO corresponds to 17.5% of a radiation length with the RF-foil introducing the bulk of the material.

The performance of the VELO can be quantified by considering the resolution on the measured vertices and the impact parameter of tracks. The impact parameter, IP, is defined as the shortest distance between a point and a line, where in this example the point is the PV and the line is the particle trajectory. A similar variable exists defining the distance of closest approach between two lines, DOCA, In this example the DOCA is the closest point between the particle trajectory and

the beam line. The resolution on the primary vertex is measured experimentally by randomly dividing the tracks from a vertex into two subsets and reconstructing the PV location for each subset. The resolution is found by subtracting one measurement from the other [47]. A minimum of five tracks are required to reconstruct a PV, and this method is capable of measuring the resolution on primary vertices containing up to 65 tracks. The PV uncertainty is strongly correlated to the number of tracks originating from the PV. The PV resolution, in z, is plotted against track multiplicity in Figure 3.6.

Tracks originating from the decays of long lived B or D mesons typically have larger impact parameters than tracks originating from the primary vertex. The impact parameter is used extensively in LHCb analyses to reduce pollution from prompt backgrounds making an understanding of its resolution essential. The resolution on the impact parameter is governed by three main factors: multiple scattering of particles due to the detector material, typically the RF-foil, the resolution on the hit location in the VELO and the distance between the PV and the first measured hit. The VELO was designed to minimise these factors.

The resolution of the impact parameter is typically displayed for a component of IP vector in the plane transverse to its flight direction, IP_x and IP_y, where,

$$IP_x = x - x_{PV} - (z - z_{PV})t_x, \tag{3.3}$$

and similarly for y, where (x, y, z) is the position of the track at the closest point to the primary vertex, and $(t_x, t_y, 1)$ is the direction vector of the track. The component of the IP parallel to the flight direction, IP_z is defind to be 0. The resolution on IP_x is plotted against $1/p_{\mathrm{T}}$ in Figure 3.6. The linear dependence with $1/p_{\mathrm{T}}$ is a consequence of multiple scattering.

3.4.3 Silicon and Straw Trackers

Tracking information is provided by the silicon tracker turicensis, TT, located downstream of RICH1 and upstream of the magnet and the three additional tracking stations, T1-T3, immediately downstream of the magnet. The three tracking stations consist of a silicon inner tracker (IT) with small acceptance close to the beam pipe and a straw tube outer tracker (OT).

A primary goal of the TT is to reconstruct tracks which originate outside the

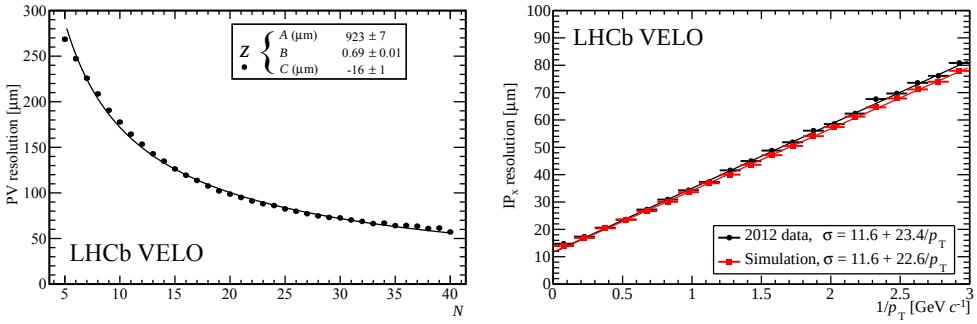


Figure 3.6 *The uncertainty on the primary vertex in the z direction is plotted against the number of tracks originating from the PV (left) and the uncertainty on the impact parameter in the x direction is plotted against* $1/p_{\mathrm{T}}$ *(right) using data collected in 2012. Images taken from [48].*

VELO, such as those originating from the decays of K^0_{S} and Λ. While the B^0_s only flies 14 mm and this decays inside the VELO, the TT still provides vital tracking information. The TT is made of four layers of silicon strip sensors with a pitch of 183 μm [49]. The outer two layers are aligned vertically and the two inner sensors are rotated by ±5° from the vertical. Each layer of the TT consists of 14 columns of silicon sensors, with adjacent modules staggered in z and gaps in acceptance are avoided by overlapping sensors by a few mm.

The IT is very similar to the TT, it consists of four individual detector boxes arranged around the beam pipe in the highest occupancy part of the detector as shown in Figure 3.7 [50].

The Outer Tracker is a drift time detector [51] providing excellent momentum resolution and a high reconstruction efficiency over a large acceptance. The OT holds an array of gas-tight straw tube modules, each containing two staggered layers of drift tubes with an internal diameter of 4.9 mm. A drift time below 50 ns is achieved by using a 70/30 mix of Argon and CO_2 gasses giving a drift resolution of 200 μm. Each station consists of four layers with the tubes in the outer tubes arranged vertically, and the inner tubes rotated by ±5°. The tracker is made of narrow columns to provide the greatest resolution in the y, bending, direction allowing the curvature of the tracks to be precisely measured.

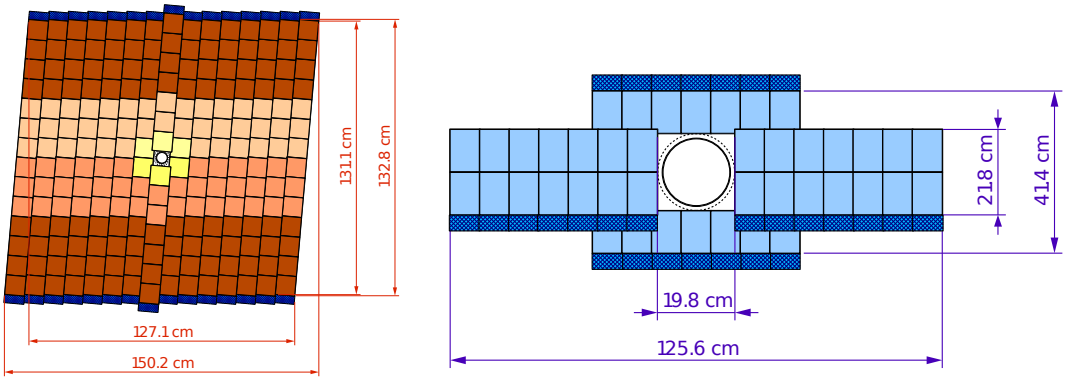


Figure 3.7 *The layout of the third Tracker Turicensis layer (left) with different readout sections indicated by different shadings, and the layout of an x detection layer of the second Inner Tracker station. Images taken from [45]*

	Tracking	SPD	PS	ECAL	HCAL	MUON
γ			Hit	Shower		
π^0				Shower		
$\pi^0 \rightarrow \gamma\gamma$			Hit	Shower		
n					Shower	
$e^{\pm}$	Hit	Hit	Hit	Shower		
$\mu^{\pm}$	Hit	Hit	Hit	Hit	Hit	Hit
$K^{\pm}$,$\pi^{\pm}$,p	Hit	Hit	Hit	Hit	Shower	

Table 3.1 *The signature in the detector left by different particle types are listed. Additional information provided by the two RICH detectors are used to separate the flavours of charged hadron.*

3.5 Particle Identification and Calorimetry

3.5.1 RICH

Two ring-imaging Cherenkov detectors (RICH1 and RICH2) provide particle identification for charged hadrons. The angle of emittance of Cherenkov radiation, θ_c, is related to the particle's mass, m, momentum, p, and the refractive index, n, of the material traversed by

$$\cos\theta_c = \frac{\sqrt{m^2 + p^2}}{np}. \tag{3.4}$$

Curved mirrors project and focus the radiated Cherenkov light onto a matrix of hybrid photon detectors, HPD. The Cherenkov radiation forms a tight circle on the HPDs, and the angle of emittance is proportional to the radius of the circle. Figure 3.8 visualises the arrangement of the two radiator materials, mirror shape and HPD locations for RICH1.

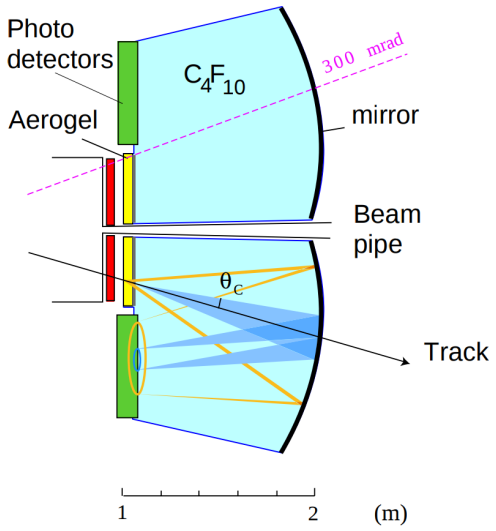


Figure 3.8 *A schematic layout of RICH1 illustrating the focussing of Cherenkov light originating from the aerogel, yellow, and C_4F_{10} gas, blue. Image taken from [52].*

In Figure 3.9 distributions of θ_c are plotted against the particle momentum for particles traversing RICH1 with radiator C_4F_{10}. Clear bands are visible corresponding to muons, pions, kaons and protons. The starting position of the bands indicates the minimum momentum required in order for Cherenkov photons to be produced. In both RICH detectors the produced Cherenkov light is focused

onto Hybrid Photon Detectors, HPDs, using a combination of spherical and flat mirrors. The HPDs are located outside the acceptance of the LHCb detector and are capable of detecting Cherenkov photons in the wavelength range 200-600 nm. The HPDs are sensitive to magnetic fields and are shielded from the magnetic fields present in the detector by mu-metal cylinders which limit the magnetic field exposed to the HPDs to 50mT.

RICH1 has an angular acceptance of 25 mrad $< \theta <$ 300 mrad covering the full acceptance of the detector and is located upstream of the magnet and uses C_4F_{10} gas as the radiator with refractive index $n = 1.0014$ giving effective separation power up to 40 GeV/c. Particle identification at momenta below these thresholds is enabled by a 50 mm layer of silica aerogel at the entrance of RICH1 with refractive index, $n = 1.03$. The aerogel was removed from RICH1 during the first long shutdown as the degradation in particle quality was worse than the additional identification performance. RICH2 uses CF_4 gas as the radiator with refractive index $n = 1.0005$ and provides effective identification in the momentum range 15 GeV/$c < p <$ 100 GeV/c. High momentum particles are typically produced with a smaller production angle and as such the RICH2 detector only has an acceptance covering the range 15 mrad $< \theta <$ 100 mrad.

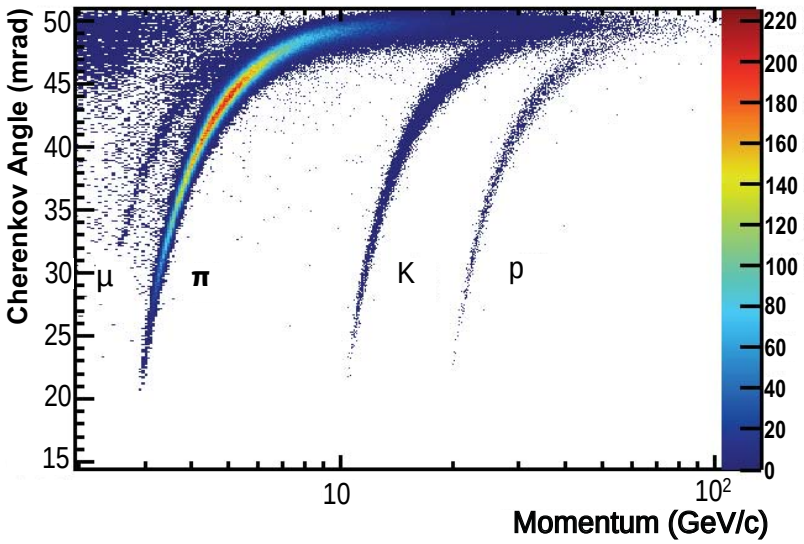


Figure 3.9 *The reconstructed Cherenkov angle, θ_c, for isolated tracks is plotted against the particle momentum, p, for radiator* C_4F_{10}*. Image taken from [53].*

3.5.2 Calorimetry and the Muon system

The calorimeters perform several crucial functions, they send low level information to the hardware trigger allowing the selection of high transverse energy, E_{T}, hadron, electron and photon candidates, provide particle identification and measures the energies and positions of electrons, photons and hadrons. The muon system provides fast information to the triggers, particle identification and space-point information.

LHCb uses a *classical* calorimetry design with the electromagnetic calorimeter, ECAL, placed in front of the hadronic calorimeter, HCAL. A double detector is placed in front of the ECAL consisting of scintillating pad detector, SPD, a thin sheet of lead and a second pad detector called the preshower detector, PS. The SPD and PS determine the electromagnetic nature of particles and whether they're charged from the calorimeter clusters allowing the vast backgrounds from charged and neutral pions to be rejected [54]. Neutral pions are identified as resolved if they decayed into two photons before the calorimeter or merged if they decay inside the calorimeter [55].

All the calorimeters follow the same principal. Wavelength shifting fibres transfer scintillation light to a photomultiplier, PMT. The fibres from the SPD/PS cells are read out using multianode photomultiplier tubes, MAPMT, and the fibres from the HCAL and ECAL are read out by individual phototubes. The ECAL has an energy resolution of $\sigma_E/E = 10\%/\sqrt{E[\ \mathrm{GeV}]} \bigoplus 1\%$ and the HCAL has a resolution of $\sigma_E/E = (69 \pm 5)\%/\sqrt{E[\ \mathrm{GeV}]} \bigoplus (9 \pm 2)\%$ [45].

The muon system is composed of five stations, M1-M5, of rectangular shape with inner and outer acceptances of 20 (16) mrad and 306 (258) in the bending (non-bending) plane respectively. A projective geometry is used for the muon stations, the dimensions scale with distance from the collision point. Muon station M1 is placed before the calorimeters and is used to improve the p_{T} measurements given to the trigger. Muon stations M2-M5 are placed downstream of the calorimeters and are separated by iron absorbers with a thickness of 80 cm and each station uses Multi-Wire Proportional Chambers (MWPCs) for detecting muons.

The energy required for a muon to traverse the entire detector is approximately 6 GeV and the total absorber thickness is approximately 20 interaction lengths. The muon stations M1-M3 have a high spatial resolution and are used to define the track direction and calculate the p_{T} of the candidate muon. The resolution

Momentum	Muon Stations
3 GeV/$c < p <$ 6 GeV/c	M2 & M3
6 GeV/$c < p <$ 10 GeV/c	M2 & M3 & (M$\|\|M5$)
$p >$ 10 GeV/c	M2 & M3 & M4 &M5

Table 3.2 *The requirements for a track to be classified as a muon are presented. The full requirements are dependant on the momentum of the candidate and which stations detect hits from the track.*

on the measured p_{T} of a muon is approximately 20% using information from the muon system only. Stations M4 and M5 have a coarser spatial resolution with their main purpose being the identification of penetrating particles.

A binary, yes/no, decision known as $isMuon$ is made based on the track momenta and which stations a track leaves hits, see Table 3.2. The muon identification method provides an excellent selection efficiency with $98.13 \pm 0.04\%$ of muons being correctly identified and less than 1% of hadrons being misidentified as muons [56].

3.5.3 Particle Likelihood

A typical event can contain several hundred particles which traverse the two RICH detectors producing many overlapping rings in the detector making the reconstruction of Cherenkov rings a challenge. A likelihood hypothesis is created for each particle ($\pi^{\pm}$, $K^{\pm}$, $\mu^{\pm}$, p) by assuming the mass of the particle and combining information from the two RICH detectors, the calorimeters and the muon system. The unique signals left by different particle types in the sub detectors are summarised in Table 3.1. When selecting the desired particle type, the logarithm of particle hypotheses are compared, i.e. when selecting kaons, the likelihood would be compared to the pion or proton, $\Delta \log \mathcal{L}(K - \pi/p)$. The kaon identification and misidentification rates are plotted in Figure 3.10 for the selections, $\Delta \log \mathcal{L}(K - \pi) > 0,5$. Simulated Monte Carlo samples fail to accurately model the rates for a given PID selection, so a data driven approach is used to calculate the rates for different $\Delta \log \mathcal{L}$ selections using clean calibration samples of pions and kaons from $D^* \to \pi^+(D^0 \to K^-\pi^+)$ decays and muons from $J/\psi \to \mu^+\mu^-$ decays, where the particle identification, PID, selection is placed on the particle of interest.

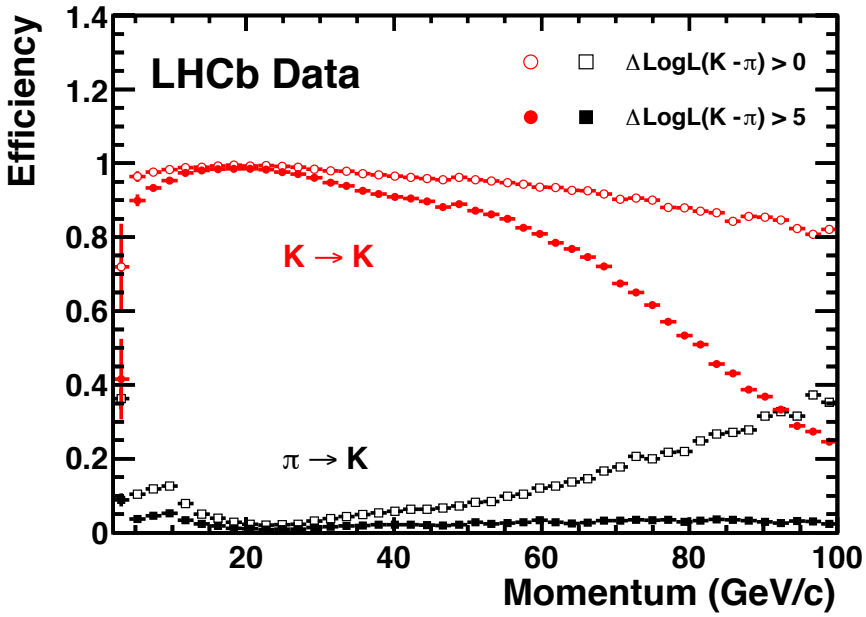


Figure 3.10 *Kaon identification efficiencies and $\pi^{\pm} \rightarrow K^{\pm}$ misidentification rates. [53]*

3.6 Trigger

During nominal running conditions in 2012 the rate of visible[2] bunch crossings was approximately 11 MHz while the maximum sustainable readout was only approximately 5 kHz. The expected $b\bar{b}$ production rate at nominal operation is approximately 100 kHz, with 15% of these events containing a B hadron with all its decay products in the LHCb acceptance. The branching fractions of *interesting* B meson decays is typically 10^{-3} [57].

A triggering system reduces the event rate by selecting events that contain potentially interesting physics and enriches the number of events saved containing b hadrons. A hardware trigger known as the Level-0 (L0) trigger provides fast, $\mathcal{O}(\mu s)$, decision making and reduces the event rate to ~ 1 MHz. Two software based triggers known as the high level triggers, HLT1 and HLT2, further reduce the rate to 40 kHz and 5 kHz respectively. The L0 trigger runs synchronously with the 40 MHz bunch crossing frequency on custom made hardware while the HLT runs asynchronously on a processor farm.

The L0 trigger takes as an input the highest E_{T} hadron, electron and photon clusters in the calorimeters and the two highest p_{T} muons in the muon chambers.

[2]A bunch crossing is defined as visible if there are at least two reconstructible charged tracks passing through the VELO.

Decision	p_{T} or E_{T} threshold	SPD hits
Single Muon	>1.76 GeV/c	<600
Dimuon $p_{\mathrm{T}1} \times p_{\mathrm{T}2}$	>(1.60 GeV/c)^2	<900
Hadron	>3.70 GeV	<600
Electron	>3.00 GeV	<600
Photon	>2.50 GeV	<600

Table 3.3 *L0 trigger thresholds used during 2012 [58].*

The calorimeters form clusters by summing the energy of 2x2 cells, and identify the clusters with highest E_{T}. Clusters are identified as γ, π^0, or hadrons using information from the calorimeters, SPD and PS, see Table 3.1. The muon chambers allow for reconstruction of the muon p_{T} with a resolution of 20% and the two highest muons in each quadrant are selected. The L0 trigger thresholds are listed in Table 3.3.

The HLT consists of a C++ application which runs on the event filter farm which contains 2000 computing nodes and makes use of the full event data to confirm the decisions made by the L0 trigger and provide further separation between signal and background. As the HLT has access to the full event information and is software defined one could implement the non-trivial offline selection algorithms, e.g. machine learning, using the trigger. The purpose of HLT1 is to reconstruct particles using information from the VELO and tracking stations and confirm the decision of the L0 trigger. The HLT2 performs a full pattern recognition and track reconstruction on the remaining events and runs a series of inclusive and exclusive trigger algorithms where the B is partially or fully reconstructed. The final trigger is the logical OR of all exclusive and inclusive triggers [59].

The total selection efficiency, ε_{tot}, is the combination of the trigger efficiency, ε_{trig}, reconstruction and selection, ε_{sel}, and efficiency for candidates to be in the detector acceptance, ε_{acc},

$$\varepsilon_{tot} = \varepsilon_{trig} \cdot \varepsilon_{sel} \cdot \varepsilon_{acc}. \tag{3.5}$$

The selection and acceptance efficiencies can be determined from simulation, and the trigger efficiency can be determined according to,

$$\varepsilon_{trig} = \frac{N_{trig|sel}}{N_{sel}}, \tag{3.6}$$

where $N_{trig|sel}$ are the number of events passing the selection and trigger and N_{sel} are the number of events passing the selection only with the absence of a trigger requirement. In Monte Carlo the number of events passing the selection in the absence of a trigger is known, however in data this cannot be known as only events which pass the trigger can be studied.

The TISTOS method [60] is used to determine the trigger efficiency from data. A candidate is labelled as TOS (Trigger On Signal) if the event was triggered using tracks from the candidate B hadron. A candidate is labelled as TIS (Trigger Independent Of Signal) if the tracks causing the event to trigger are independent of the signal candidate. The trigger efficiency can be rewritten as

$$\begin{aligned} \varepsilon_{trig|sel} &= \frac{N_{trig|sel}}{N_{TIS|sel}} \times \frac{N_{TIS|sel}}{N_{sel}} \\ &= \frac{N_{trig|sel}}{N_{TIS|sel}} \times \varepsilon_{TIS} \end{aligned} \tag{3.7}$$

where $N_{TIS|sel}$ is the number of events passing the TIS trigger and the full selection, and the efficiency of the TIS trigger, ε_{TIS}, on signal candidates which can be approximated using the number of events which pass the TOS trigger, N_{TOS}, and both TIS and TOS triggers, $N_{TIS\&TOS}$,

$$\varepsilon_{TIS} \approx \varepsilon_{TISTOS} \equiv \frac{N_{TIS\&TOS}}{N_{TOS}}. \tag{3.8}$$

This assumes that the TIS and TISTOS triggers are uncorrelated. This assumption is shown to be valid in Reference [60]

3.7 Simulation

Simulated events are used to model signal decays of b hadrons as well as various backgrounds. The simulation is divided into three packages, each of which uses additional third party libraries. The Gauss [61, 62] package simulates the pp collisions, hadronisation, decay and passage of particles through the detector. The Boole [63] package simulates the detector response and provides data in the same format as the LHCb readout electronics and the Moore [63] package provides a full simulation of the trigger.

The Gauss package uses Pythia 6.4 [64] and 8.1 [65] to simulate $pp \to b\bar{b}X$

interactions. After the $b\bar{b}$ pair have been produced they are repeatedly hadronised until the desired B hadron is created. The EvtGen [66] package simulates the decay of the B hadron and the PHOTOS [67] package models final state electromagnetic radiation. EvtGen was initially developed by the BaBar collaboration, and modified by LHCb to simulate B meson production with proton collisions. The GEANT4 [68, 69] package is used to simulate interactions between particles and the LHCb detector.

Chapter 4

The Strategy for $|V_{ub}|$ at LHCb

This chapter outlines the strategy used for the measurement of $|V_{ub}|$ at the LHCb experiment presented in this thesis. This includes the motivation for the choice to measure $|V_{ub}|$ with the decay $B_s^0 \to K^-\mu^+\nu_\mu$.

An exclusive approach is used to measure $|V_{ub}|$ at LHCb instead of an inclusive approach for several reasons. The LHCb environment contains huge amounts $b \to c$ decays which completely mask the $b \to u$ inclusive signal, in addition at the LHCb experiment it is not possible to exploit the $b \to u$ kinematic endpoint, the approach used by the B factories. The $B_s^0 \to K^-\mu^+\nu_\mu$ decay was chosen over the decay $B_s^0 \to \pi^-\mu^+\nu_\mu$ as it is easier to positively identify a kaon and there are fewer backgrounds. The decay $B_s^0 \to D_s^-\mu^+\nu_\mu$ was chosen as it most closely resembles the $B_s^0 \to K^-\mu^+\nu_\mu$ decay, but with a $b \to c$ transition. By placing a selection on the invariant mass of the final state particles decaying from the D_s^+ a data sample is produced containing very few additional backgrounds.

Over 600 billion $b\bar{b}$ pairs were produced at LHCb during 2012, with $\approx 8.2\%$ of b quarks fragmenting into B_s^0 mesons [40, 70]. The high branching fraction of $b \to u\ell\nu$ processes, $\approx 10^{-4}$, creates a high statistics environment in which $b \to u$ transitions can be measured allowing for novel determinations of $|V_{ub}|$ from the decay $B_s^0 \to K^-\mu^+\nu_\mu$.

To measure $|V_{ub}|$ solely using the decay $B_s^0 \to K^-\mu^+\nu_\mu$ a precise measurement of the $b\bar{b}$ cross section at the LHC is needed in addition to a precise measurement of the integrated luminosity. While both of these measurements have been performed they are not precise enough to perform a competitive measurement of $|V_{ub}|$. Instead a normalisation is made to $B_s^0 \to D_s^-\mu^+\nu_\mu$ decays and a

measurement of the ratio of branching fractions is performed restricting the $B_s^0 \to K^-\mu^+\nu_\mu$ decays to a region in q^2. This measurement, when combined with form factor predictions from lattice QCD and light-cone sum rules, allows for a determination of $|V_{ub}|^2/|V_{cb}|^2$. Lattice QCD form factor predictions are used when restricting to a high q^2, $q^2 > 7$ GeV$^2/c^4$ and light-cone sum rules are used when restricting to low q^2, $q^2 < 7$ GeV$^2/c^4$. An experimental measurement of the ratio of branching fractions is performed. There are three driving factors which determine the ratio of branching fractions, the ratio of CKM matrix elements $|V_{ub}|/|V_{cb}|$, the kinematics and phase space of the decays, and the calculation of form factors for the decays. The kinematic and phase space dependencies and form factors are combined into a single term which is referred to as the ratio of form factors, R_{FF}, which fully encapsulates the theoretical contributions to the ratio of branching fractions. $|V_{ub}|^2/|V_{cb}|^2$ is determined according to:

$$\frac{|V_{ub}|^2}{|V_{cb}|^2} = \frac{\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)_{|q^2>7\ \text{GeV}^2/c^4}}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} \times R_{FF}^{LQCD} \tag{4.1}$$

and

$$\frac{|V_{ub}|^2}{|V_{cb}|^2} = \frac{\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)_{|q^2<7\ \text{GeV}^2/c^4}}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} \times R_{FF}^{LCSR} \tag{4.2}$$

where the branching fractions, $\mathcal{B}$, and form factor ratio, R_{FF}, are written using equations 2.75 and 2.82,

$$R_{FF}^{LQCD} = \frac{\int_{m_\mu^2}^{(m_{B_s^0}-m_{D_s^+})^2} \frac{1}{|V_{cb}|^2}\frac{d\Gamma}{dq^2}_{B_s^0\to D_s^-\mu^+\nu_\mu} dq^2}{\int_{7\ \text{GeV}^2/c^4}^{(m_{B_s^0}-m_{K^-})^2} \frac{1}{|V_{ub}|^2}\frac{d\Gamma}{dq^2}_{B_s^0\to K^-\mu^+\nu_\mu} dq^2}, \tag{4.3}$$

and

$$R_{FF}^{LCSR} = \frac{\int_{m_\mu^2}^{(m_{B_s^0}-m_{D_s^+})^2} \frac{1}{|V_{cb}|^2}\frac{d\Gamma}{dq^2}_{B_s^0\to D_s^-\mu^+\nu_\mu} dq^2}{\int_{m_\mu^2}^{7\ \text{GeV}^2/c^4} \frac{1}{|V_{ub}|^2}\frac{d\Gamma}{dq^2}_{B_s^0\to K^-\mu^+\nu_\mu} dq^2}. \tag{4.4}$$

The values of $|V_{ub}|$ and $|V_{cb}|$ in equations 4.3 and 4.4 cancel with the terms in equations 2.75 and 2.82, producing a quantity which may be derived from theoretical calculations. The calculated form factors are,

$$R_{FF}^{LQCD} = 0.46 \pm 0.11, \tag{4.5}$$

and

$$R_{FF}^{LCSR} = 0.48 \pm 0.06. \tag{4.6}$$

When considering the decay $B_s^0 \to D_s^- \mu^+ \nu_\mu$ the integration ranges are $m^2_{\mu^+} \to (m_{B_s^0} - m_{D_s^+})^2$.

The choice to measure the decay using two bins in q^2 is motived by the fact that the form factor predictions from LQCD are most precise at high q^2 (models differ by an order of magnitude at low q^2) and the predictions from LCSR are most precise at low q^2.

The ratio of branching fractions is measured experimentally by taking the ratio of the yields of signal events and combining with their relative efficiencies and branching fractions of intermediate decays,

$$\begin{aligned} &\frac{\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} \\ &= \frac{N_{B_s^0 \to K^- \mu^+ \nu_\mu}}{N_{B_s^0 \to (D_s^- \to K^+ K^- \pi^-) \mu^+ \nu_\mu}} \cdot \frac{\varepsilon_{B_s^0 \to (D_s^- \to K^+ K^- \pi^-) \mu^+ \nu_\mu}}{\varepsilon_{B_s^0 \to K^- \mu^+ \nu_\mu}} \cdot \mathcal{B}(D_s^- \to K^+ K^- \pi^-) \end{aligned} \tag{4.7}$$

where $N_{B_s^0 \to K^- \mu^+ \nu_\mu}$ and $N_{B_s^0 \to (D_s^- \to K^+ K^- \pi^-) \mu^+ \nu_\mu}$ are the signal yields after applying all selections, $\varepsilon_{B_s^0 \to K^- \mu^+ \nu_\mu}$ and $\varepsilon_{B_s^0 \to (D_s^- \to K^+ K^- \pi^-) \mu^+ \nu_\mu}$ are the selection efficiencies for the decays $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to (D_s^- \to K^+ K^- \pi^-) \mu^+ \nu_\mu$ respectively.

The signal yields are determined by performing fits to the corrected mass distributions of selected $K^- \mu^+$ and $D_s^- \mu^+$ candidates and the efficiencies are determined from simulation after a series of data driven corrections are applied. All efficiencies are taken with respect to the specified q^2 selection.

The $D_s^- \to K^+ K^- \pi^-$ branching fraction is taken from the PDG and is a weighted average of three measurements from the CLEO, Belle, and BaBar experiments [37],

$$\mathcal{B}(D_s^- \to K^+ K^- \pi^-) = (5.44 \pm 0.18) \times 10^{-2} \tag{4.8}$$

Chapter 5

Finding $b \to u\ell\nu_\ell$ at a hadron collider

It was long thought that a measurement of $|V_{ub}|$ at a hadron collider would be impossible due to the invisible neutrino and the challenge of isolating the $b \to u$ signal from the crowded hadronic environment containing many decays with similar decay topologies. The $|V_{ub}|$ measurement using the decay $\Lambda_b^0 \to p\mu^-\overline{\nu}_\mu$ demonstrated that this was not the case [71, 72]. This chapter details the search and process for finding signal $B_s^0 \to K^-\mu^+\nu_\mu$ decays at the large hadron collider.

A discussion of semileptonic kinematics and their reconstruction are given in Section 5.1. A summary of backgrounds and the techniques used to reject them is given in Section 5.2. The use of calibration samples to model signal decays is discussed in Section 5.3. The modelling of combinatoric candidates is discussed in Section 5.4 and the simulated samples used to model the signal background are given in Section 5.5 The data processing pipeline detailing the selections is given in Section 5.6.

5.1 Kinematics

Semileptonic decays present a unique challenge at LHCb. The invisible neutrino requires that all events are partially reconstructed making it impossible to reconstruct the invariant mass of the parent decay particle. Fortunately the LHCb experiment has excellent vertex resolution allowing the B_s^0 production and decay vertices to be measured. With the knowledge of the B_s^0 flight direction

one can use the geometry of the event to measure the transverse momentum of the invisible neutrino and calculate a lower limit on the mass of the B_s^0 meson. This is called the corrected mass, and the details of its calculation are given in Section 5.1.1. Alternatively one can use the knowledge of the true mass of the B_s^0 meson to reconstruct the full kinematics of the invisible neutrino with a two fold ambiguity. The details of neutrino reconstruction are given in Section 5.1.2.

5.1.1 Corrected Mass

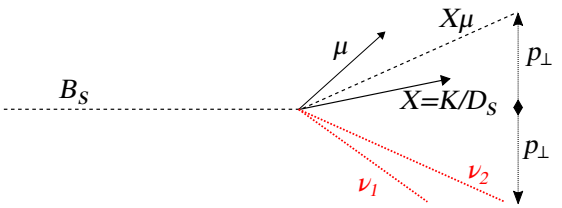


Figure 5.1 *Visualisation of conservation of momentum with respect to the B_s^0 flight direction. Two neutrino solutions are compatible with the reconstructed decay. In the B_s^0 rest frame the two solutions are back to back in the z direction, but after boosting they both travel in the positive z direction.*

The corrected mass is a lower limit on the mass of the B_s^0 momentum. As visualised in Figure 5.1, the event is rotated such that the B_s^0 meson flies in the z direction, and from the symmetric geometry of the event the transverse momentum of the neutrino must be equal and opposite to the transverse momentum of the visible system,

$$\begin{aligned} \vec{p_\perp}(K^-\mu^+) &= -\vec{p_\perp}(\nu_\mu) \\ p_\perp &\equiv |\vec{p_\perp}(K^-\mu^+)| \end{aligned} \tag{5.1}$$

The corrected mass is defined as

$$M_{corr} = \sqrt{M_{X\mu}^2 + p_\perp^2} + p_\perp, \tag{5.2}$$

with uncertainty

$$\sigma_{M_{corr}} = \left(\frac{p_\perp}{\sqrt{M_{X\mu}^2 + p_\perp^2}} + 1 \right) \sigma_{p_\perp} \tag{5.3}$$

where $M_{X\mu}$ is the invariant mass of the visible final state particles, and $p_\perp$[1] is

[1] $p_\perp$ is the momentum transverse to the B_s^0 flight directions and p_T is the momentum

the visible momentum transverse to the B_s^0 flight direction. If the only missing particle is a neutrino the corrected mass distribution will peak at the B_s^0 mass with a wide tail to the left and an immediate cut off above the mass of the B_s^0. The effects of resolution on the measurement of the B_s^0 flight direction result in tails forming above the mass of the B_s^0, the corrected mass distributions for $B_s^0 \to K^-\mu^+\nu_\mu$ and several backgrounds are plotted in Figure 5.2 before and after the simulation of resolution effects.

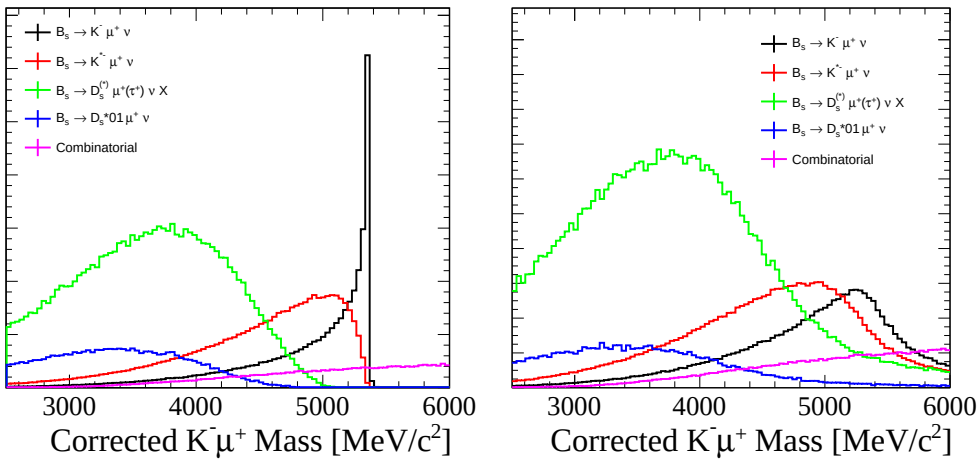


Figure 5.2 *The corrected mass distribution for simulated signal and background events reconstructed as $B_s^0 \to K^-\mu^+\nu_\mu$ before (left) and after (right) the modelling of vertex resolution.*

The dominant source of uncertainty on the corrected mass comes from an uncertainty in the B flight direction which results in a large uncertainty $p_\perp$. The uncertainty on the B flight direction must be propagated through to the uncertainty in $p_\perp$. The propagation of uncertainties to $\sigma_{p_\perp}$ is non trivial and the calculation has not been included in this thesis. The full derivation may be found in Reference [72]. The corrected mass distributions for simulated events reconstructed as $B_s^0 \to K^-\mu^+\nu_\mu$ are plotted in Figure 5.2 before and after the simulation of resolution effects, the very sharp signal peak becomes significantly broader and harder to resolve with the addition of resolution effects.

The resolution on the plotted corrected mass is significantly improved if one rejects events with a large corrected mass uncertainty. The distributions of signal Monte Carlo decays and same sign data candidates are plotted in Figure 5.3. The signal events passing the selection have a significantly sharper peak while the background sample of reconstructed $K^-\mu^+$ candidates is shifted to the left away from the signal peak. The additional resolution and separating power

transverse to the z axis.

obtained by rejecting events with a high corrected mass uncertainty result in reduced systematics when performing a fit to the corrected mass.

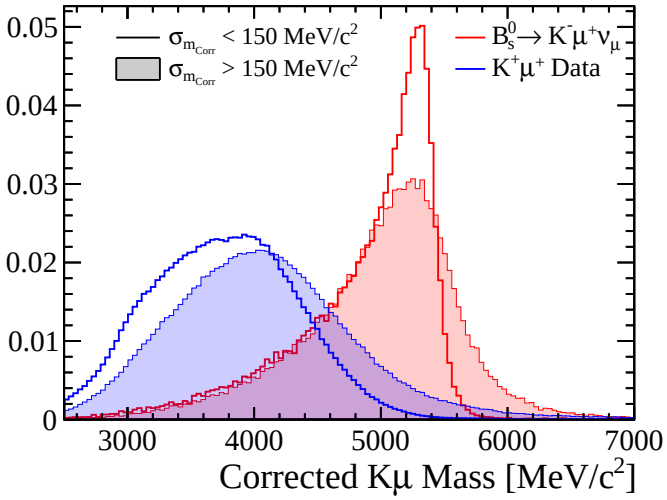


Figure 5.3 *The corrected mass distributions of Monte Carlo signal decays (red) and same sign candidates from data (blue). Events passing the $\sigma_{m_{\rm corr}}$ selection are unshaded and the events failing the selection are shaded.*

For this analysis candidates with a corrected mass uncertainty greater than $m_{\rm corr} = 150$ MeV/c^2 are rejected. This selection has an efficiency of approximately 45% for both signal and partially reconstructed background decays alike while backgrounds from combinatoric combinations are significantly reduced. Although this selection doesn't significantly increase signal purity the separation between signal and background decays is improved in the corrected mass distribution resulting in a fit with significantly reduced systematics. The efficiency of this selection is verified using a kaon and muon combination from the decay $B^+ \to (J/\psi \to \mu^+\mu^-)K^+$ and is quantified later in Section 6.5.4. The distribution of the corrected mass uncertainty is plotted in Figure 5.4 for signal Monte Carlo and the $K^-\mu^+$ combination from $B^+ \to J/\psi\, K^+$ using Monte Carlo and data. As the dominant source of uncertainty on the B_s^0 flight direction originates from the precision on the primary and secondary vertices a B_s^0 meson with a longer flight distance will have a lower corrected mass uncertainty, consequently the application of this selection will introduce a bias on the measured flight distance or calculated decay time. This selection is very effective at rejecting backgrounds from combinatoric combinations of a $K^-\mu^+$ pair. Combinatorics originate from two sources, the combination of prompt tracks originating from the primary vertex or from $b\bar{b}$ production with one b decaying semileptonicaly producing a muon and the other decaying hadronically producing a kaon.

When considering the decay $B_s^0 \to D_s^- \mu^+ \nu_\mu$ no selection is made on the corrected mass uncertainty of the $K^-\mu^+$ pair or the $D_s^-\mu^+$ pair. The dominant background to the decay $B_s^0 \to D_s^- \mu^+ \nu_\mu$ is $B_s^0 \to D_s^{*-} \mu^+ \nu_\mu$ and a selection on the corrected mass uncertainty does little to further separate the two decays.

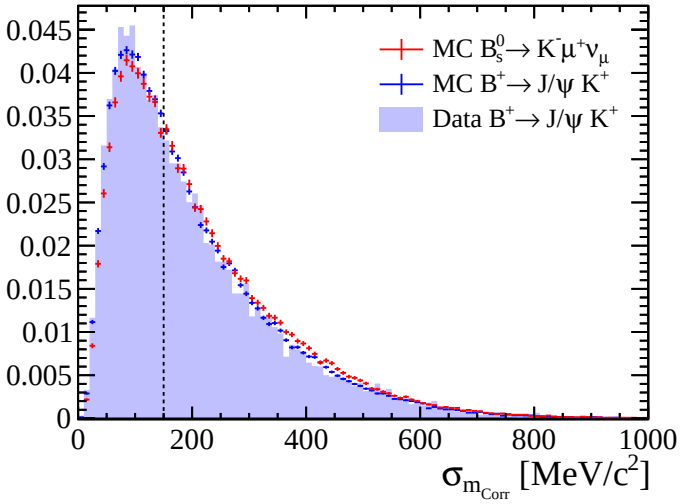


Figure 5.4 *The corrected mass uncertainty for signal decays and $B^+ \to J/\psi K^+$ decays reconstructed as $B_s^0 \to K^-\mu^+\nu_\mu$.*

5.1.2 Neutrino Reconstruction and q^2

The $|V_{ub}|$ measurement will be performed by measuring the signal yield of $B_s^0 \to K^-\mu^+\nu_\mu$ candidates in two regions of phase space, separated by $q^2 = 7$ GeV^2, where q^2 is the squared four vector momentum recoiling off the B_s^0, which is equal to the four momentum squared of the $\mu^+\nu_\mu$ combination. A calculation of q^2 first requires the neutrino momentum be reconstructed. The component of the neutrino momentum transverse to the B flight direction, $p_\perp$ is equal and opposite the transverse momentum of the $K^-\mu^+$ pair. The longitudinal component, $p_\parallel$, may be determined to a two-fold ambiguity with the quadratic equation

$$p_\perp = p_\perp(K^-\mu^+) \tag{5.4}$$

$$p_\parallel = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \tag{5.5}$$

where a, b and c are determined as

$$\begin{aligned} a &= |2p_{\parallel} m_{K\mu}|^2, \\ b &= 4p_{\parallel}(2p_{\perp}p_{\parallel} - m^2_{\text{miss}}), \\ c &= 4p_{\perp}(p^2_{\parallel} + m^2_{B^0_s}) - |m^2_{\text{miss}}|^2, \\ m^2_{\text{miss}} &= m^2_{B^0_s} - m^2_{K\mu}. \end{aligned} \tag{5.6}$$

The kinematics of the B^0_s meson and q^2 of the event may now be calculated with a two fold ambiguity [73, 74]. When performing a physics analysis it is desirable to resolve this ambiguity without the introduction of a bias in q^2. A choice must be made on which of the two solutions of q^2 will be used when performing an analysis.

The simplest approach is to randomly select one of the two solutions which while unbiased has a poor resolution in q^2. A significantly improved method uses a linear regression model to predict the B^0_s momentum and the ambiguity is resolved by selecting the solution most consistent with the regression value. The full details of the regression method are given below.

Due to the detector resolution effects approximately 20% of the candidates have an unphysical solution (i.e. $b^2 < 4ac$) for $P_{\parallel}$. The unphysical events fall into corrected mass region above the B^0_s invariant mass, $m_{corr}(B^0_s) > m(B^0_s)$, and are removed when restricting events to a specific region in q^2.

5.1.3 Linear Regression to Reconstruct q^2

Linear regression analysis is a statistical technique for predicting the value of a target or response variable based on relationships with predictor or regressor variables [75–77]. For this analysis the momentum of the B^0_s is inferred from the flight distance and polar angle of the B^0_s with a resolution of 60% which is sufficient to select the correct solution of the quadratic equation 70% of the time [78], compared to the random selection which selects the correct solution 50% of the time.

The B^0_s momentum is weakly correlated to its polar angle, θ_{flight},

$$P = \frac{p_T}{\sin\theta_{\text{flight}}}, \tag{5.7}$$

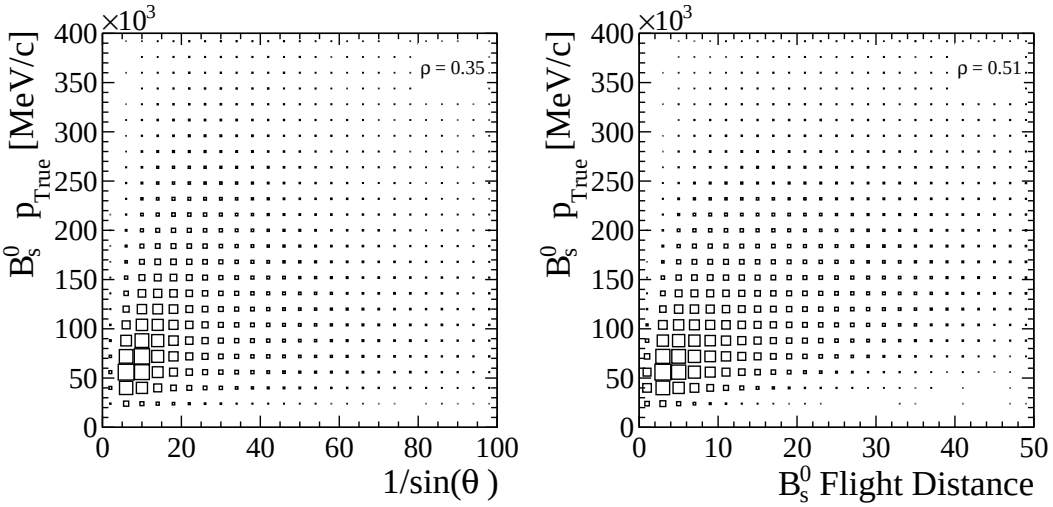


Figure 5.5 *Input variables used to predict the B_s^0 momentum with a linear regression model.*

and flight distance, $|\vec{F}|$, and decay time, t,

$$P = \frac{M|\vec{F}|}{t}, \tag{5.8}$$

as shown in Figure 5.5. The two flight variables are considered in a least squares linear regression model [79]

$$P = \beta_0 + \beta_1 / \sin\theta_{\text{flight}} + \beta_2 |\vec{F}| + \varepsilon, \tag{5.9}$$

where β_n are parameters to be determined, and ε is a random component with a mean of 0 and variance equal to the variance of the predicted momentum. The predicted value of the B_s^0 momentum is compared to the two solutions derived from the quadratic equation defined in Section 5.1.2 and the solution closest to the regression value, ${q^2}_{\text{Best}}$, is selected. The use of regression in the selection of a solution to the quadratic equation significantly improves the resolution on the reconstructed true q^2 as plotted in Figure 5.6. The resolution on the reconstructed q^2 for different methods of selecting a solution is given in Table 5.1. Using the output of the linear regression model to select a solution improves the resolution on the reconstructed q^2 by 38% when compared to a random selection.

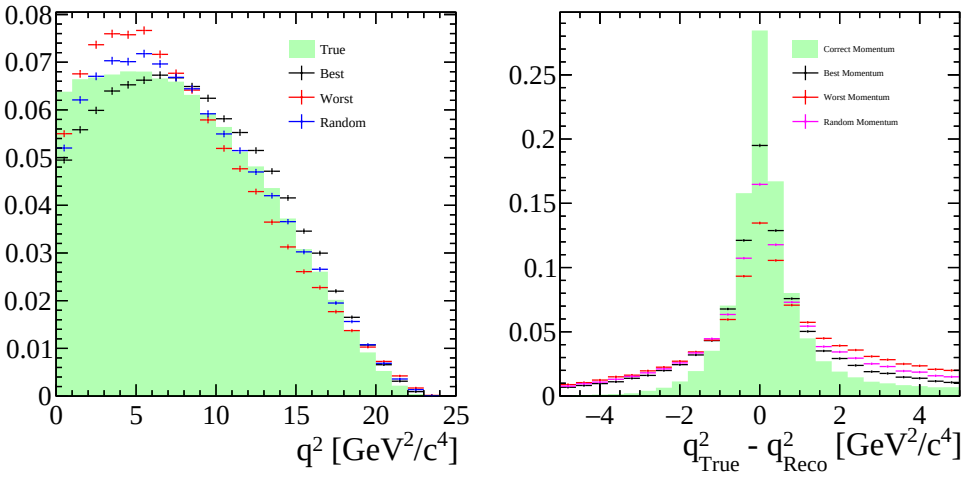


Figure 5.6 *True q^2 distributions from Monte Carlo (shaded green) with the reconstructed q^2 with different methods of selecting the B_s^0 momentum solution (left) and the resolution on q^2 is shown using different selections (right). The best solution is the solution closest to the regression value, and the worst solution is furthest from the regression value.*

Solution	RMS
Correct	1.07 $\mathrm{GeV}^2/\mathrm{c}^4$
Regression	2.21 $\mathrm{GeV}^2/\mathrm{c}^4$
Random	3.06 $\mathrm{GeV}^2/\mathrm{c}^4$
Incorrect	4.23 $\mathrm{GeV}^2/\mathrm{c}^4$

Table 5.1 *Resolution on reconstructed q^2 after selecting one of the two solutions. Resolutions are given for the correct solution, solution obtained from regression, randomly selecting a solution and the incorrect solution.*

5.2 Backgrounds

Reconstruction of the decay $B_s^0 \to K^-\mu^+\nu_\mu$ faces large backgrounds at LHCb. A significant number of events contain an opposite sign $K^-\mu^+$ pair which may be reconstucted as a $B_s^0 \to K^-\mu^+\nu_\mu$ candidate. These backgrounds include partially reconstructed B hadron decays with additional charged or neutral final state particles, random combinatoric combinations of a $K^-\mu^+$ pair and decays with a misidentified particle. These are simulated using Monte Carlo or estimated using background data samples.

The dominant source of backgrounds that are selected when constructing $B_s^0 \to K^-\mu^+\nu_\mu$ candidates are partially reconstructed B hadron decays with additional unreconstructed charged tracks. As visualised in Figure 5.7, the most concerning of these backgrounds is $B^+ \to J/\psi\, K^+$ with an unreconstructed muon

which has a topology almost identical to $B_s^0 \to K^-\mu^+\nu_\mu$ with an unreconstructed neutrino, with the main difference being the invariant mass of the reconstructed particle and the q^2 of the two decays. The largest source of backgrounds with additional charged tracks correspond to semileptonic $b \to c\ell\nu_\ell$ transitions of a B hadron decaying into a charm meson with the decay of the charm containing a charged kaon. The inclusive semileptonic B branching fraction is approximately 11% [37] compared to the $B_s^0 \to K^-\mu^+\nu_\mu$ branching fraction of approximately 0.015%. Backgrounds containing additional charged tracks may be significantly reduced by searching through the other charged tracks in the event for tracks compatible with the candidate vertex. This process is fully detailed in Section 5.6.7.

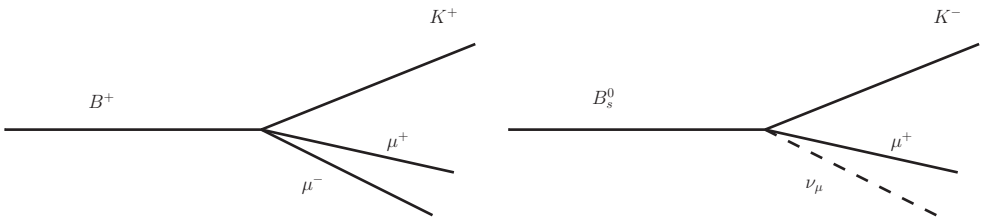


Figure 5.7 *The topology of $B^+ \to J/\psi K^+$ (left) and $B_s^0 \to K^-\mu^+\nu_\mu$ (right). When partially reconstructed with one missing lepton the decays are almost identical.*

Backgrounds containing unreconstructed neutral final state particles present a greater challenge. The reconstruction efficiency of low transverse momentum neutral tracks is low, approximately 20% [80], which makes the reduction of such backgrounds a challenge. Higher mass resonances of the K^- and D_s^- for the signal and normalisation decays produce soft, low momentum, neutral particles. These backgrounds may be reduced using a cone isolation procedure. A cone is drawn around the candidate tracks in $\Delta R = \sqrt{\Delta\eta + \Delta\phi}$, where η is the pseudorapidity and ϕ is the radial angle, and the activity of the ECAL and HCAL within the cone is investigated. It is expected that tracks originating from the decay of higher mass resonances will have increased deposits in the calorimeters close to the candidate track. The use of cone isolation is detailed fully in Section 5.6.7.

Given that the rate of pions from B decays is much larger than the rate of kaons there is a substantial background of pions, and to a lesser extent protons, electrons and muons, which will be falsely identified as a kaon. backgrounds from misidentified particles are substantially reduced by requiring that the muon and kaon candidates have high muon and kaon likelihood respectively when reconstructing as $B_s^0 \to K^-\mu^+\nu_\mu$.

The final major background under consideration is the combinatoric combinations of kaon and muon candidates. Combinatoric backgrounds are effectively reduced with selections using topological information and vertex quality criteria.

5.3 Calibration Samples

In order to verify the accuracy of simulated Monte Carlo samples and evaluate systematics and corrections due to mismodelling in the simulation, calibration samples analogous to the signal decay are used. When considering the decay $B_s^0 \to K^-\mu^+\nu_\mu$ the decay $B^+ \to J/\psi\, K^+$ is used as a calibration sample. The B^+ decay chain may be reconstructed either by explicitly searching for a $K^+\mu^+\mu^-$ final state compatible with a $B^+ \to J/\psi\, K^+$ decay or through the reconstruction of a $K^-\mu^+$ pair analogous to the decay $B_s^0 \to K^-\mu^+\nu_\mu$, with the additional muon found through the use of the isolation BDT detailed in Section 5.6.7. The former method of reconstruction is useful for validating the efficiencies of selections dependant on the underlying event as the $B_s^0 \to K^-\mu^+\nu_\mu$ and $B^+ \to J/\psi\, K^+$ are exclusively reconstructed and the underlying event will contain no tracks compatible with the signal decays. The latter method is useful when validating the kinematics of the decay because a $B_s^0 \to K^-\mu^+\nu_\mu$ decay with a non reconstructed neutrino is very difficult to distinguish from a $B^+ \to J/\psi\, K^+$ decay with a non reconstructed muon. The decay $B^+ \to J/\psi\, K^+$ has a high yield and it is possible to generate a highly pure data sample with very few backgrounds making it ideal for comparisons with pure Monte Carlo. The decay $B^+ \to J/\psi\, K^+$ is used to validate the Monte Carlo simulations by comparing the distributions of kinematic variables and the calculation of selection efficiencies.

5.4 Combinatoric Modelling

Due to the nature of partially reconstructed B hadron decays there is no calibration sample from which a representative sample of combinatoric combinations of kaons and muons may be obtained, although there are regions of phase space where a pure sample of combinatorics may be obtained. By requiring that $m_{K\mu} > m_{B_s^0}$, a pure combinatoric sample may be produced however the corrected mass cannot be extrapolated to lower values below the mass of the B_s^0 making the sample irrelevant when considering the full phase space. For this analysis

combinatorics are modelled using a procedure termed *event mixing*. A candidate kaon track from one event is combined with a candidate muon from another event and a new B_s^0 candidate is reconstructed from this combination. The effectiveness of this method is validated by comparing the kinematics of the mixed events with candidates in data in the region $m_{K\mu} > m_{B_s^0}$.

A new combinatoric candidate is constructed by *mixing* a kaon and muon candidate from different events forming a B_s^0 candidate simulating a combinatoric combination of the kaon and muon. When reconstructing the *mixed* B_s^0 the momenta of the kaon and muon and primary vertex location are taken from different events. The secondary vertex location is chosen by randomly sampling the flight distance of reconstructed B_s^0 candidates with $m_{K\mu} > m_{B_s^0}$ and placing it at that distance downstream of the primary vertex. The corrected mass is determined from these quantities. The uncertainty on the corrected mass is determined by randomly displacing the secondary vertex 500 times, for each displacement the corrected mass is calculated and the uncertainty is the standard deviation on the 500 corrected mass values.

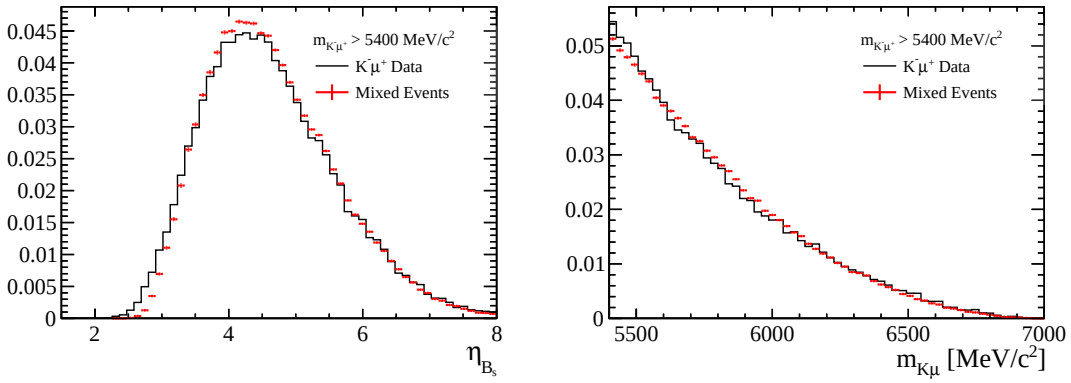


Figure 5.8 *The distributions of $K^-\mu^+$ candidates in data (black) are plotted alongside simulated combinatoric candidates (red). All distributions are drawn requiring that the reconstructed $K^-\mu^+$ mass is greater than $m_{B_s^0}$.*

This method of modelling combinatorics does not accurately reproduce the kinematics of the B_s^0 meson in the region $m_{K\mu} > m_{B_s^0}$. A two dimensional reweighting is used to correct the momentum and transverse momentum of the B_s^0 candidate. The distributions of the $K^-\mu^+$ invariant mass and corrected mass do not change as a result of the reweighting indicating the shapes of the distributions are robust. The invariant mass and corrected mass distributions with $m_{K\mu} > 5400$ MeV/c^2 are plotted in Figure 5.8 for mixed events and data.

The details of the reweighting procedure and validation plots may be found in Appendix B.

Combinatoric candidates originating from the decay of a $b\bar{b}$ pair will produce tracks with a large opening angle, when looking down the beam line or z axis, the b and $\bar{b}$ will be produced *back to back* and will have opposite momenta in the transverse plane. The large opening angle obtained when selecting a kaon and muon candidate from different quarks in the $b\bar{b}$ pair produces B_s^0 candidates with a large invariant mass. This feature may also be used to effectively reject combinatoric backgrounds and is discussed further in Section 5.6.3.

The corrected mass of the $K^-\mu^+$ pair obtained from the event mixing procedure is plotted in Figure 5.9 alongside data obtained from two different triggers. Therein lies a sensitive topic with respect to the trigger. As will be discussed later in Section 5.6.2, the trigger used for this analysis is a topological trigger which uses a multivariate selection trained to select partially reconstructed B hadron candidates with two visible final state particles. The topological trigger uses the corrected mass of the candidate decay in its multivariate selection resulting in a reduction of reconstructed candidates with $m_{\text{corr}} > 5800$ MeV/c^2. Without access to the trigger software it is impossible to fully reproduce the behaviour of the topological trigger above $m_{\text{corr}} = 5800$ MeV/c^2. An additional trigger is available which selects candidates based solely on the muon and is unbiased in m_{corr}. This trigger has significantly lower statistics. The corrected mass distribution of the mixed events and two trigger are plotted in Figure 5.9. Above the region in m_{corr} where no decays from B hadrons are present the corrected mass of the mixed samples agrees incredibly well with the data originating from the muon trigger. In addition the m_{corr} distributions for both triggers are very similar below 5800 MeV/c^2 indicating that the topological trigger does not significantly impact the shape of the corrected mass distribution. By combining the above two arguments it is decided that the use of mixing to model combinatoric combinations of a $K^-\mu^+$ pair selected using the topological trigger is valid up to $m_{\text{corr}} = 5800$ MeV/c^2. Therefore all corrected mass distributions used in this analysis will end at $m_{\text{corr}} = 5800$ MeV/c^2.

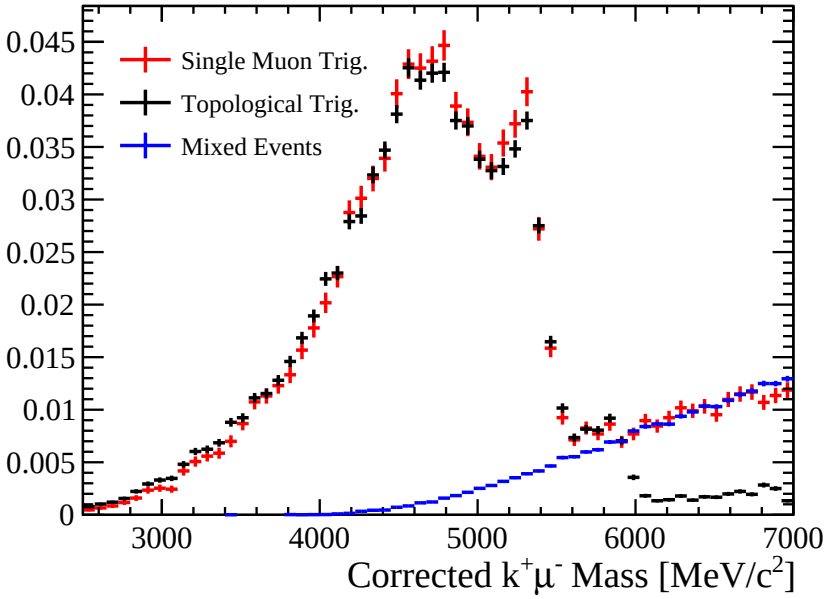


Figure 5.9 *The corrected mass of a $K^-\mu^+$ pair is plotted for candidates modelling combinatoric backgrounds.*

5.5 Simulated Samples

A huge volume of simulated Monte Carlo decays were produced for this analysis in order to model the signal, normalisation and background decays. These are detailed in Table 5.2. Signal $B_s^0 \to K^-\mu^+\nu_\mu$ Monte Carlo samples are generated exclusively using the Isgur-Scora-Grinstein-Wise updated model [81, 82] to model the form factors. $B_s^0 \to K^-\mu^+\nu_\mu$ background samples include excited K^+ resonances decaying with additional neutral tracks. Monte Carlo samples are used to model the massive contributions from B hadrons decaying to charm hadrons which ultimately decay producing a kaon. The $B_s^0 \to D_s^-\mu^+\nu_\mu$ simulated samples contain a cocktail of $B_s^0 \to D_s^-\mu^+\nu_\mu X$ decays with each event given a flag corresponding to the correct decay type. The cocktail contains the exclusive $B_s^0 \to D_s^-\mu^+\nu_\mu$ decay in addition to decays including excited D_s^+ resonances and tauonic decay modes. Two cocktail $B_s^0 \to D_s^-\mu^+\nu_\mu$ samples were produced, one which had been accidentally produced with an incorrect modelling of the form factors.

Both samples are used in this analysis. For partially reconstructed decays with more than one missing particle the corrected mass is, to first order, independent of q^2. The form factors are defined in terms of q^2, resulting in a corrected mass distribution which is independent of the form factor parametrisation. Both

samples are used when drawing histograms of the corrected mass distribution for the background sample. When calculating efficiencies and generating template shapes of the signal mode, only the model with the correct form factor parametrisation is used.

5.6 Selections for $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$

5.6.1 Data Pipeline

In order for a physics analysis to return meaningful results, the data collected must go through multiple stages of processing. The first stage of processing is the hardware trigger which identifies the events containing potentially interesting physics. All events passing at least one hardware trigger are temporarily stored before processing by the software trigger. The software trigger combines information from multiple subsystems of the detector and reconstructs candidate tracks and B hadrons. Events passing at least one software trigger are stored and all others are permanently deleted. The dataset passing the trigger selection for 2012 contains 28×10^9 events, and due to size of this dataset it is inaccessible to analysts and is processed centrally every two years. The central processing of the triggered data is referred to as the stripping. During the stripping, candidate signal decays are built by combining different tracks from the event. For $B_s^0 \to K^-\mu^+\nu_\mu$ the reconstruction will search through all combinations of opposite sign kaon and muon pairs and reconstruct a candidate B_s^0. The selections applied during the stripping are designed to be *loose* allowing a highly efficient selection of signal events with sufficient background to ensure meaningful background studies may be performed.

The DaVinci software package [63] performs the next stage of offline processing, iterating through the events passing the desired stripping selections and saving the selected events to a local file. The DaVinci package takes as an input the raw reconstructed tracks and candidates of the event and outputs an organised, formatted NTuple containing event by event data of the candidate tracks and underlying event. The next stage of processing applies the first round of selections, applies multivariate classifications and calculates weights for the kinematic correction of Monte Carlo. During the final stage of processing input histograms for use during fitting and validations are drawn, final tight selections

Decay	Year	Size	
Signal $B_s^0 \to K^- \mu^+ \nu_\mu$ decays plus excited K^- resonances.			
$B_s^0 \to K^- \mu^+ \nu_\mu$	2012/2011	8M - 3 fb^{-1}	
$B_s^0 \to K^{*-} \mu^+ \nu_\mu$	2012	4M - 1 fb^{-1}	
$B_s^0 \to K_2^{*-} \mu^+ \nu_\mu$	2012/2011	6M - 1 fb^{-1}	
$B_s^0 \to K_0^{*-} \mu^+ \nu_\mu$	2012/2011	6M - 1 fb^{-1}	
Signal $B_s^0 \to D_s^- \mu^+ \nu_\mu$ decays. Samples contain a cocktail of $B_s^0 \to D_s^+$ decays.			
$B_s^0 \to D_s^+ \mu^+ \nu_\mu X$	2012/2011	9M - 0.2 fb^{-1}	
$B_s^0 \to D_s^+ \mu^+ \nu_\mu X$	2012	20M - 0.4 fb^{-1}	Incorrect FF
Charged backgrounds to $B_s^0 \to K^- \mu^+ \nu_\mu$ used as calibration samples.			
$B^+ \to J/\psi K^+$	2012	20M - 2 fb^{-1}	
$B_s^0 \to J/\psi \phi$	2012	100M - 40 fb^{-1}	
Inclusive $b \to c$ (s) decays modelling $K^- \mu^+$ backgrounds.			
$b \to c \to K^+ \mu X$	2012/2011	6M - 0.01 fb^{-1}	
$b \to c \to K \mu X$	2012/2011	250k - 0.01 fb^{-1}	Filtered MC
$b \to K^+ \mu^+ X$	2012/2011	640k - 0.03 fb^{-1}	Filtered MC
$b \to c \to K^+ \mu X$	2012/2011	1.5M - 0.05 fb^{-1}	Filtered MC
Background decays reconstructible as $B_s^0 \to K^- \mu^+ \nu_\mu$ with misidentification.			
$\Lambda_b^0 \to p \mu \nu$	2012	15M	
$B^0 \to \pi \mu \nu$	2012	4M	
$B^+ \to \rho \mu \nu$	2012	5M	
$B^0 \to \rho^- \mu^+ \nu$	2012	4M	
$B_s^0 \to K^- \mu^+ \nu_\mu$ backgrounds containing additional charged tracks.			
$B^0 \to J/\psi K^*$	2012	10M	
$B^+ \to J/\psi K^*$	2012	24M	
$B^0 \to D^* \pi, D^* \to D^0 \pi$	2012	140k	
$B^0 \to D \pi, D \to K \mu \nu$	2012	160k	
$B^0 \to D^* \mu \nu X, D \to K \pi \pi \pi$	2012	5M	
$B^0 \to D^* \mu \nu X, D \to K \pi$	2012	7M	
$B^0 \to D \mu \nu X, D \to K \pi \pi$	2012	20M	
$B^+ \to D^0 \pi$	2012	140k	
$B^+ \to D^{*0} \pi, D^{*0} \to D^0 \pi$	2012	140k	
$B^+ \to D^{*0} \pi, D^{*0} \to D^0 \gamma$	2012	140k	
$B^+ \to D^0 \rho$	2012	140k	
$B^+ \to D^0 \mu \nu$	2012	15M	
$B_s^0 \to D^0 K^{*0}$	2012	150k	
$B_c^+ \to D^0 \mu \nu$	2012	1M	
$B_s^0 \to D_s^- \mu^+ \nu_\mu$ backgrounds containing additional charged tracks.			
$B^+ \to D_s^{*+} D^*$	2012	5M	
$B^+ \to DD$	2011	3M	
$B_s^0 \to D^0 D_s^+ K$	2012	5M	
$B_s^0 \to D_s^{*+} D_s^{*+}$	2012	5M	
$B_s^0 \to DD$	2011	1M	
$B^0 \to D_s^+ D^*$	2012	5M	
$B^0 \to DD$	2011	2M	
$B_c^+ \to J/\psi D_s^+$	2012	3M	

Table 5.2 *A summary of the simulated samples used in this Analysis. The sample size for filtered events is counted after stripping selections are applied. An approximate luminosity is included for the most significant samples.*

are applied and additional Monte Carlo corrections are calculated and applied.

5.6.2 Preselection

The selections applied to signal candidates are optimised to maximise signal efficiency and background rejection by exploiting topological differences between signal and background events. Signal events containing long lived particles have final state particles originating a significant distance from the interaction point with a high transverse momentum. Background events and candidates typically contain prompt tracks originating from the interaction point.

For a full understanding of the selections applied several variables must be defined:

DOCA	Distance of closest approach of two lines or particle tracks.
IP	Impact parameter. The distance between a track and vertex at closest point.
IPχ^2	Impact parameter chi-squared. The difference in χ^2 of the primary vertex reconstructed with and without the candidate track.
DIRA	The cosine of the angle formed between the direction of the measured momentum of a decaying particle and the line formed from the production and decay vertices.
FD	The flight distance, or distance between production and decay vertex.
FDχ^2	The flight distance chi-squared. The difference in χ^2 of the SV fit reconstructed with or without the requirement of zero flight distance.

The HLT2 trigger `TopoMu2BodyBBDT` is used to select candidates for both signal and normalisation decays and requires as an input a muon having passed the L0 and HLT1 single muon triggers with a minimum transverse momentum of 1.76 GeV [57]. The `TopoMu2BodyBBDT` trigger is designed to select partially reconstructed decays of B hadrons containing a muon candidate [59, 83]. The trigger algorithm requires a displaced secondary vertex and a candidate is built from the muon and the additional particle. A bonsai boosted decision tree (BBDT) is employed to efficiently select signal events using discretised kinematics of the candidate [84]. BDTs are detailed in Section 5.6.5. A BDT is used

Variable	Selection	BBDT Intervals
$\sum \lvert p_{\mathrm{T}} \rvert$[GeV/$c$]	> 3	3.5, 4, 4.5, 5, 6, 7, 8, 9, 10, 15, 20
${p_{\mathrm{T}}}^{\mathrm{min}}$[GeV/$c$]	> 0.5	0.6, 0.7, 0.8, 0.9, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 5, 10
m[GeV/c^2]	< 7	2.5, 4.75
m_{corr}[GeV/c^2]		2, 3, 4, 5, 6, 7, 8, 9, 10, 15
DOCA [mm]	< 0.2	0.05, 0.1, 0.15
IPχ^2		20
FDχ^2/100	> 1	2, 3, 4, 5, 6, 7, 8, 9, 10, 25, 50, 100

Table 5.3 *The variables and intervals used in the BBDT for trigger* `TopoMu2BodyBBDT` *selecting 2 body decays. Table taken from [84].*

in the trigger to identify signal candidates as it has a higher signal efficiency and improved background rejection compared to a simple cut based trigger. To prevent the BDT from making a series of overly complicated selections the input variables to the trigger are discretised and the BDT may only introduce selections at the specific intervals listed in Table 5.3. The discretisation earns the BBDT it's bonsai name. During optimisation of the BBDT the number of intervals was gradually reduced until a decrease in performance was observed.

Two stripping lines were developed and used for this analysis designed to efficiently select $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ candidates, the stripping selections are detailed in Tables 5.5 and 5.6. In addition to the two lines a suite of lines were based on the lines to select events in background enriched regions by removing the likelihood hypothesis selections on certain particles, or inverting the charge on the muon. The stripping lines are summarised in Table 5.4. Selections are made on the track properties by requiring a high track quality, low ghost probability and high particle identification (PID) likelihood ($\mathcal{L}$). Topological cuts require that the final state tracks are well isolated from the primary vertex, the fit quality on the reconstructed vertex is high, the secondary vertex is well separated from the primary vertex and the cosine of the angle between the reconstructed flight direction of the B_s^0 and its measured momentum, DIRA, is close to 1. It is often required to reduce the rate of a stripping line if the output rate is exceptionally high, the selection algorithm is computationally intensive or the perceived value of the selected data is low. The rate of a line is reduced by applying a prescale which runs the stripping algorithms on a random subset of the data.

Line	Purpose	Prescale
B2DMuNuX_Ds	Signal $B_s^0 \to D_s^- \mu^+ \nu_\mu$ candidates	1.0
B2DMuNuX_Ds_FakeMu	$B_s^0 \to D_s^- \mu^+ \nu_\mu$ with misidentified muon	0.02
B2XuMuNuBs2K	Signal $B_s^0 \to K^- \mu^+ \nu_\mu$ candidates	1.0
B2XuMuNuBs2KSS	$K^\pm$ $\mu^\pm$ candidates to model backgrounds	0.1
B2XuMuNuBs2K_FakeMu	$B_s^0 \to K^- \mu^+ \nu_\mu$ with misidentified muon	0.02
B2XuMuNuBs2K_FakeK	$B_s^0 \to K^- \mu^+ \nu_\mu$ with misidentified kaon	0.02
B2XuMuNuBs2K_FakeKMu	$B_s^0 \to K^- \mu^+ \nu_\mu$ with misidentified K/μ	0.02

Table 5.4 *A summary of the stripping lines used to select signal and background candidates. The $D_s^- \mu$ lines select same sign and opposite sign candidates. The Fake lines are identical to the candidate lines except the likelihood selections are removed.*

5.6.3 Background Vetoes

A number of vetoes are applied to explicitly remove certain backgrounds in an efficient manner. These backgrounds may take the form of combinatorics, misidentified decays with no additional tracks or reconstructible decays from decays of higher excitations, e.g. $B_s^0 \to K^{*-} \mu^+ \nu_\mu$.

Decays where the K^- and μ^+ originate from the same charm meson decay are rejected by requiring that y./the invariant mass of the $K^- \mu^+$ pair is greater than the mass of D mesons. The decay $J/\psi \to \mu^+ \mu^-$ may be selected and reconstructed as a signal $B_s^0 \to K^- \mu^+ \nu_\mu$ decay if one muon is misidentified as a kaon. Candidates containing a kaon which penetrates muon the system are rejected if the $K^- \mu^+$ invariant mass reconstructed under the $\mu^+ \mu^-$ mass hypothesis is consistent with the J/ψ mass. Kaons produced from the decays of excited kaons, $K^{*-} \to K^- \pi^0$, are rejected by searching for neutral pions in a cone around the kaon track. The candidate is rejected if a pion is found and the invariant mass of the $K^- \pi^0$ pair is consistent with the K^{*-} or $K^{*-}(1430)$. This selection only rejects $\approx 20\%$ of the background from higher excited resonances due to the low reconstruction efficiency of soft pions [55].

Combinatorics arising from $b\bar{b}$ production with the kaon and muon originating from the decay of the different b quarks may be rejected by exploiting the topology of $b\bar{b}$ production. The two b quarks fragment into B hadrons and in the rest frame recoil off one another resulting in two B mesons with opposite momenta. If the $b\bar{b}$ pair is boosted in the longitudinal direction, which is approximately the case at LHCb, the recoil of fragmentation will result in the B mesons having opposite momenta in the transverse plane. The event is rejected if the kaon and muon

Candidate	Selection
Stripping Preselection	
Event	Long track multiplicity < 250
muon	Track quality $\chi^2/N_{dof} < 4.0$
muon	$p > 6000$ MeV
muon	$p_{\mathrm{T}} > 1500$ MeV
muon	Track ghost probability< 0.35
muon	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_\pi > 3.0$
muon	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_p > 0.0$
muon	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_K > 0.0$
muon	Primary vertex $\mathrm{IP}_{\chi^2} > 16$
kaon	Track quality $\chi^2/N_{dof} < 6.0$
kaon	$p > 10000$ MeV
kaon	$p_{\mathrm{T}} > 500$ MeV
kaon	Track ghost probability< 0.5
kaon	$\ln \mathcal{L}_K - \ln \mathcal{L}_\pi > 5.0$
kaon	$\ln \mathcal{L}_K - \ln \mathcal{L}_p > 5.0$
kaon	$\ln \mathcal{L}_K - \ln \mathcal{L}_\mu > 5.0$
kaon	Primary vertex $\mathrm{IP}_{\chi^2} > 16$
B_s^0	SV quality $\chi^2/N_{dof} < 4.0$
B_s^0	DIRA > 0.994
B_s^0	SV separation from PV $\chi^2 > 120$
B_s^0	2500 MeV $< m_{\mathrm{corr}} <$ 7000 MeV
Vetoes	
Candidates per event	1
$J/\psi \to \mu^+\mu^-$ misid veto	$\lvert m_{(K\to\mu)\mu} - m_{J/\psi}\rvert < 30$ MeV and K_{IsMuon}
$D \to \to K^-\mu^+$ X rejection	$m_{K\mu} > 1900$ MeV
$K^{*-}(892) \to K^-\pi^0$ veto	$\lvert m_{K\pi^0} - m_{K^*}\rvert < 65$ MeV
$K^{*-}(1430) \to K^-\pi^0$ veto	$\lvert m_{K\pi^0} - m_{K^*(1430)}\rvert < 90$ MeV
Combinatoric quadrant veto	$P_x(K) \times p_x(\mu) < 0$ and $P_y(K) \times p_y(\mu) < 0$
BDT Selections	
Isolation preselection	`min(IsoMinBDT_K, IsoMinBDT_Mu) > -0.9`
Charged background BDT	`Charge_BDT > 0.1`
Same sign BDT	`SameSign_BDT > 0.05`
Additional Selections	
Corrected mass error	$\sigma_{m_{corr}} < 150$ MeV/c^2
q^2 boundary	$q^2 \gtrless 7$ GeV$^2/c^4$

Table 5.5 *All selections applied to the $B_s^0 \to K^-\mu^+\nu_\mu$ candidates using the* `B2XuMuNuBs2K` *stripping line.*

Candidate	Selection
Stripping Preselection	
Event	Long track multiplicity < 250
muon	$p_{\mathrm{T}} > 1000$ MeV
muon	$p > 6000$ MeV
muon	Track ghost probability< 0.35
muon	Track quality $\chi^2/N_{dof} < 3.0$
muon	Primary vertex $\mathrm{IP}_{\chi^2} > 12$
muon	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_\pi > 3.0$
kaon	$p_{\mathrm{T}} > 250$ MeV
kaon	$p > 2000$ MeV
kaon	Track ghost probability< 0.35
kaon	Track quality $\chi^2/N_{dof} < 3.0$
kaon	Primary vertex $\mathrm{IP}_{\chi^2} > 4$
kaon	$\ln \mathcal{L}_K - \ln \mathcal{L}_\pi > -2.0$
pion	$p_{\mathrm{T}} > 250$ MeV
pion	$p > 2000$ MeV
pion	Track ghost probability< 0.35
pion	Track quality $\chi^2/N_{dof} < 3.0$
pion	Primary vertex $\mathrm{IP}_{\chi^2} > 4$
pion	$\ln \mathcal{L}_K - \ln \mathcal{L}_\pi < 20.0$
D_s^+	$\lvert m_{\mathrm{Cand.}} - m_{D_s^+} \rvert < 80$ MeV
D_s^+	DOCA $\chi^2 < 20$
D_s^+	Vertex $\chi^2/N_{dof} < 6.0$
D_s^+	$FD_{\chi^2} > 25$
D_s^+	DIRA> 0.99
B_s^0	2200 MeV $< m_{\mathrm{Cand}} < 8000$ MeV
B_s^0	Vertex $\chi^2/N_{dof} < 9.0$
B_s^0	DIRA> 0.999
B_s^0	$\mathrm{Vertex}(D_s^+)_Z - \mathrm{Vertex}(B_s^0)_Z > -0.05$
Additional Selections	
K^-	$\ln \mathcal{L}_K - \ln \mathcal{L}_\pi > 5.0$
K^-	$\ln \mathcal{L}_K - \ln \mathcal{L}_p > 5.0$
K^-	$\ln \mathcal{L}_K - \ln \mathcal{L}_\mu > 5.0$
K^-	$p > 10000$ MeV
μ^+	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_\pi > 3.0$
μ^+	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_p > 0$
μ^+	$\ln \mathcal{L}_\mu - \ln \mathcal{L}_K > 5.0$
Vetoes	
$D^{*-} \to D^0\pi^+$ Veto	$m_{KK\pi} - m_{KK} > 138$ MeV
$B_s^0 \to D_s^-\pi^+$ Veto	$\lvert m_{Ds(\mu\to\pi)} - m_{B_s^0} \rvert > 70$ MeV
BDTs	
Isolation	`min(IsoMinBDT_K, IsoMinBDT_Mu) > -0.8`

Table 5.6 *All selections applied to $B_s^0 \to D_s^-\mu^+\nu_\mu$ candidates using the `B2DMuNuX_Ds` line. Selections are aligned with those for $B_s^0 \to K^-\mu^+\nu_\mu$ as closely as possible.*

candidates are in opposite quadrants of the xy plane as visualised in Figure 5.10. All vetoes are summarised in Table 5.5.

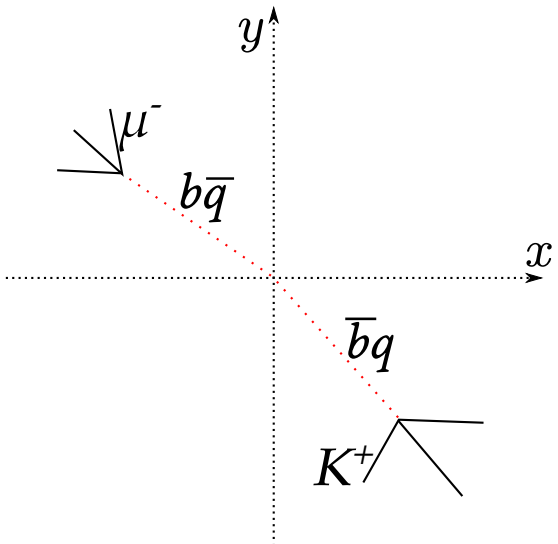


Figure 5.10 *The topology of a combinatoric candidate looking down the z axis, beam line axis, with kaon and muon originating from the decay of different B mesons. The two B mesons are produced back to back in the transverse plane. The kaon and muon are visualised in opposite quadrants of the xy plane.*

Vetoes are applied to the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ candidates rejecting events compatible with the decays $B \to (D^* \to (D^0 \to K^+K^-)\pi)\mu\nu X$ and $B_s^0 \to D_s^-\pi^+$. For the former decay the mass difference between the D^0 and D^* is only slightly higher than the mass of a pion. A selection is placed on the mass difference, $m_{KK\pi} - m_{KK} > 138$ MeV, which efficiently rejects all $B^0 \to D^*\mu^+\nu_\mu$ decays. The latter decay is rejected by reconstructing the muon under the mass hypothesis of a pion and rejecting $D_s^-\mu^+$ candidates with an invariant mass consistent with $|m_{B_s^0} - m_{D_s^-\pi^+}| < 70$ MeV.

5.6.4 *sPlot* Unfolding

The *sPlot* technique is a statistical tool used to unfold data distributions consisting of several sources merged into a single sample [85]. The most frequent use case containing two data sources classified as data and background, typically there is a region in the data containing pure background, e.g. the sidebands of a mass peak in an invariant mass distribution, and a region containing an inseparable mix of signal and background, e.g. the region containing the signal peak. The *sPlot* procedure unfolds the contributions of different sources of a data

sample in the context of a likelihood fit to a data distribution, e.g the invariant mass distribution, obtaining the yields of the signal and background components. The *sPlot* procedure asignes each event a weight calculated from the likelihood obtained from the fit. When assigning *signal* weights, events at the centre of the mass peak will have a large positive weight and events in the sidebands will have a large negative weight as plotted in Figure 5.11. In the case of an invariant mass peak the *sPlot* method of subtracting backgrounds specialises to a sideband subtraction. Two sets of weights are provided by the *sPlot* method allowing the full dataset to be viewed as either signal or background. When weighting as signal, the sum of weights will equal the yield of the signal sample. This is also true when considering a subset of the events, such as those contained in the bin of a histogram. By drawing histograms and weighting events by the *sPlot* signal weights the unfolded signal distribution is obtained. It is essential that the variables drawn using the *sPlot* weights are uncorrelated with the variable used to obtain the *sPlot* weights.

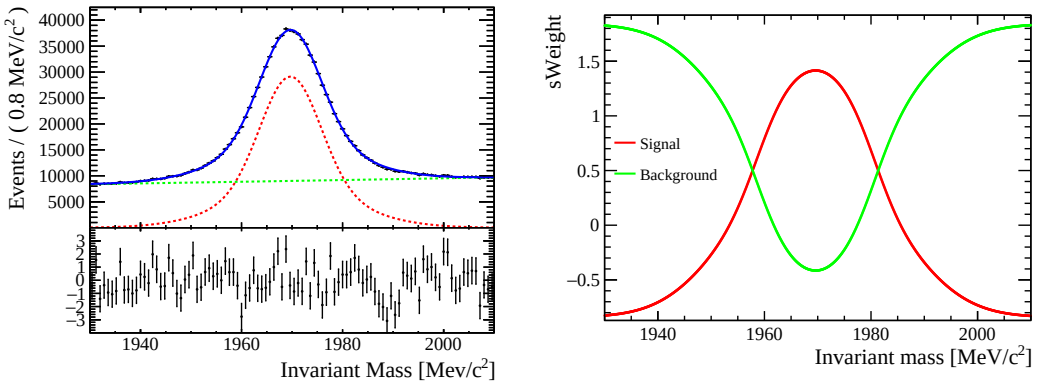


Figure 5.11 *Fit results of a maximum likelihood to the* $D_s^- \to K^-K^+\pi^-$ *invariant mass distribution and* sPlot *weights.*

The *sPlot* method is used in this analysis to obtain the true signal distributions of $B^+ \to J/\psi K^+$ and $D_s^- \to K^-K^+\pi^-$ candidates. The true distributions of signal decays are necessary in order to compare signal Monte Carlo decays with candidates in data. In all cases an extended maximum likelihood fit is performed to the invariant mass spectrum of the final state particles. The signal decay is modelled using a double Gaussian shape and the combinatoric background is modelled using and exponential and all parameters are left free. The invariant mass distribution of $D_s^- \to K^-K^+\pi^-$ candidates with the fit results overlaid is plotted in Figure 5.11 (left) and the sWeights calculated from the fit results are plotted on the right. The results of all fits from which the *sPlot* method is used are given in Appendix C.

5.6.5 Boosted Decision Tree

A binary decision tree is a decision support tool designed to efficiently separate data into categories [86–88]. At each node in the tree a selection, c_N is made on one of the N variables, $\vec{x} = \{x_1, x_2, x_3, ..., x_N\}$, separating the data into two subsets. The selections are chosen such that the Gini index, $p(1-p)$, is minimised, where p is the signal purity. For a binary decision tree categorising data into signal and background categories, the classifier output, $h(\vec{x})$, is calculated from the final subsets, or leaves. Leaves classified as signal have an output, $h(\vec{x}) = 1$, and leaves classified as background have an output, $h(\vec{x}) = -1$. Decision trees may be fully trained, with leaves containing a fully pure sample of signal or background events, or they may be partially trained with the leaves containing a mix of signal and background events as visualised in Figure 5.12. The use of a fully trained tree may introduce over training effects while the use of partial trees may introduce biases. When training a decision tree it is essential that bias and variance are minimised. A biased decision tree is typically under trained and the response variable will have systematic shifts away from the optimal value whereas an over trained decision tree should be unbiased but the response variable will have a large resolution.

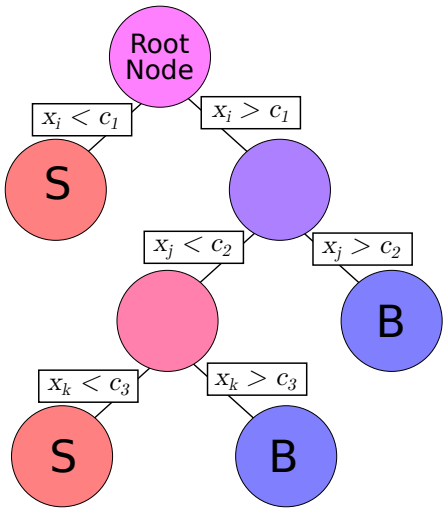


Figure 5.12 *A schematic of a small decision tree applying sequential selections to data maximising the signal to background separation. The final nodes are classified as signal and background.*

A boosted decision tree (BDT) uses many partially trained trees; the partially trained trees are weak learners and the BDT gains its power from the combination of many weak learners. A *typical* BDT would contain several hundred weak decision trained with a maximum depth of three to five. A node is no longer

divided when it contains a critical number of events. Each tree is is given a boost weight, α_i after training calculated from the fraction of misclassified events, ρ_i,

$$\alpha_i = \frac{1-\rho_i}{\rho_i}. \tag{5.10}$$

After a tree is trained the misclassified events are weighted by α_i. The re-weighted and renormalised data is used to train the next tree. The final output of the BDT classifier is

$$h_{\mathrm{BDT}}(\vec{x}) = \frac{1}{N_{\mathrm{Tree}}} \sum_{i}^{N_{\mathrm{Tree}}} \ln(\alpha_i) h_i(\vec{x}). \tag{5.11}$$

As with all classifiers, BDTs can be susceptible to over training. Over training occurs when the decision making algorithm makes decisions due to statistical fluctuations rather than true differences in the data. An over trained classifier will quote a greater separating power than is truly achieved. Over training is remedied using k-fold cross validation [89] which is a technique used to quantify the amount of over training in the classifier and test the model's ability to predict new data. The training data is divided into k sub samples and the classifier is trained k times using $k-1$ sub samples. The classifier is tested with the sub sample independent of the training samples.

Due to inaccuracies in simulation or a lack of knowledge on the underlying physics, there may be fundamental differences between the data used in training and classification may introduce biases into the BDT. This is of particular importance for particle physics analyses where the signal sample used in training is usually simulated Monte Carlo and the background is taken from data in a region with no signal. Fundamental differences between Monte Carlo and data arise due to the simulation algorithms mismodelling the data; the classifier may detect these differences and falsely classify events due to these differences. Correcting Monte Carlo by reweighting the mismodelled variables can go a long way towards solving biases due to mismodelling.

Boosted decision trees are frequently used in this analysis to separate signal candidates from background candidates and to provide a method of quantifying and correcting for differences between ideally identical datasets.

5.6.6 Kinematic Corrections

The simulation of Monte Carlo data is not perfect and the distributions of several variables show disagreements between Monte Carlo and data. It is essential that differences between Monte Carlo and data are corrected to ensure that the distributions of variables used for fitting are correct, and that the determination of efficiencies are accurate. A simple approach to reweighting is to plot the histogram of a variable for Monte Carlo and data, and assign each Monte Carlo event a weight corresponding to the ratio of data and Monte Carlo yields at that point. The simple method fails for multidimensional reweighting as a multidimensional histogram with granular bins will face problems due to low bin statistics and a histogram with coarse bins will produce biases due to variations within the bins. For this analysis a novel approach is taken; a boosted decision tree (BDT) is trained to detect differences between pure data and Monte Carlo. If the simulation is perfectly modelled, the BDT will return an output variable with no separating power[2]. If, however the simulation does not perfectly model the data the BDT will return an output variable with a significant separating power. The driving assumption behind this method for correcting the simulation is that the output variable of the BDT will completely encapsulate all Monte Carlo/data differences in a single discriminating variable. By performing the simple one dimensional correction on the BDT output it is possible to correct all the variables used in training the BDT simultaneously [90].

In order to perform an effective comparison between simulated Monte Carlo and data a pure, signal only data is sample is needed. The decay $B^+ \to J/\psi K^+$ is partially reconstructed using the $K^-\mu^+$ pair and is used to correct $B_s^0 \to K^-\mu^+\nu_\mu$, also reconstructed using the $K^-\mu^+$ pair. An *sPlot* background subtraction is performed on the $B^+ \to J/\psi K^+$ invariant mass peak by combining the $K^-\mu^+$ pair with a muon found using the isolation BDT detailed in Section 5.6.7 allowing the signal distributions to be obtained. When correcting the $B_s^0 \to D_s^-\mu^+\nu_\mu$ simulation, simulated cocktail $B_s^0 \to D_s^-\mu^+\nu_\mu X$ data is compared to data containing a well reconstructed D_s^+ in association with a muon. Backgrounds are reduced by applying a selection on the output of the isolation BDT and an *sPlot* background subtraction is performed on the D_s^+ mass peak. For both the signal and normalisation modes a BDT containing 200 trees with a maximum depth of 3 and a minimum leaf size of 6% for $B_s^0 \to D_s^-\mu^+\nu_\mu$ and 4% for $B_s^0 \to K^-\mu^+\nu_\mu$ is trained to separate simulated Monte Carlo and data. For

[2]Over training effects and statistical fluctuations will return some *false* separating power.

$B_s^0 \to K^-\mu^+\nu_\mu$	$B_s^0 \to D_s^-\mu^+\nu_\mu$
Track Multiplicity	Track Multiplicity
η B_s^0	η B_s^0
p_{T} B_s^0	p_{T} B_s^0
p_{T} K^-	p_{T} D_s^-
p_{T} μ^+	p_{T} μ^+

Table 5.7 *Monte Carlo distributions corrected in Monte Carlo using a BDT reweighting.*

the training and evaluation a $k = 2$ $k-$folding is used. The BDT is trained to separate, and hence correct the variables listed in Table 5.7. Figure 5.13 plots the BDT response obtained (left) when training to separate Monte Carlo from data and the correction weights (right) which are applied to Monte Carlo. Kinematic distributions of $B^+ \to J/\psi\, K^+$ Monte Carlo before and after correction are plotted alongside background subtracted data in Figure 5.14 demonstrating the effectiveness of this method. Additional validation plots are provided in Appendix D.

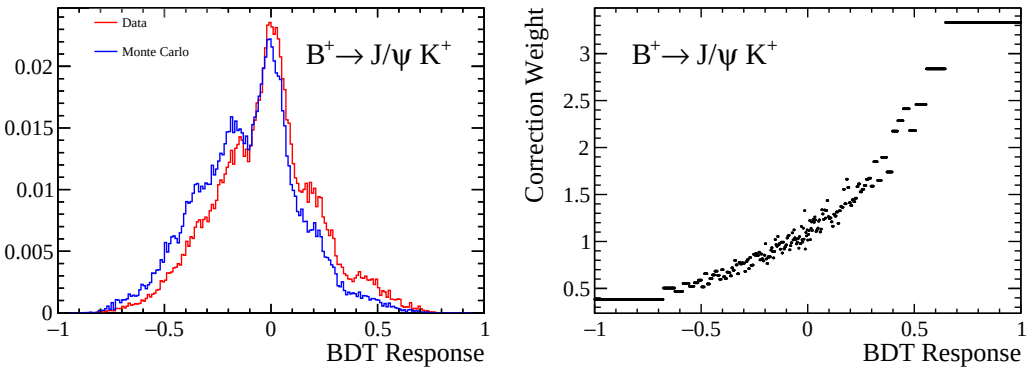


Figure 5.13 *The BDT response used to separate Monte Carlo and data (left) and the weights used to correct the simulation (right). The $J/\psi\, K^+$ is reconstructed using only the $K^+\mu^-$ pair.*

5.6.7 Charged Track Isolation BDT

Charged track isolation variables have been used by analyses at LHCb since the very beginning to identify backgrounds containing additional charged tracks [91, 92]. The purpose of isolation algorithms is to identify events containing partially reconstructed decays. Consider the topology of the two decays in Figure 5.15, on the left is a signal $B_s^0 \to K^-\mu^+\nu_\mu$ decay produced via $b\bar{b}$ production with the second quark decaying to produce an additional muon and the figure on the right

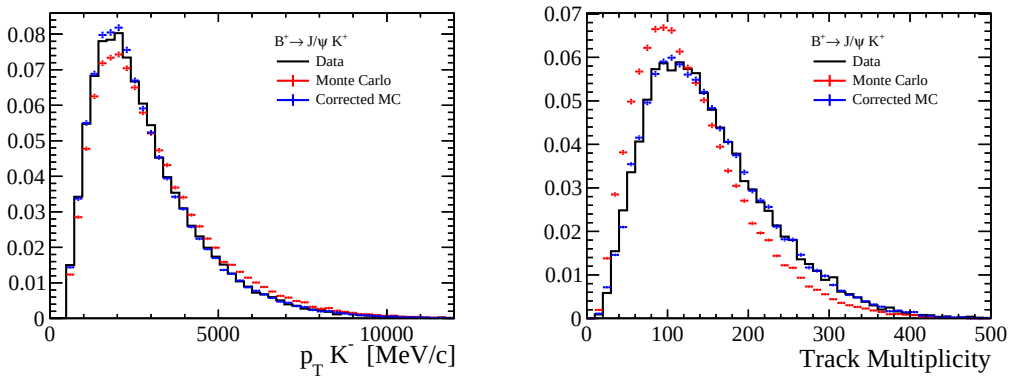


Figure 5.14 *Kinematic distributions of $B^+ \to J/\psi K^+$ reconstructed using the $K^+\mu^-$ pair are plotted for data and simulation before and after correction.*

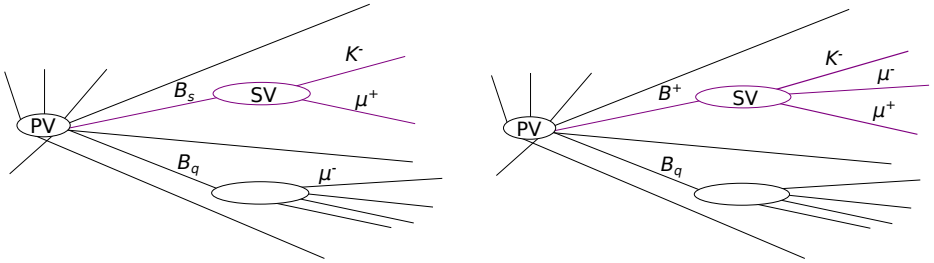


Figure 5.15 *The topology of $B_s^0 \to K^-\mu^+\nu_\mu$ (left) and $B^+ \to J/\psi K^+$ (right). The isolation BDT is trained to reject events containing tracks compatible with the $B_s^0 \to K^-\mu^+\nu_\mu$ candidate decay.*

containing a background $B^+ \to J/\psi K^+$ decay. On the left plot the negative muon is well isolated from the candidate B_s^0 while on the right the negative muon is coupled with the candidate B_s^0. The isolation tools are used to reject events similar to those on the right of the figure.

The first iteration of isolation algorithms use cone isolation, a cone in ΔR,

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}, \tag{5.12}$$

where ϕ is the azimuthal angle and η is the pseudorapidity is drawn around the candidate track. $\Delta\eta$ and ΔR are both Lorentz invariant quantities. The isolation algorithm investigates the tracks and energy deposits in the calorimeters within this cone and returns a series of variables detailing the activity around the track. For a well isolated track one would expect very little detector activity within the cone around the candidate track. When investigating the contents of the cone one may choose to consider activity from neutral and/or charged particles depending on the expected backgrounds. Several variables may be defined by considering the kinematics of the cone and the candidate track. The momentum of a cone, $p_{\mathrm{T}}(Cone)$ is defined as the vector sum of the transverse momentum, p_{T}, of tracks within the cone,

$$p_{\mathrm{T}}(Cone) = \sum_i \vec{p_{\mathrm{T}}} \tag{5.13}$$

The momentum asymmetry between the cone and the candidate track, CT, is defined

$$A_{p_{\mathrm{T}}} = \frac{|\vec{p_{\mathrm{T}}}(CT) - \vec{p_{\mathrm{T}}}(Cone)|}{|\vec{p_{\mathrm{T}}}(CT) + \vec{p_{\mathrm{T}}}(Cone)|}, \tag{5.14}$$

and the transverse isolation of the cone is defined as

$$TI = \frac{p_{\mathrm{T}}(CT)}{p_{\mathrm{T}}(CT + Cone)}. \tag{5.15}$$

In addition one may attempt to reconstruct partially reconstructed decays by combining the candidate track with the highest p_{T} charged track within the cone. Decays containing additional neutral particles may be reconstructed by searching for a cluster of hits in the electromagnetic calorimeter corresponding to a neutral particle and combining the candidate track with the reconstructed photon. Two photon clusters may be combined allowing the π^0 to be reconstructed and combined with the candidate track, which is especially useful when searching for the partially reconstructed decay $B_s^0 \to (K^{*-} \to K^-\pi^0)\mu^+\nu_\mu$.

A more sophisticated approach to identifying partially reconstructed backgrounds is to search through all reconstructed tracks in the event to determine if it is compatible with the B_s^0 candidate vertex by investigating the topologies of the candidate tracks and the remaining tracks within the event. A true signal candidate should have no additional tracks compatible with the B_s^0 decay vertex. Before continuing the terminology must be defined:

Candidate track:	A reconstructed track originating from the decay of a B_s^0.
Underlying track:	All tracks in the event which are not candidate tracks
Isolated track:	A track which is truly independent of the B_s^0 decay, e.g. a prompt track originating from the primary vertex.
Non isolated track:	A track which truly originates from the candidate decay but is not reconstructed as such, e.g. the additional muon in $B^+ \to J/\psi K^+$.
Least isolated track:	The underlying track with the highest probability of originating from the candidate decay.

A boosted decision tree developed for a different analysis [93] takes as an input the kinematics and topologies of a candidate and underlying track and returns an output correlated to the likelihood that the two tracks originate from the same vertex. The BDT is trained using tracks from Monte Carlo decays of $B^+ \to D^{*-}\pi^+\mu^+\nu_\mu$ reconstructed as $B^0 \to D^{*-}\mu^+\nu_\mu$ resulting in the underlying event containing an additional pion coupled to the reconstructed candidate. During training of the BDT the background sample is composed of the combinations of the B^0 candidate tracks and the additional pion, and the signal sample is composed of the combinations of candidate tracks with another non isolated track. As there are approximately 100 additional tracks in each event the signal sample which combines the candidate decay with another track is approximately 100 times larger than the background sample. The BDT is trained to separate signal and background samples using the following variables:

- $\chi^2$of the minimum impact parameter of the track with respect to any PV
- Distance between Vertex(candidate track, track) and the SV
- Distance between Vertex(candidate track, track) and the PV
- Distance of closest approach between the track and PV

- Difference in ϕ between track and reconstructed D^0
- The angle between the sum of momenta $p_{\text{Candidate track}} + p_{\text{track}}$ and the line from PV to SV

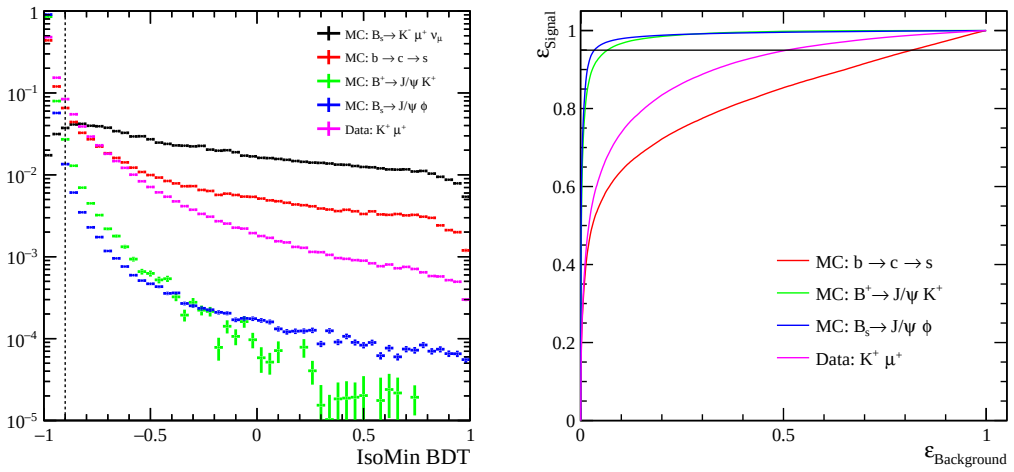


Figure 5.16 *The per event minimum output of the isolation BDT (left) and the corresponding receiver operating characteristic, ROC, curves. A selection of* IsoMinBDT > -0.9 *is applied. The solid horizontal line on the ROC curve indicates the signal efficiency.*

When processing an event, every combination of candidate and underlying track is processed with the BDT producing ≈ 300 variables per event. Two combined variables are created for each candidate track. `IsoMinBDT`, plotted in Figure 5.16 (left), is the BDT output for the underlying track most likely to originate from the same vertex as the candidate, and `IsoSumBDT` is the average of the BDT outputs for all underlying tracks when compared to the candidate track. In addition, the kinematics of the least isolated track are saved. The invariant mass spectrum of the combination of the candidate and least isolated underling track in data is plotted in Figure 5.17 for $K^-\mu^+$ (left) and $K^+\mu^+$ data (right). Mass peaks corresponding to the ϕ, D^0, J/ψ, $\psi(2S)$ and $f_0(500)$ are clearly visible and represent partially reconstructed backgrounds. These peaks may be explicitly rejected by making a selection on the invariant mass or a selection on the BDT may be used to reject almost all partially reconstructed decays. The isolation BDT also provides exceptional rejection power for combinatoric decays.

A loose selection, `min(kaon_m_IsoMinBDT,muon_p_IsoMinBDT) > -0.9`, is applied to the isolation BDT.

For this analysis variables from both the cone isolation and BDT isolation tools

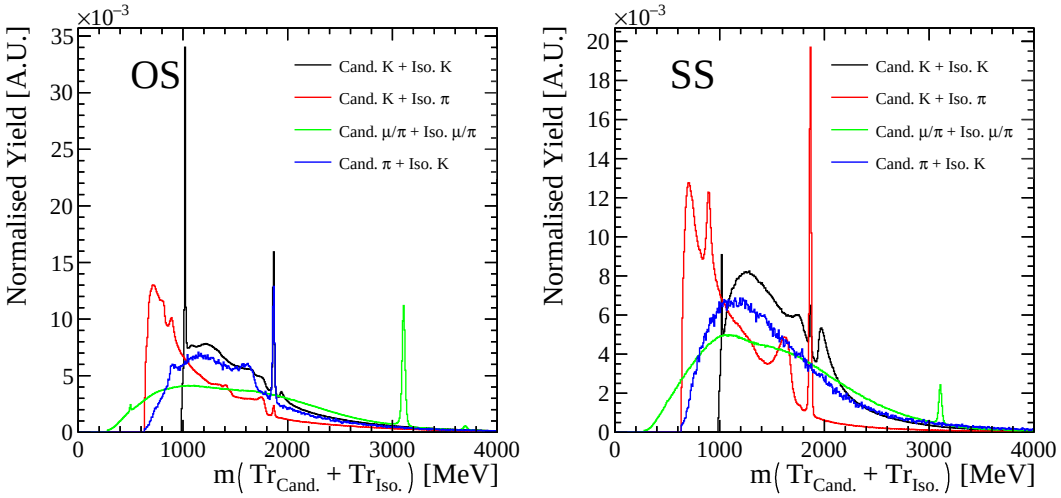


Figure 5.17 *The invariant mass distribution of a candidate track combined with the least isolated track under various mass hypotheses. Data is reconstructed as opposite sign (OS) $K^-\mu^+$ (left) and same sign (SS) $K^-\mu^-$ (right).*

are used in rejecting and reconstructing partially reconstructed backgrounds. The cone isolation is used for both charged and neutral backgrounds and the BDT isolation is used for charged backgrounds.

5.6.8 Selection BDT

Two BDTs are trained to separate signal from background for this analysis. The first BDT is trained to separate signal events from partially reconstructed backgrounds modelled using Monte Carlo and the second is trained to remove backgrounds found in the same sign data sample. Both BDTs use Monte Carlo $B_s^0 \to K^-\mu^+\nu_\mu$ decays as the signal sample during training. The background sample used when training the first BDT is a cocktail mix of background Monte Carlo decays reconstructed as $B_s^0 \to K^-\mu^+\nu_\mu$ and the second BDT uses same sign (SS) data reconstructed as $B_s^0 \to K^+\mu^+\nu_\mu$. The strategy is to apply a loose selection to the output of the isolation BDT, `min(kaon_m_IsoMinBDT,muon_p_IsoMinBDT) > -0.9`, train the first BDT to reduce backgrounds from partially reconstructed decays containing additional charged tracks, apply a selection on the output of this BDT and train the second BDT to provide additional discriminating power and reduce additional backgrounds seen in data such as combinatorics and *trickle down* decays of higher excited particle states. The BDTS are applied sequentially, i.e.the first BDT is

trained and a selection applied to all data samples, then the second BDT is trained and a selection applied to all samples. For the isolation BDT the choice of location for the selection is fairly arbitrary as the BDT response variable is used as an input for a later BDT, and the applied selection only removes a small region with very high background purity. The signal efficiency for the isolation BDT is 95%.

The first BDT, refereed to as the *charged BDT* is trained to separate signal Monte Carlo from a cocktail of Monte Carlo backgrounds, the events used in the training are detailed in Table 5.10 alongside their separating power $\langle S^2 \rangle$. Separating power between signal, $\hat{y}_S$, and background, $\hat{y}_B$, distributions is defined as,

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy. \tag{5.16}$$

The separation is zero for identical distributions and is one if the distributions do not overlap [94]. The charged BDT uses the variables listed in Table 5.8 to separate signal from background. All Monte Carlo samples used in the training have their kinematics corrected using the reweighting procedure detailed in Section 5.6.6. During the training 850 trees with a maximum depth of 3 levels and a minimum node size of 2.5% are trained using the AdaBoost boosting method [95] and the variables listed in Table 5.8. The effects of over training are removed through the use of 2 factor k-folding with the data divided by magnet polarity for training and testing.

The second BDT is referred to as the same sign (SS) BDT since it is trained with $K^+\mu^+$ data candidates as the background sample. The variables used in training the SS BDT are listed in Table 5.9. The SS BDT follows the same training procedure as the charged BDT, and a selection is placed on the output of the charged BDT before training minimising correlations between the two BDTs.

Optimising the point at which a selection is made on the two BDTs is non trivial. The BDTs are incredibly effective at removing backgrounds and producing a data sample with an impressive signal peak, however due to the limited size of the available Monte Carlo samples modelling the backgrounds, such a selection would result in Monte Carlo samples containing almost no events making a fit to the data impossible. Thus a compromise must be reached whereby backgrounds are reduced to such an extent that the signal peak is well identifiable but the background Monte Carlo samples have a large enough yield such that a Fit to the data will return meaningful results. Consequently no quantitative optimisation is

Variable	Separation
Minimum isolation BDT response	1.90×10^{-1}
Invariant Mass, K^- +$\mathrm{T}_{\text{Least Iso.}}$	1.00×10^{-1}
Isolation BDT summed over all underlying tracks	9.39×10^{-2}
B_s^0 Transverse Isolation	3.54×10^{-2}
Transverse isolation between K^- and cone	3.00×10^{-2}
Kaon transverse momentum	2.84×10^{-2}
Transverse isolation between K^- and charged cone	2.58×10^{-2}
$p_{\mathrm{T}}(B_s^0) - 1.5 \times p_{\mathrm{T}}(\mu^+)$	2.17×10^{-2}
$p_{\mathrm{T}}(B_s^0)$	1.94×10^{-2}
$\Delta\eta$ between K^- and cone	1.89×10^{-2}
Momentum asymmetry between μ^+ and cone	1.85×10^{-2}
K^- Isolation BDT response - μ^+ Isolation BDT response	1.63×10^{-2}
B_s^0 Decay vertex fit χ^2	1.11×10^{-2}
Invariant Mass, μ^+ +$\mathrm{T}_{\text{Least Iso.}}$	7.67×10^{-3}
B_s^0 helicity angle	2.03×10^{-3}

Table 5.8 *The input variables used as inputs for the BDT trained to reject charged backgrounds are listed with their separating power. Several variables use information obtained from a cone draw around candidate tracks with $\Delta R = 0.5$. The least isolated track is referred to as* $\mathrm{T}_{\text{Least Iso.}}$.

Variable	Separation
$p_{\mathrm{T}}(K^-)$	2.76×10^{-02}
$\mathrm{DIRA}_{B_s^0}$	1.98×10^{-02}
Momentum asymmetry between K^- and neutral cone	1.36×10^{-02}
Transverse isolation between K^- and neutral cone	1.30×10^{-02}
Invariant mass K^- and additional π^0	1.02×10^{-02}
$p_{\mathrm{T}}(B_s^0) - 1.5 \times p_{\mathrm{T}}(\mu^+)$	9.31×10^{-03}
B_s^0 Flight distance significance	7.56×10^{-03}
B_s^0 transverse momentum	6.20×10^{-03}
B_s^0 helicity angle	5.96×10^{-03}
B_s^0 Decay vertex fit χ^2	3.68×10^{-04}

Table 5.9 *The input variables used as inputs for the BDT trained to reject backgrounds found within the same sign sample. Several variables use information obtained from a cone draw around candidate tracks with $\Delta R = 0.5$.*

Sample	Entries/1000
$B_s^0 \to K^- \mu^+ \nu_\mu$	165
Background Samples	
$b \to (c \to K\mu X)X$	47
$b \to K^\pm \mu^\pm X$	190
$b \to K\mu X$	550
$B^0 \to J/\psi\, K^{*0}$	226
$B^0 \to (D^* \to K\pi\pi\pi X)\mu\nu$	50
$B^0 \to (D^* \to K\pi X)\mu\nu$	1210
$B^0 \to (D^* \to K\pi)\mu\nu$	34
$B^0 \to (D \to K\pi\pi)\mu\nu$	75
$B^+ \to J/\psi\, K^+$	445
$B^+ \to J/\psi\, K^{*+}$	135
$B^+ \to D^0 \mu^+ \nu_\mu X$	230
$B_s^0 \to J/\psi\, \phi$	2500
$B_s^0 \to D_s^- \mu^+ \nu_\mu X$	96

Table 5.10 *The simulated decays and number of events used during the training of the BDTs.*

performed on the BDT as there is no trivial choice of parameter t optimize. The BDT selections used in this analysis were found using a more qualitative method. The two BDTs were iteratively tightened, by incrementally tightening the BDT that would give the greatest increase of the signal to noise ratio, S_{NR},

$$S_{NR} = \frac{S}{\sqrt{S+B}}, \tag{5.17}$$

where S is the signal yield, and B is the yield of same-sign data, a sample consisting purely of background events. At several points during the tightening process, histograms of the B_s^0 corrected mass were drawn for the Monte Carlo and data samples. A final selection was chosen by analysing the histograms *by eye* to find a compromise between clarity of signal peak and background Monte Carlo statistics.

The response of both BDTs with the selection are plotted in Figures 5.18 and 5.19 with the corresponding Receiver operating characteristic (ROC) curves plotting the signal efficiency against the background efficiency. The ROC curve plots the signal efficiency against background efficiency providing a visual method of quantifying the performance of a selection method. ROC curves with larger areas under the curve have improved background rejection, and the integral of the ROC curve is often used as a performance metric. A selection of `BDT_Charged > 0.10` is placed on the charged BDT and `BDT_SS > 0.05` on the same sign BDT.

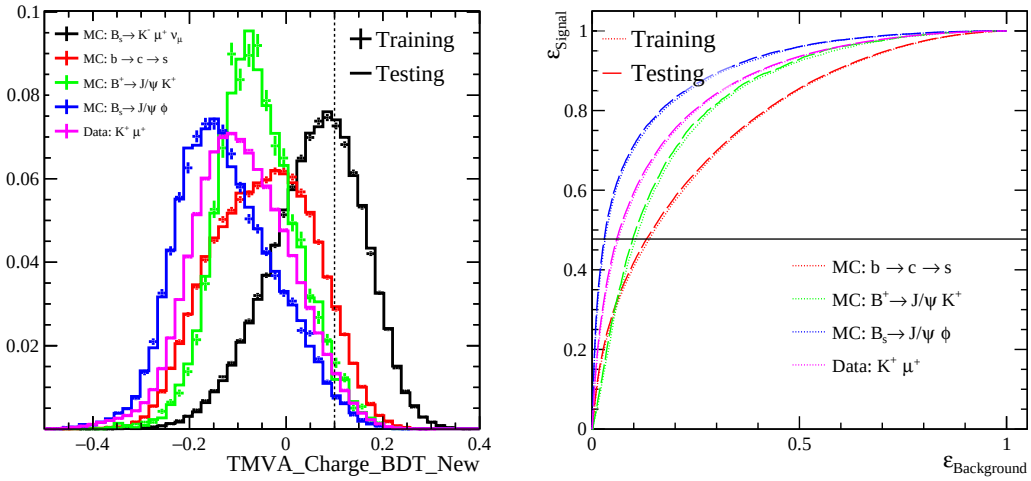


Figure 5.18 *The response for the charged BDT (left) is plotted with the corresponding ROC curves (right). The data used for training is superimposed with the data used for testing. Events to the left of the dashed vertical line are rejected and the solid horizontal line on the ROC curve indicates the signal efficiency.*

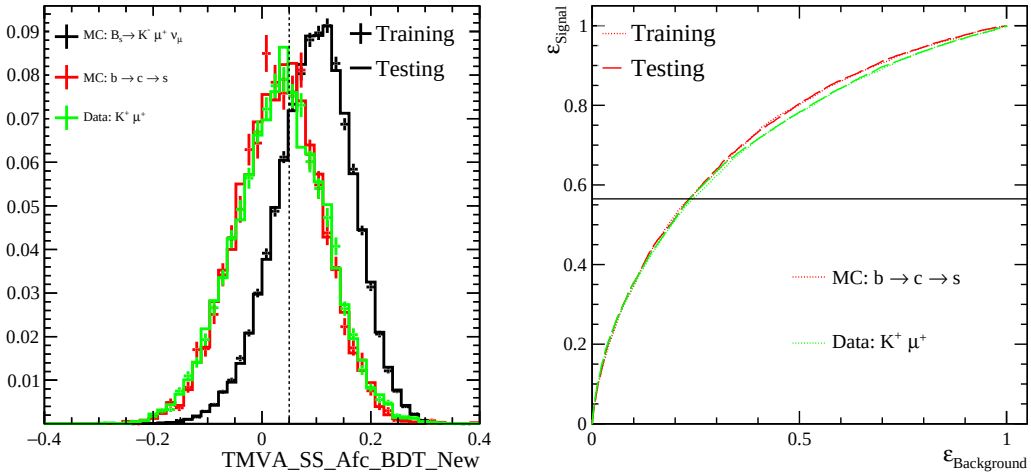


Figure 5.19 *The response for the same sign BDT (left) is plotted with the corresponding ROC curves (right). The data used for training is superimposed with the data used for testing. Events to the left of the dashed vertical line are rejected and the solid horizontal line on the ROC curve indicates the signal efficiency.*

The BDTs are validated by comparing the BDT response in *sPlot* unfolded data with the BDT response in Monte Carlo for $B^+ \to J/\psi K^+$ decays using a fully reconstructed $K^+\mu^-\mu^+$ triad or the $K^+\mu^-$ pair. The validation using $B^+ \to J/\psi K^-$ is plotted in Figure 5.20 and a slight discrepancy is seen in the BDT response between Monte Carlo simulation and data. A correction factor and systematic uncertainty are applied to the calculated BDT efficiency using the $B^+ \to J/\psi K^+$ decay, the details of which are in Section 6.5.4.

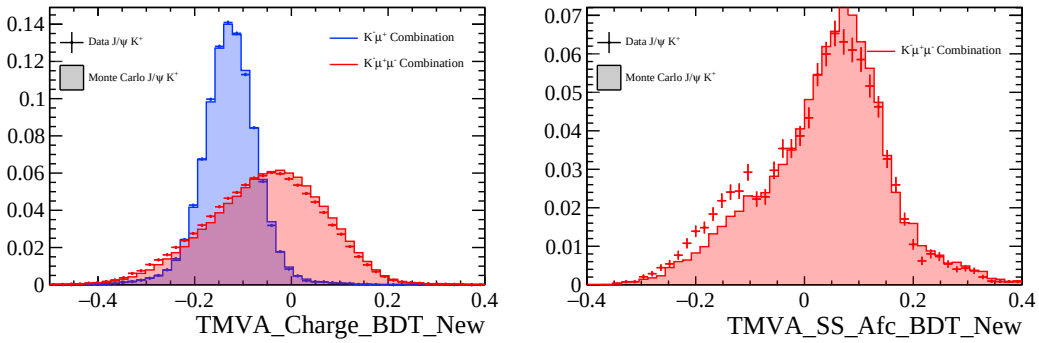


Figure 5.20 *The response of both selection BDTs is plotted for $B^+ \to J/\psi K^+$ candidates reconstructed as $B^+ \to J/\psi K^+$ (red) and $B_s^0 \to K^-\mu^+\nu_\mu$ (blue) for Monte Carlo (line) and background subtracted data (points).*

5.7 Selection on Data

The distributions of $B_s^0 \to K^-\mu^+\nu_\mu$ candidates are plotted in Figure 5.21 as successive seelctions are applied. Events reconstructed using a same sign kaon and muon (SS) combination completely overshadow the opposite sign (OS) combinations. This is due in part to the massive source of same sign kaon and muon pairs from the decay $B^+ \to (D^0 \to K^+\pi^-)\mu^+\nu_\mu$ and friends. As successive selections from vetoes, BDTs and the corrected mass uncertainty are applied a significant structure appears in the $K^-\mu^+$ distribution at the mass of the B_s^0. This structure is the decay $B_s^0 \to K^-\mu^+\nu_\mu$.

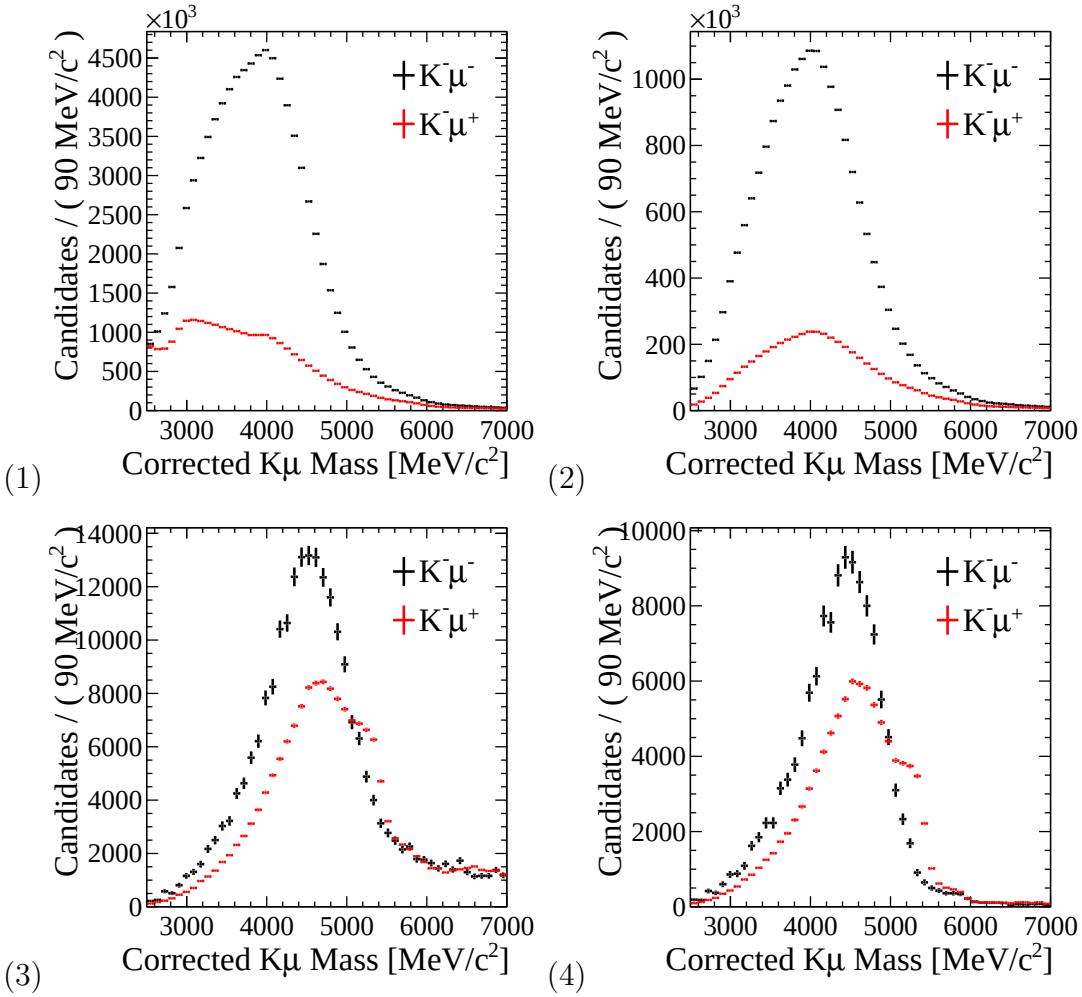


Figure 5.21 *Corrected mass distributions for kaon and muon candidates using same sign (SS) and opposite sign (OS) combinations with successive selections applied. (1) The stripping preselection. (2) Vetoes. (3) BDTs. (4) Corrected mass error Cuts. After the full selection is applied a clear structure containing signal events is visible at the* B_s^0 *mass.*

Chapter 6

Determining $|V_{ub}|/|V_{cb}|$ and $\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)$ at LHCb

This chapter presents the analysis performed to measure the ratio of branching fractions of $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$, and the ratio of CKM matrix elements $|V_{ub}|/|V_{cb}|$. A template fit is performed on the $D_s^-\mu^+$ invariant mass to determine the $B_s^0 \to D_s^-\mu^+\nu_\mu$ yield, and two fits are performed on the corrected $K^-\mu^+$ mass to determine the $B_s^0 \to K^-\mu^+\nu_\mu$ yield in order to measure the ratio of branching fractions, $\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)/\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)$, and the ratio of CKM matrix elements, $|V_{ub}|/|V_{cb}|$. The yields from the fits are combined with efficiencies determined from Monte Carlo and data driven methods and the systematic uncertainties on the final values are determined.

The Beeston Barlow method for fitting with finite statistics is discussed in Section 6.1, and the fits to determine the $B_s^0 \to D_s^-\mu^+\nu_\mu$ and $B_s^0 \to K^-\mu^+\nu_\mu$ yields are discussed in Sections 6.2 and 6.3 respectively. Systematic uncertainties are discussed in Section 6.4 and efficiency calculations are discussed in Section 6.5. The final determinations of $\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)/\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)$ and $|V_{ub}|/|V_{cb}|$ are discussed in Section 6.6.

6.1 Fit Method

6.1.1 Beeston Barlow Fit Method

The Beeston-Barlow method for fitting using finite Monte Carlo samples [96, 97] is a binned template fit method used to extract the yields of different components from a data sample. Instead of using an analytical form for the shapes of the contributions a discrete histogram is used, dividing the distribution into n bins.

The total number of events in data, N_D, and total number of events in the j^{th} Monte Carlo template are found by summing over the n bins of the template histogram,

$$N_D = \sum_{i=1}^{n} d_i, \qquad N_{MC_j} = \sum_{i=1}^{n} a_{ji}, \tag{6.1}$$

where d_i and a_{ji} are the number of data and Monte Carlo events in bin i respectively. Given m fit components with fractional proportions, f_j, The predicted number of events in the i^{th} bin of the data template, $n_i(f_1, f_2, ..., f_m)$, is

$$n_i = N_D \sum_{j=1}^{m} f_j a_{ji}/N_{MC_j}, \tag{6.2}$$

where N_D is the total data yield, and a_{ji} the number of Monte Carlo events from source j in bin i. The fractional proportions must sum to unity,

$$\sum_{j=1}^{m} f_j = 1. \tag{6.3}$$

The proportions of each component, $p_j = N_D f_j/N_{MC_j}$, can be used, allowing Equation 6.2 to be rewritten,

$$n_i = \sum_{j=1}^{m} p_j a_{ji}, \tag{6.4}$$

where the sum of proportions need not equal unity. The proportions scale the Monte Carlo template to its size in data.

The fractional proportions of each component, f_j, sum to unity and can be estimated by minimising

$$\chi^2 = \sum_i \frac{(d_i - n_i)^2}{d_i}, \tag{6.5}$$

which assumes d_i follows a Gaussian distribution. Truly d_i follows a Poisson distribution, and with large numbers of events this is not a bad approximation, however there are often many bins with a low number of events making this approximation invalid. One approach would be to use a binned maximum likelihood fit where the probability of observing a particular d_i multiplied over all n bins is

$$\mathcal{L} = \prod_{i=1}^{n} e^{-n_i} \frac{n_i^{d_i}}{d_i!}, \tag{6.6}$$

and the estimates of the fractions, f_j can be found my maximising the likelihood, or for convenience, the logarithm of the likelihood,

$$\ln(\mathcal{L}) = \sum_{i=1}^{n} d_i ln(n_i) - n_i. \tag{6.7}$$

The methods detailed above only consider the statistical uncertainties in the data sample and neglects any variation in the bin contents of the Monte Carlo templates. As a rule of thumb the Monte Carlo samples should contain ≈ 10 times the number of events in the data, however due to the computational cost of generating simulated Monte Carlo events, the space required to store the data, and impracticalities handling massive datasets this rule is rarely observed. An approach is therefore needed which considers the statistical fluctuations in the Monte Carlo datasets.

The uncertainty parameter of the χ^2 formalism shown in Equation 6.5 can be modified to include the Monte Carlo uncertainty,

$$\chi^2 = \sum_i \frac{(d_i - n_i)^2}{d_i + N_D^2 \sum_j a_{ji}/N_{MC_j}^2}, \tag{6.8}$$

however this still uses the Gaussian approximation which is invalid when bins contain a low number of events.

In order to fully consider the statistical uncertainty from both the data and Monte Carlo, Equation 6.4 can be rewritten replacing a_{ji} with Poisson distributions, A_{ji},

$$n_i = \sum_{j=1}^{m} p_j A_{ji}. \tag{6.9}$$

The total likelihood is now the combined probability of the observed d_i and a_{ji},

$$\ln(\mathcal{L}) = \sum_{i=1}^{n} d_i ln(n_i) - n_i + \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji} ln(A_{ji}) - A_{ji}. \tag{6.10}$$

During the construction of the likelihood the Poisson distributions of all template histograms are combined into a single Poisson distribution, thus there is only one Poisson distribution for each bin. For this step to work correctly the initial values of the proportions must be close to the values determined by the fit.

6.2 $B_s^0 \to D_s^- \mu^+ \nu_\mu$ Fit Results

6.2.1 Normalisation Fit Model

A maximum likelihood, Beeston-Barlow binned template fit is performed on the corrected $D_s^-\mu^+$ mass distribution to extract the $B_s^0 \to D_s^-\mu^+\nu_\mu$ signal yield. A bias free background subtraction is performed using the $K^-K^+\pi^-$ invariant mass distribution to remove the $K^-K^+\pi^-$ combinatoric contribution, the plots of which are shown in Figure 6.2. The yield in each bin of the corrected $D_s^-\mu^+$ input histogram is the result of a fit to the $K^-K^+\mu^-$ invariant mass to determine the D_s^+ yield. The fit to the corrected $D_s^-\mu^+$ mass distribution is used to separate the signal $B_s^0 \to D_s^-\mu^+\nu_\mu$ signal component from the background contributions. The backgrounds predominantly originate from semileptonic B_s^0 decays containing higher excited D_s^- resonances. Backgrounds consisting of partially reconstructed $B \to D_s^+ DX$ candidates and tauonic decays are considered, as are candidates containing misidentified muons. Combinatoric combinations of real muons with real D_s^+ mesons are neglected; when investigating the $K^-K^+\pi^-$ invariant mass distribution using the same sign, $D_s^-\mu^-$, sample no D_s^- peak is seen. The same sign sample may be assumed to be purely combinatoric as very few decays contain a same sign D_s^- and muon. The fit components and the sources of templates used in the fit are summarised in Table 6.1.

The templates used in the fit contain 40 bins in corrected mass ranging from 3000 MeV to 6500 MeV with an equal bin width. Sections 6.5.2 onwards discuss the corrections applied to Monte Carlo simulation to correct the selection efficiencies. It is important to note that when producing histograms for the fit

Component	Source
$B_s^0 \to D_s^- \mu^+ \nu_\mu$	Monte Carlo
$B_s^0 \to D_s^{*-} \mu^+ \nu_\mu$, with $D_s^{*-} \to D_s^- \gamma$	Monte Carlo
$B_s^0 \to D_s^{*-} \mu^+ \nu_\mu$, with Higher D_s^+ Excitations	Monte Carlo
$B_s^0 \to D_s^- \tau^+ \nu_\mu X$, with $\tau^+ \to \mu^+ \nu_\mu \overline{\nu}_\tau$	Monte Carlo
$B_q^0 \to D_s^- D_q^{(*)} X$, with $D_q \to \mu^+ \nu_\mu X$	Monte Carlo
Misidentified Muons	Fake Muon Data

Table 6.1 *Fit components for the normalisation fit and sources of data used when drawing templates.*

the entries in each bin are weighted by kinematic, PID and tracking corrections detailed in Sections 5.6.6, 6.5.2 and 6.5.3. Backgrounds originating from similar decays with low yields are combined into a common template; all $B \to D_s^+ D$ backgrounds are combined into a single template and the higher excitations of the D_s^+ above the D_s^{*+} are combined into single template. The combinations are plotted in Figure 6.1. When creating the Monte Carlo templates for the fit all events are weighted by the product of the weights obtained from the BDT kinematic reweighting, PID correction and tracking correction. When performing the fit to obtain the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yield all component yields are left free.

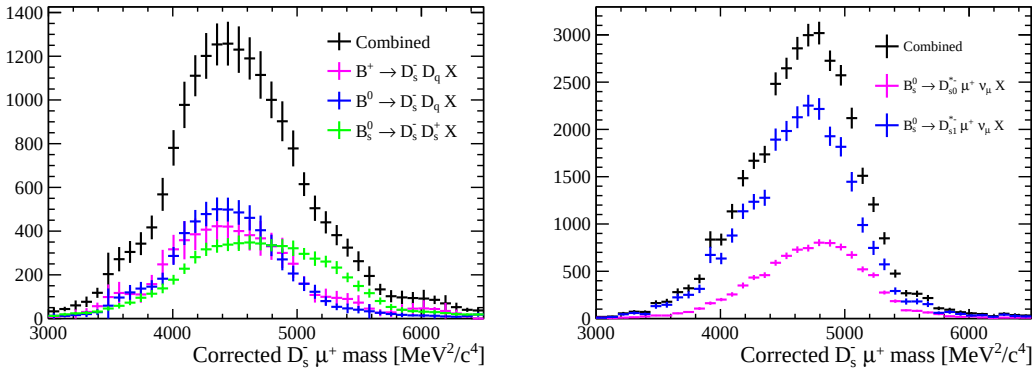


Figure 6.1 *The $B \to DD\mu^+\nu_\mu X$ templates are merged into a single template, and the decays of further excitations of the D_s^+ are combined into a single template.*

6.2.2 Background Subtraction

A signal extraction is performed to remove the $K^+K^-\pi^+$ combinatoric contribution from the data. Each data point in the $D_s^-\mu^+$ corected mass input histogram for candidates in data, plotted in Figure 6.3, is the result of a fit to the $K^+K^-\pi^-$ invariant mass distribution plotted in Figure 6.2 to determine the

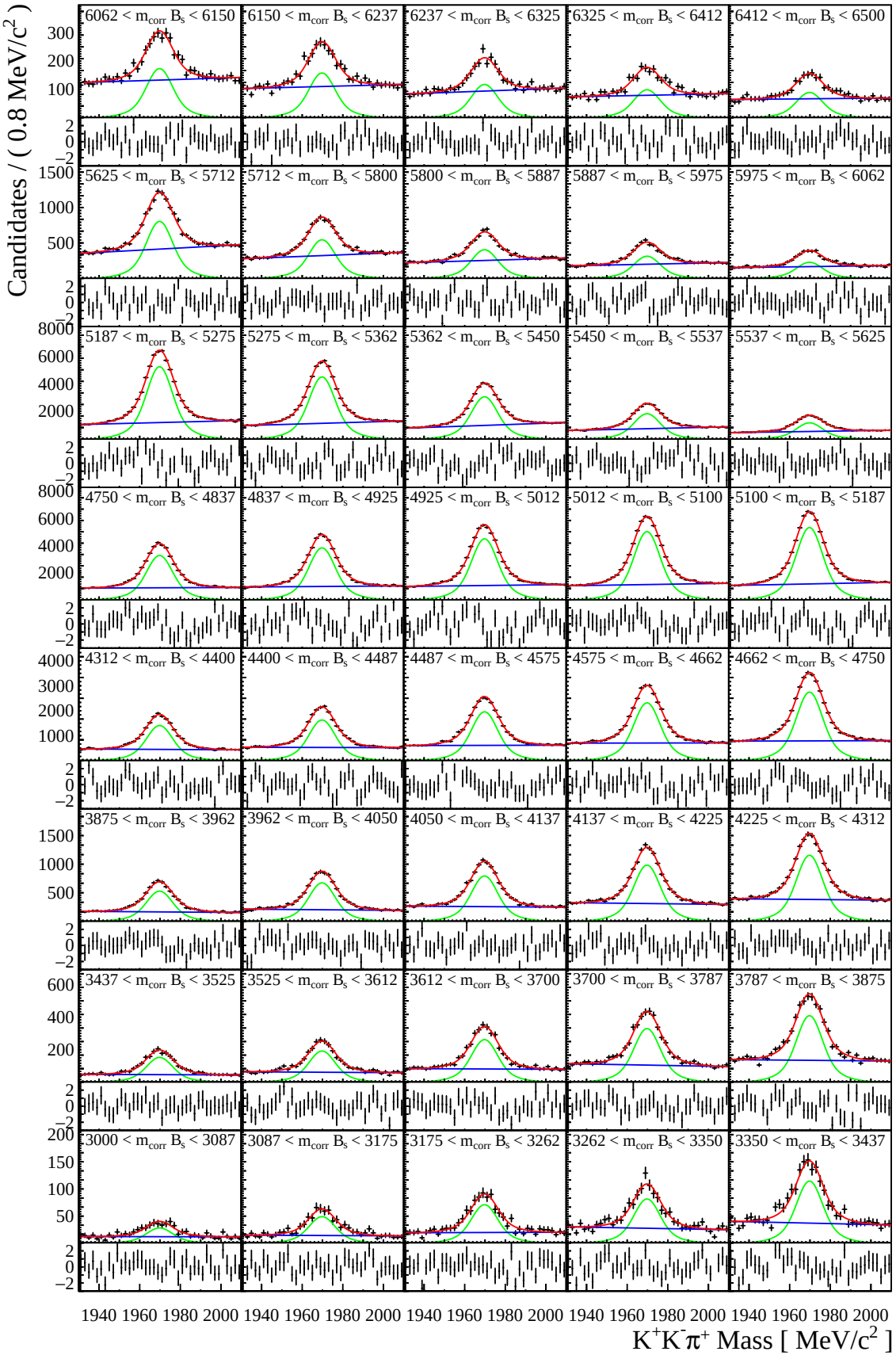


Figure 6.2 *Fits performed as part of a combinatoric background subtraction on the $K^+K^-\pi^+$ invariant mass with pulls underneath.*

D_s^- yield. Correlations between the $K^+K^-\pi^+$ invariant mass and the $D_s^-\mu^+$ corrected mass mean that the *sPlot* method for subtracting backgrounds cannot be used. Instead a *divide and fit* method is used whereby the data is divided into n smaller subsets, each corresponding to a specific bin in the $D_s^-\mu^+$ corrected mass spectrum. A binned maximum likelihood fit is performed to the $K^+K^-\pi^-$ invariant mass distribution for each dataset. A double-Gaussian models the D_s^- shape and an exponential models the combinatoric background shape. The $D_s^+\mu^+$ yield in the corrected mass histogram for each bin is set as the signal yield from the fit. The fits from the *divide and fit* method are plotted in Figure 6.2. No background subtraction is required for the Monte Carlo samples.

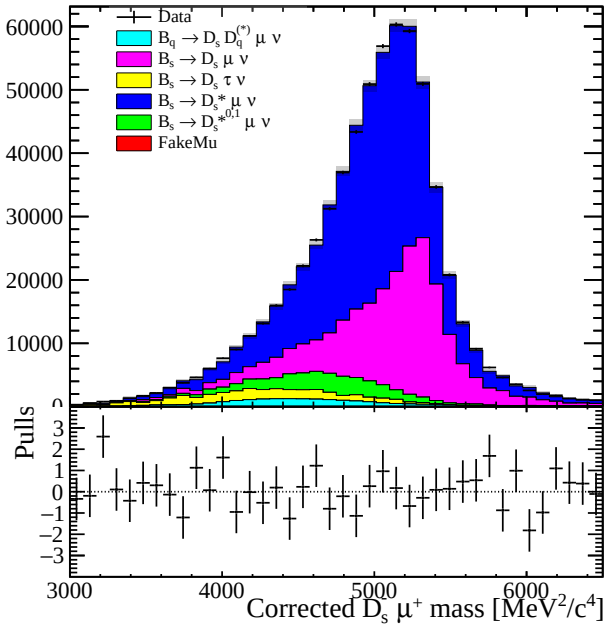


Figure 6.3 *A fit to the $D_s^-\mu^+$ corrected mass for candidates in data passing the selections. The grey shaded boxes display the uncertainty in the fit model's predicted yield due to the finite Monte Carlo statistics.*

6.2.3 Fit Results

The results of the maximum likelihood fit to the $D_s^+\mu^+$ corrected mass are plotted in Figure 6.3, fitting to all events passing the signal selection. The pulls, defined for the i^{th} bin as the difference between data and model predictions, n^i_{data}, and n^i_{model} respectively and their uncertainties, σ^i_{data} and σ^i_{model} respectively, are defined as,

$$\frac{n^i_{\text{data}} - n^i_{\text{model}}}{\sqrt{(\sigma^i_{\text{data}})^2 + (\sigma^i_{\text{model}})^2}}. \tag{6.11}$$

Sample	Yield / 10^3
$B_s^0 \to D_s^+ \mu^+ \nu_\mu$	197.9 ± 11.9
$B_s^0 \to D_s^{*-} \mu^+ \nu_\mu$	366.0 ± 17.5
$B_s^0 \to D_{s0,1,2}^{*-} \mu^+ \nu_\mu$	21.1 ± 14.5
$B_s^0 \to D_s^+ \tau^+ \nu_\mu$	21.3 ± 4.1
$B \to D_s^+ DX$	37.0 ± 13.7
$B_s^0 \to D_s^+$ Fake$(\mu^+)X$	0.6 ± 1.3

Table 6.2 *Fit results for all components of the fit used to obtain the* $B_s^0 \to D_s^- \mu^+ \nu_\mu$ *yield.*

The $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yield is found to be $(197.9 \pm 11.9) \times 10^3$. The signal and background yields obtained from the fit are provided in Table 6.2. The results of the $D_s^- \mu^+$ fit are validated by performing 1000 fits to pseudo-data. The data template in each pseudo-data fit is replaced with a *toy template* generated by randomly selecting points from the fit templates. Consequently the yields of each fit component are known precisely. The fit templates used in the fits to pseudo-data are *statistically compatible copies*, i.e. the contents of each bin is replaced by a random number sampled from a Gaussian distribution centred on the bin contents with width equal to the bin uncertainty. The $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yield in the pseudo-data is fixed at 197.9×10^3 while the yields of all backgrounds are chosen by randomly sampling a Gaussian distribution centred on the yield determined from the fit and a width set to the component's uncertainty. The mean and width of the Gaussians are provided in Table 6.2. The distribution of the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yield for all 1000 fits to the pseudo-data is plotted in Figure 6.4 alongside the pull distribution. The pull is defined as $(N_{\text{Fit}} - N_{\text{In}})/\sigma_{\text{Fit}}$. where N_{Fit} and σ_{Fit} are the yield and uncertainty obtained from the fit to pseudo-data and N_{In} is the true number of $B_s^0 \to D_s^- \mu^+ \nu_\mu$ events in the pseudo-data. The pulls should be centred at 0 and follow a Gaussian distribution with a width of 1 As seen in Figure 6.4 the pull distribution of the toy fits is well fit by a Gaussian with a which is slightly offset and has a width slightly less than 1 implying that the fit uncertainties are overestimated. A conservative approach is taken and the narrow width is ignored. The offset is treated as a systematic error and detailed later in Section 6.4.

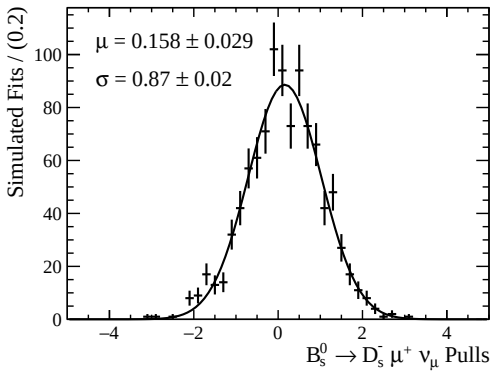


Figure 6.4 *A distribution of pulls from fits to 1000 pseudo datasets. The input $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yield is fixed for each toy fit and the variation in results is due to statistical uncertainties in the fit.*

6.3 Signal Fit

A two stage fit is is used to determine the signal and background yields in the $K^- \mu^+$ data. The results of the first fit are used to calculate the $B_s^0 \to K^- \mu^+ \nu_\mu$ branching fraction and the second fit uses the results of the first as a constraint with the results used to measure $|V_{ub}|/|V_{cb}|$ in two bins of q^2. An initial fit is performed with no selection on the q^2 of the B_s^0 candidate and uses a corrected mass range of 2500 MeV/c^2 $< m_{\rm corr} <$ 5700 MeV/c^2. Sections 6.5.2 onwards discuss the corrections applied to Monte Carlo simulation to correct the selection efficiencies. It is important to note that when producing input histograms for the fit the entries in each bin are weighted by kinematic, PID and tracking corrections detailed in Sections 5.6.6, 6.5.2 and 6.5.3.

The purpose of the first fit is to accurately determine the B_s^0 yield in order to measure the ratio of branching fractions. The second fit requires the q^2 solution to be valid, reducing the fit range to 2500 MeV/c^2 $< m_{\rm corr} < m_{B_s^0}$, and a simultaneous fit is performed in two bins of q^2, with the bin boundary placed at $q^2 = 7 \mathrm{GeV}^2/\mathrm{c}^4$. The combined yields of the fits to the high and low q^2 samples are constrained to the values obtained from the first fit,

$$N_{q^2<7} + N_{q^2>7} = N \times \varepsilon_{q^2>0}, \tag{6.12}$$

where $N_{q^2 \gtrless x}$ is the yield given a q^2 selection, N is the yield with no selection and $\varepsilon_{q^2>0}$ is the efficiency of requiring a q^2 selection. $N_{q^2 \gtrless x}$ is determined from the second fit, N is determined from the first fit, and $\varepsilon_{q^2>0}$ is determined from Monte Carlo.

The purpose of the second fit is to take our pre-existing knowledge of the yields obtained from the first fit and precisely determine the fractions of the yields in the high and low q^2 bins, thus allowing for a measurement of $|V_{ub}|$ to be performed.

The dominant backgrounds in the fits to extract the $B_s^0 \to K^-\mu^+\nu_\mu$ yields include decays from the excited K^* resonances, many $b \to c$ decays, combinatoric combinations of a kaon and muon, and candidates containing misidentified kaons. The most concerning background is the partially reconstructed decay $B^+ \to J/\psi K^+$, which has a fit distribution almost identical to signal. The yields of many backgrounds may be constrained using data driven methods; the yields of misidentified kaons are constrained by measuring the efficiency of selecting/vetoing misidentified particles using a calibration sample. While the selections used in this analysis reject almost all reconstructible $B^+ \to J/\psi K^+$ candidates, there is still a significant contribution of decays where the additional muon falls outside the acceptance of the detector. Using a combination of a J/ψ mass constraint and the geometry of the decay, the B^+ peak is reconstructed and a fit is performed allowing the true yield of $B^+ \to J/\psi K^+$ events to be determined.

6.3.1 Components and Templates

The signal fit extracts the $B_s^0 \to K^-\mu^+\nu_\mu$ yield, separating it from a variety of backgrounds. The components and data sources used to generate fit templates are summarised in Table 6.3. The templates used in the fits are one dimensional histograms of the corrected $K^-\mu^+$ mass, a binning scheme is chosen with variable bin widths, the bin boundaries chosen such that the number of candidates in each bin is approximately equal. When fitting to the sample with no selection on q^2 the template contains 30 bins and covers a range $2500 < m_{\rm corr}[\ {\rm MeV}/c^2] < 5750$. When fitting to the high and low q^2 regions the templates contain 25 and 20 bins respectively, in both cases the fit range is $2500 < m_{\rm corr}[\ {\rm MeV}/c^2] < m_{B_s^0}$. The candidates for the template input histograms originate from a variety of sources including simulated Monte Carlo decays and background enriched data. The fit templates are summarised in Table 6.3 along with the source of data used in the creation of the templates. Events originating from Simulated Monte Carlo decays are weighted to correct for differences in kinematics and mismodelling between the simulation and data.

Three excited resonances of the kaon are considered as backgrounds to the $B_s^0 \to K^-\mu^+\nu_\mu$ decay, the $K^{*-}(892)$, $K_2^{*-}(1430)$ and $K_0^{*-}(1430)$. The background

templates corresponding to the excited resonances of the kaon are combined into a single template with equal contributions. The merging of templates is motivated by a lack of knowledge on the relative branching fractions of the different decays and in part by the similarity of the template shapes. As the template shapes of K^{*-} decays are almost identical it is impossible to distinguish between the different excited resonances, therefore the $B_s^0 \to K^{*-}\mu^+\nu_\mu$ decay is taken to mean the combination of all excited K^{*-} decays. The corrected $K^-\mu^+$ mass for each template and the combination is plotted in Figure 6.5 (left).

To aid in plotting, the templates containing candidates with a misidentified particle are combined into a single template. As the yields of misidentified particles are determined externally to the signal fit in Section 6.3.3 and constrained, the merging of templates has no impact on the overall quality of the fit. The corrected $K^-\mu^+$ distributions for the misidentified particles are plotted in Figure 6.5 (right) with the combined template.

	Component	Source
Signal	$B_s^0 \to K^-\mu^+\nu_\mu$	Monte Carlo
	$B_s^0 \to K^{*-}\mu^+\nu_\mu$	Monte Carlo
	$B^+ \to J/\psi\, K^+$	Monte Carlo
	$B^+ \to J/\psi\ \phi$	Monte Carlo
	Combinatorics	*Mixing* of K^- and μ^+ from different events
	Misidentified particles	Signal candidates in data failing PID selections
	$b \to c \to s$	Monte Carlo

Table 6.3 *Sources of data used to generate the corrected mass histograms for each $B_s^0 \to K^-\mu^+\nu_\mu$ fit component.*

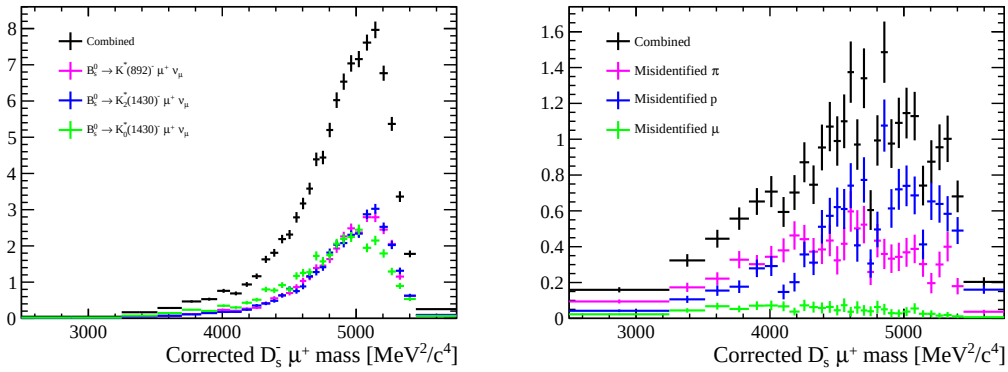


Figure 6.5 *The three $B_s^0 \to K^{*-}\mu^+\nu_\mu$ templates (left) are combined into a single template with all contributions given an equal weight. The three sources of misidentified kaons are combined into a single template (right) with the yields for each template determined using the PIDCalib package.*

6.3.2 $B^+ \to J/\psi K^+$ Yield Constraint

The $B^+ \to J/\psi K^+$ background is very effectively removed through the use of charged isolation and BDTs. These selections, however, are only effective when the additional muon is reconstructible, which is often not the case. The most likely case occurring when the muon is produced outside the acceptance of the detector. $B^+ \to J/\psi K^+$ decays with a non reconstructible muon present a major concern, as the efficiency of selecting events is similar to that of signal decays and shape of the reconstructed corrected mass is almost identical to the signal shape.

Using an approach similar to that detailed in Section 5.1.2 it is possible to reconstruct a B^+ mass peak. The momentum of the invisible muon perpendicular to the B^+ flight direction, $p_\perp$, must be equal and opposite the momentum of the visible particles. The momentum of the invisible muon parallel to the B^+ flight direction, or longitudinal momentum, $p_\parallel$, may be found from a knowledge of the J/ψ mass. When calculating the J/ψ mass from the $K^+\mu^+\mu^-$ four vectors, the only unknown component is the longitudinal momentum of the invisible muon. Solving the four vector equation yields $p_\parallel$ with a two fold ambiguity,

$$\begin{aligned} p_\perp(\mu^-) &= -p_\perp(K^-\mu^+), \\ p_\parallel(\mu^-) &= \frac{\pm\sqrt{A^2 + BC^2 - B} - AC}{C^2 - 1}, \end{aligned} \tag{6.13}$$

where,

$$\begin{aligned} A &= \frac{m^2_{J/\psi} - m^2_\mu + {p_\perp}^2(K^-) + E^2(\mu^+) + {p_\parallel}^2(\mu^+) - {p_\perp}^2(\mu^+K^-)}{2E(\mu^+)}, \\ B &= m^2_\mu + {p_\perp}^2(\mu^+K^-), \\ C &= \frac{p_\parallel(\mu^+)}{E(\mu^+)}. \end{aligned} \tag{6.14}$$

This method of reconstructing the B^+ peak will be referred to as *the neutrino method*[1]. Due to an imperfect vertex resolution approximately 15% of events have unphysical solutions for $p_\parallel$, i.e. when $A^2 + BC^2 - B < 0$.

The $B^+ \to J/\psi K^+$ yield is obtained by performing a binned maximum likelihood fit to a histogram containing both solutions of the B^+ invariant mass. The signal peak is modelled using a double Gaussian and the background shape is modelled

[1]Using a mass constraint in combination with momentum asymmetries has traditionally been used to reconstruct neutrinos.

		(1) Full Sel.	(2) No BDT Iso.	(3) No BDT *sPlot*	$\varepsilon_{\text{Reco.}}$ [%]
No q^2 Sel.	M.C.	2257 ± 51	127800 ± 400	105800 ± 400	82.8 ± 0.4
	Data	2220 ± 100	78700 ± 500	58400 ± 300	74.2 ± 0.6
$q^2 < 7$ GeV$^2/c^4$	M.C.	217 ± 15	8700 ± 100	7470 ± 130	85.9 ± 1.8
	Data	270 ± 33	5100 ± 100	4290 ± 70	84.1 ± 2.1
$q^2 > 7$ GeV$^2/c^4$	M.C.	2116 ± 50	112800 ± 400	96900 ± 400	85.9 ± 0.6
	Data	2104 ± 100	69800 ± 400	54500 ± 300	78.1 ± 0.6

Table 6.4 *The $B^+ \to J/\psi K^+$ yields obtained from a maximum likelihood fit to the B^+ invariant mass. Fits are performed to the B^+ invariant mass, 1, after a full selection using the neutrino method, 2, before BDT selections using the least isolated track and 3, before BDT selections using the neutrino method with a* sPlot *background subtraction performed from the results of fit 2. The efficiency is the ratio of events from fit 3 and fit 2.*

with the sum of a Gaussian and a Crystal Ball function. The Crystal Ball function consists of a Gaussian in the central region with a power-law end tail [98]. The function is given by:

$$f(x;\alpha,n,\overline{x},\sigma) = N \cdot \begin{cases} \exp(-\dfrac{(x-\overline{x})^2}{2\sigma^2}), & \text{for } \dfrac{x-\overline{x}}{\sigma} > -\alpha \\ A \cdot (B - \dfrac{(x-\overline{x})}{\sigma})^{-n}, & \text{for } \dfrac{x-\overline{x}}{\sigma} \leq -\alpha \end{cases} \tag{6.15}$$

where,

$$\begin{aligned} A &= \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), & B &= \frac{n}{|\alpha|} - |\alpha| \\ C &= \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), & N &= \frac{1}{\sigma(C+D)}, \\ D &= \sqrt{\frac{\pi}{2}}\left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right). \end{aligned} \tag{6.16}$$

The fit is performed in a two stage process. An initial fit is performed to the Monte Carlo distribution to determine the signal shape and a second fit is performed on the data distribution to determine the signal and background yields. The signal shape is fixed when fitting the data using the results from the Monte Carlo fit. Uncertainties in the signal and background shape make up the dominant systematic uncertainty when determining the $B^+ \to J/\psi K^+$ yield. To determine the true $B^+ \to J/\psi K^+$ yield independent of the additional muon the fit results must be divided by the efficiency of reconstructing the additional muon. Before applying BDT selections a B^+ peak may be reconstructed by calculating the

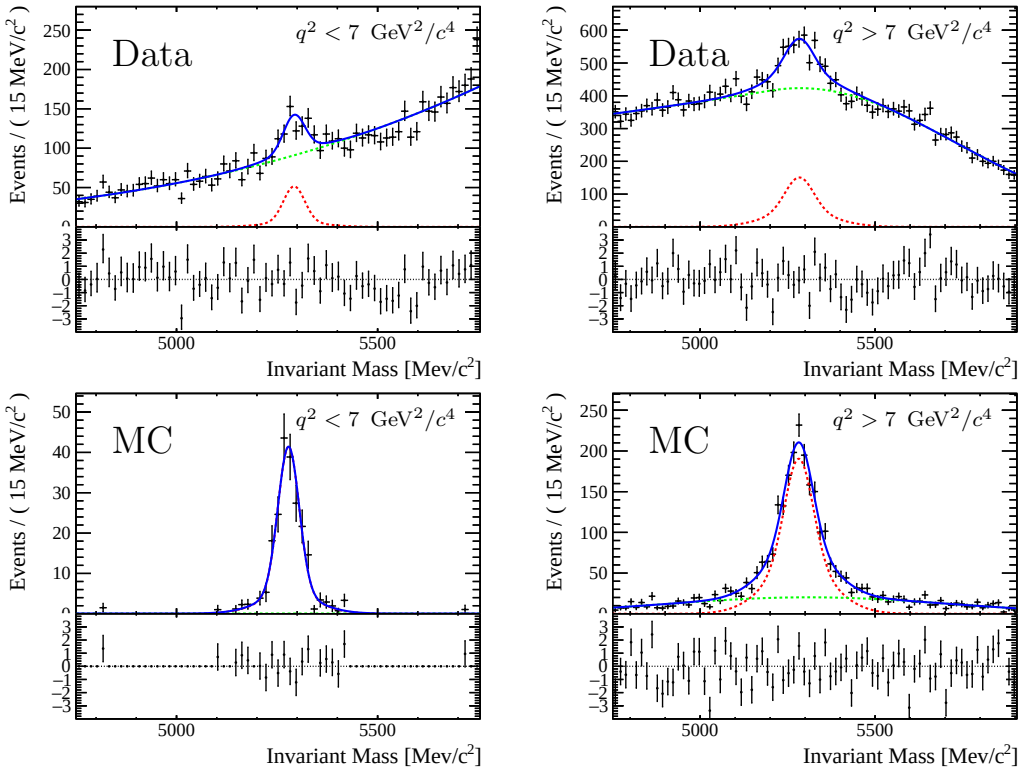


Figure 6.6 *The $K^-\mu^+\mu^-$ invariant mass reconstructed from a $K^-\mu^+$ pair using a knowledge of the B flight direction. Fit results are plotted for Monte Carlo and data in both q^2 bins. The Monte Carlo background is from the incorrect muon solution.*

invariant mass of the $B_s^0 \to K^-\mu^+\nu_\mu$ candidate combined with the least isolated track providing a relatively pure sample of B^+ candidates. The efficiency of reconstructing the additional muon using *the neutrino method* is determined by measuring the B^+ yield reconstructed using Equation 6.13 given that the candidate truly originates from a B^+. A fit is performed to the B^+ invariant mass obtained from the least isolated track from which a *sPlot* background subtraction is performed. By plotting the B^+ invariant mass distribution calculated using *the neutrino method* with the *sPlot* background subtraction applied, the ratio of B^+ yields gives the efficiency of reconstruction. The results of fits to the B^+ invariant mass are given in Table 6.4. The B^+ invariant mass calculated using *the neutrino method* is plotted in Figure 6.6 for the high and low q^2 bins. Maximum likelihood fits to the B^+ invariant mass distributions are used to obtain the $B^+ \to J/\psi\, K^+$ yield in data after applying a full selection, and the yields are divided by the efficiencies quoted in Table 6.4. A discrepancy between Monte Carlo and data of $\approx 8\%$ in the measured reconstruction efficiency is applied as a systematic. When performing a fit to determine the $B_s^0 \to K^-\mu^+\nu_\mu$ yields the $B^+ \to J/\psi\, K^+$ yield is constrained using a Gaussian constraint centred on the yield with a width set tot he statistical error given in Table 6.5.

	B^+ Yield	Statistical	Systematic
No q^2	2680	22	215
$q^2 < 7\text{Gev}^2/\text{c}^4$	314	39	22
$q^2 > 7\text{Gev}^2/\text{c}^4$	2450	20	195

Table 6.5 *The yields, constraints and systematic uncertainties for the $B^+ \to J/\psi K^+$ yield used in fits to determine the $B_s^0 \to K^-\mu^+\nu_\mu$ yield. The constraint applied to the $B_s^0 \to K^-\mu^+\nu_\mu$ fit originates from the statistical uncertainty of the fits and the systematic originates from Monte Carlo discrepancies.*

6.3.3 Misidentified Particle Yield Constraints

Despite tight selections on the likelihood criteria of the candidate kaon, some protons, pions and muons will pass the selections and be falsely reconstructed as kaons. The yields and fit distributions of particles misidentified as kaons must be determined. The rate of misidentification as a muon is considerably lower than that of the kaon, thus contributions from fake muons are not considered in the fit. A fake kaon is any particle falsely reconstructed as a kaon. A misidentified particle refers to a particle which has been misidentified, e.g. a misidentified pion

	Likelihood Selection	MisID Rate [%]	Efficiency [%]
$K^{\pm}$	$\mathcal{L}_{K/\pi} > 5$ and $\mathcal{L}_{K/p} > 5$ and $\mathcal{L}_{K/\mu} > 5$	N/A	50.8
$\pi^{\pm}$	$\mathcal{L}_{K/\pi} < 0$ and $\mathcal{L}_{p/\pi} < 0$ and $\mathcal{L}_{\mu/\pi} < 0$	0.975	71.38
p	$\mathcal{L}_{K/\pi} < 5$ and $\mathcal{L}_{p/\pi} > 0$ and $\mathcal{L}_{\mu/\pi} < 0$	1.446	29.8
$\mu^{\pm}$	$\mathcal{L}_{K/\pi} < 0$ and $\mathcal{L}_{p/\pi} < 0$ and $\mathcal{L}_{\mu/\pi} > 0$	0.325	77.5

Table 6.6 *The likelihood selection used to enrich the Fake Kaon sample with the desired particle type is given. The MisID rate is defined as the percentage of particles passing the kaon likelihood selection and the efficiency is defined as the percentage of particles passing the likelihood selection.*

is truly a pion which has been identified as e.g. a kaon.

This section details the procedure used to determine the yields of each misidentified particle type. A dedicated stripping line is written with selections identical to the $B_s^0 \to K^-\mu^+\nu_\mu$ line, except the likelihood criteria on the kaon is removed. A prescale[2] of 0.02 is applied. It is implied that all yields with prescales have been correctly scaled. This is referred to as the fake kaon sample. A full selection is applied to the fake kaon sample. The kaon candidates in data are a blend of misidentified particles and true kaons. One may produce background samples containing misidentified particles with a high purity by simultaneous requiring that the candidate kaon has a low kaon likelihood, $\mathcal{L}_{K^\pm}$, and a high likelihood for the desired particle under investigation. When searching for misidentified $\Lambda_b^0 \to p\mu^-\overline{\nu}_\mu$ decays, a sample of high purity protons misidentified as kaons may be created by requiring the candidate kaon has a low kaon likelihood and a high proton likelihood.

The selections used to produce enriched samples of the different particle types are listed in Table 6.6 with the rates of misidentification and efficiency of the enrichment selection. The misidentification rates and efficiency of selection are calculated using the PIDCalib package [99].

The yields of events passing the enrichment selections are listed in Table 6.7 alongside the scaling used to convert the enriched yield into the yield in data. The uncertainties quoted for the data yields originate from the limited yields in the fake muon sample. When performing a fit to the $K^-\mu^+$ corrected mass a Gaussian constrained is applied to the yields of the fake samples with a mean at the derived data yield and a width equal to the uncertainty. The template shapes for the misidentified kaons are plotted in Figure 6.5. During plotting of fit results

[2]A random scaling used to reduce the rate. Discussed in Section 5.6.2

	Enriched *Yield*	Scaling	Data *Yield*	Data *Yield* $q^2 < 7$	Data *Yield* $q^2 > 7$
$\pi^{\pm}$	55650	0.0137	762 ± 23	496 ± 18	243 ± 13
p	18800	0.0485	911 ± 47	320 ± 28	512 ± 35
$\mu^{\pm}$	28900	0.0042	121 ± 5.0	86 ± 4	33 ± 3

Table 6.7 *The yields of particles within the enriched regions selected using the selection in Table 6.6. The Data Yield is the yield of misidentified particles passing the full selection. The scaling converts the enriched yield to the data yield, and is the ratio of columns two and three in Table 6.6. The dominant uncertainty on the data yields originates from the limited statistics in the enriched sample.*

the templates for misidentified particles are merged into a single template.

6.3.4 Fit Model

The same Beeston Barlow fit method detailed in Section 6.1 and used to determine the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ yield is used to extract the $B_s^0 \to K^- \mu^+ \nu_\mu$ yield. However a two stage fit is performed in order to first extract the $B_s^0 \to K^- \mu^+ \nu_\mu$ yield in data, and then determine the relative fractions in the high and low q^2 bins.

The corrected mass distribution for combinatoric $K^- \mu^+$ combinations in the region below $m_{B_s^0}$ is incredibly similar to the $B_s^0 \to K^- \mu^+ \nu_\mu$ corrected mass shape. Both have disappearing tails at low corrected mass, however the $B_s^0 \to K^- \mu^+ \nu_\mu$ peaks at $m_{\text{corr}} = m_{B_s^0}$ while the combinatoric sample continues to rise. See Figures 5.9 and 6.12. Removing events with no q^2 solution has the unfortunate effect of removing all events with $m_{\text{corr}} > m_{B_s^0}$, thus producing almost identical fit distributions. To solve this the two stage fit is used and the results of the first fit are used to constrain yields in the second fit. The first fit is performed over the full corrected mass range with no selection on the q^2 allowing the signal and combinatoric distributions to be clearly distinguished in the high corrected mass region, see Figure 6.7. The second fit is a simultaneous fit in high and low bins of q^2, with all unphysical solutions removed, and uses the results of the first fit to constrain the different yields, see Figure 6.8. To summarise, the first fit determines the absolute yields and the second fit determines their fractions in the high and low bins of q^2. Performing a simultaneous fit in both bins of q^2 without initially Constraining the total yields results in significantly larger uncertainties due to similarity in fit shapes of the signal and combinatoric shapes.

The uncertainties from the first fit are propagated through to the second fit as Gaussian constraints on the combined yield in the high and low q^2 bins. Take for example $B_s^0 \to K^-\mu^+\nu_\mu$, in the second fit a Gaussian constraint is applied to the combined yield in both bins with a value equal to the yield from the first fit multiplied by the efficiency of requiring a valid q^2 solution and a width equal to the uncertainty from the first fit. All parameters in the fit are listed in Table 6.8. As detailed in previous sections the $B^+ \to J/\psi\, K^+$ yield and yields of misidentified particles are obtained externally to the fit and the yields of the components are given a Gaussian constraint. Additional constraints are used to constrain some yields relative to others, most notably the $B_s^0 \to J/\psi\ \phi$ yield is constrained to the $B^+ \to J/\psi\, K^+$ yield using the knowledge of relative fragmentation fractions, branching fractions and efficiencies. All constraints used in the fit are listed in Table 6.9.

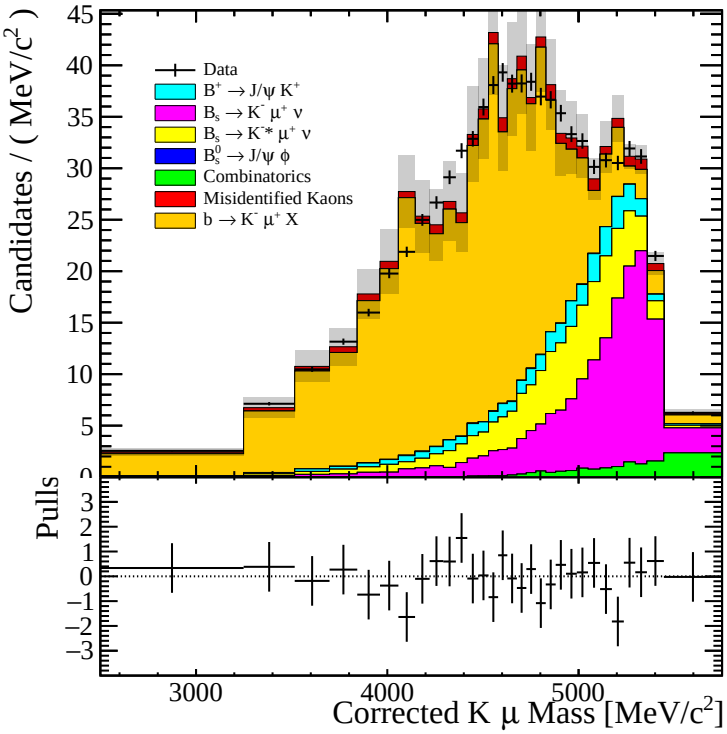


Figure 6.7 *A fit to the corrected $K^-\mu^+$ mass distribution for data candidates passing the selections. The uncertainty in the predicted data yield for each bin is shaded in grey. The pulls for each bin, i, are shown underneath the fit.*

	Fit #1	Fit #2	
	No q^2 sel.	$q^2 < 7$ GeV$^2/c^4$	$q^2 > 7$ GeV$^2/c^4$
$B_s^0 \to K^-\mu^+\nu_\mu$	$Y_{K^-\mu^+}$	$Y_{K^-\mu^+} \times \varepsilon_{q^2>0}$	
		$f_{q^2<7}$	$1 - f_{q^2<7}$
$B_s^0 \to K^{*-}\mu^+\nu_\mu$	$Y_{K^{*-}\mu^+}$	$Y_{K^{*-}\mu^+} \times \varepsilon_{q^2>0}$	
		$f_{q^2<7}$	$1 - f_{q^2<7}$
$B^+ \to J/\psi K^+$	$Y_{J/\psi K^+}$	$Y_{J/\psi K^+ \ \|q^2<7}$	$Y_{J/\psi K^+ \ \|q^2>7}$
$B_s^0 \to J/\psi \ \phi$	$R \times Y_{J/\psi K^+}$	$R \times Y_{J/\psi K^+ \ \|q^2<7}$	$R \times Y_{J/\psi K^+ \ \|q^2>7}$
$b \to c \to s$	$Y_{\text{inc.}}$	$Y_{\text{inc.}} \times \varepsilon_{q^2>0}$	
		$f_{q^2<7}$	$1 - f_{q^2<7}$
Combinatorics	$Y_{\text{Combi.}}$	$Y_{\text{Combi.}} \times \varepsilon_{q^2<7}$	$Y_{\text{Combi.}} \times \varepsilon_{q^2>7}$
$\pi \to K$ MisID	$Y_{\pi\to K}$	$Y_{\pi\to K \ \|q^2<7}$	$Y_{\pi\to K \ \|q^2>7}$
$p \to K$ MisID	$Y_{p\to K}$	$Y_{p\to K \ \|q^2<7}$	$Y_{p\to K \ \|q^2>7}$
$\mu \to K$ MisID	$Y_{\mu\to K}$	$Y_{\mu\to K \ \|q^2<7}$	$Y_{\mu\to K \ \|q^2>7}$

Table 6.8 *Components of the two fits used to determine the $B_s^0 \to K^-\mu^+\nu_\mu$ yields, Y, in data are presented. Yields shaded in yellow are determined from the fit and left completely free, yields shaded in green are determined externally to the fit and their values are Gaussian constrained, and yields shaded in blue for Fit #2 are Gaussian constrained to the results obtained from the fit #1. The $B_s^0 \to J/\psi \ \phi$ yield is determined by scaling the $B^+ \to J/\psi K^+$ yield by the relative yields, $R = f_s/f_d \times \varepsilon_{J/\psi\phi}/\varepsilon_{J/\psi K^+} \times \mathcal{B}(J/\psi K^+)/\mathcal{B}(J/\psi\phi)$. All efficiencies, ε, are determined from corrected Monte Carlo simulations. Considering the $B_s^0 \to K^-\mu^+\nu_\mu$ component, fit #1 determines the yield and fit #2 determines the distribution of the yield in the high and low q^2 bins.*

	Physics constraints		
f_s/f_d	0.252 ± 0.012 [100, 101]		
$\mathcal{B}(B^+ \to J/\psi K^+)$	$(1.01 \pm 0.03) \times 10^{-3}$ [37]		
$\mathcal{B}(B_s^0 \to J/\psi \phi)$	$(1.08 \pm 0.08) \times 10^{-3}$ [37]		
$\mathcal{B}(\phi \to K^- K^+)$	0.492 ± 0.005 [37]		
	Yield constraints		
	No q^2 sel.	$q^2 < 7$ GeV$^2/c^4$	$q^2 > 7$ GeV$^2/c^4$
$B^+ \to J/\psi K^+$	2357 ± 127	279 ± 43	1607 ± 154
$B_s^0 \to J/\psi \phi$	62 ± 3	6 ± 3	47 ± 5
$\pi \to K$ MisId	762 ± 23	496 ± 18	243 ± 13
$p \to K$ MisId	911 ± 47	320 ± 28	512 ± 35
$\mu \to K$ MisId	121 ± 5	86 ± 4	33 ± 3
	Simultaneous Fit yield constraints		
$B_s^0 \to K^- \mu^+ \nu_\mu$	8550 ± 760		
$B_s^0 \to K^{*-} \mu^+ \nu_\mu$	1760 ± 350		
$b \to c \to s$	35580 ± 740		
Combinatorics		790 ± 160	622 ± 120

Table 6.9 *A summary of the constraints and fit values entering the signal fit. During the second fit some values are constrained in both the high and low q^2 bins, e.g. the combinatoric yield, while for other components the combined sum of entries in both the high and low q^2 bins is constrained, e.g. the $B_s^0 \to K^- \mu^+ \nu_\mu$ yield.*

6.3.5 Fit Results

Results from the first signal fit to determine the $B_s^0 \to K^- \mu^+ \nu_\mu$ yield using the corrected $K^- \mu^+$ mass are plotted in Figure 6.7. The observed number of $B_s^0 \to K^- \mu^+ \nu_\mu$ events is 10050 ± 880. A significant peaking structure is observed in the corrected $K^- \mu^+$ mass distribution, at the mass of the B_s^0 meson. This corresponds to the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ and this peaking structure is the discovery of $B_s^0 \to K^- \mu^+ \nu_\mu$ decays. The family of decays $B_s^0 \to K^{*-} \mu^+ \nu_\mu$ is also observed for the first time although the individual contributions from the $K^{*-}(892)$, $K_2^{*-}(1430)$ and $K_0^{*-}(1430)$ are not individually measured. The results of the second fit are plotted in Figure 6.8, for the low (left) and high (right) q^2 bins, the signal purity is considerably higher in the low q^2 bin and the $B_s^0 \to K^- \mu^+ \nu_\mu$ contribution is clearly required in order to account for the large number of events in the high corrected mass region. The signal yield in the low q^2 bin is 5160 ± 470, and in the high q^2 bin is 3280 ± 430.

A Monte Carlo method is used to validate the signal fit results, 1000 or 500 pseudo-datasets are generated by randomly sampling the Monte Carlo input input

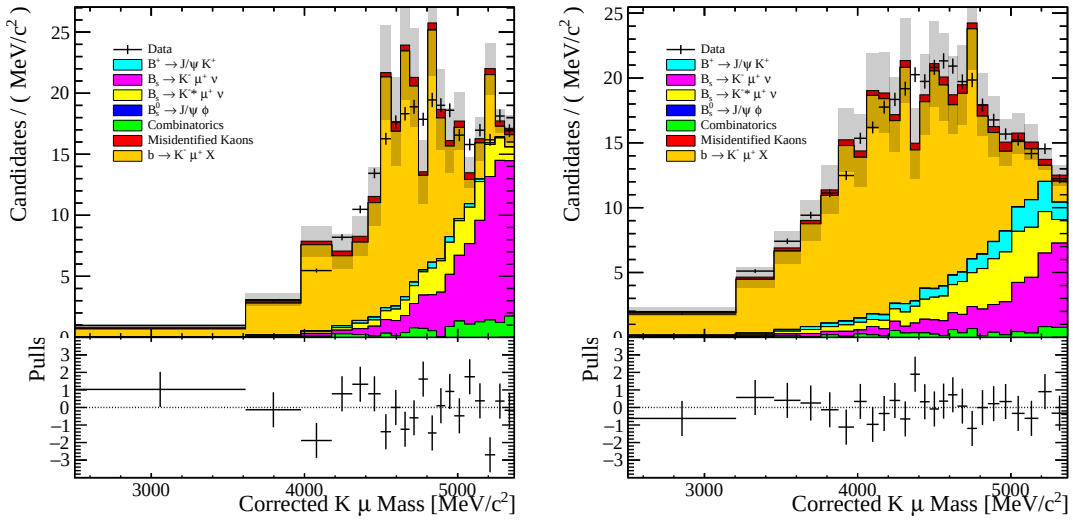


Figure 6.8 *A simultaneous fit in two bins of q^2 performed on the corrected $K^-\mu^+$ mass distribution for data candidates passing the selections.*

histograms. The $B_s^0 \to K^-\mu^+\nu_\mu$ yield in the pseudo-dataset is constant and set to the value obtained in the fits, the yields of all other components are randomly varied by selecting a point on a Gaussian distribution with a mean centred on the fit result with a width set to the fit uncertainty. The distribution of toy pulls should follow the normal distribution, be centred at zero and have a width of one. An offset distribution is indicative of biases present in the fit and a width differing from one indicates that the uncertainty on the fitted yield is being incorrectly estimated.

The pull distributions of the first fit are plotted in Figure 6.9, a slight offset of 0.14σ is observed and the width is slightly less than one indicating that the uncertainty on the signal yield is being overestimated. The pulls for second fit are plotted in Figure 6.10 for both the $B_s^0 \to K^-\mu^+\nu_\mu$ and inclusive $b \to c \to s$ contributions, an offset of 0.50σ and 0.62σ is observed in the low and high q^2 bins respectively indicating significant biasing. The widths are 0.85 and 1.01 indicating that the fit uncertainty is being overestimated in the low q^2 bin. The inclusive $b \to c \to s$ sample shows an offset of -0.37σ and -0.40σ in the high and low bins respectively indicating that the fit is unable to fully distinguish the two samples. Systematic uncertainties are assigned from the biases observed here.

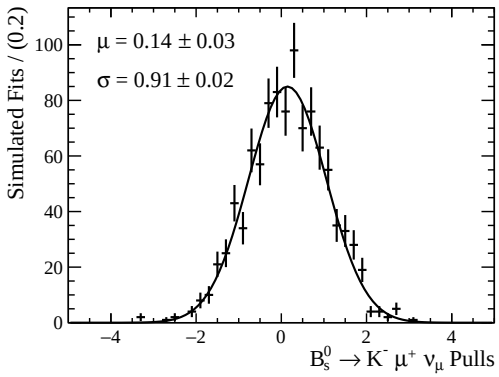


Figure 6.9 *Distributions of pulls obtained from 1000 fits to pseudo-datasets. The pull is defined as the difference between the true number of $B_s^0 \to K^-\mu^+\nu_\mu$ candidates and the yield obtained from the fit divided by the fit uncertainty.*

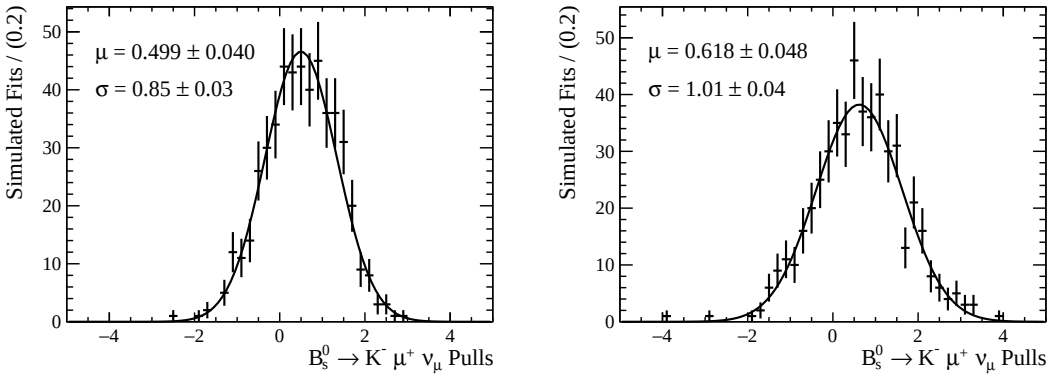


Figure 6.10 *Distributions of pulls obtained from 500 fits to pseudo-datasets for the low (left) and high (right) q^2 bins.*

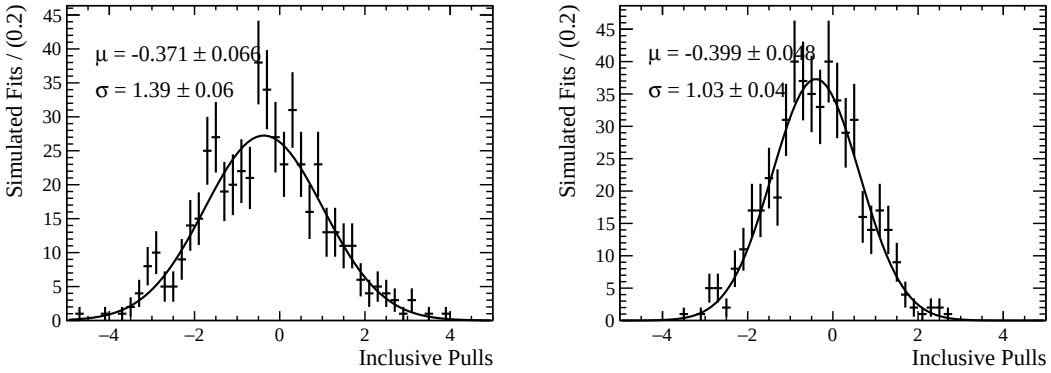


Figure 6.11 *Distributions of pulls for the $b \to c \to s$ template obtained from 500 fits to pseudo-datasets for the low (left) and high (right) q^2 bins.*

6.4 Systematic Uncertainties

The uncertainty on the $B_s^0 \to K^-\mu^+\nu_\mu$ yield obtained from the signal fit contains several systematic uncertainties. The fit does not account for systematics originating from the variation of the corrected mass shape associated with varying form factor models, the uncertainty of the $B^+ \to J/\psi K^+$ yield due to a limited knowledge of the reconstruction efficiency. Additionally the uncertainty on the signal and background yields does not consider the fact that the signal fit may be biased. For the fit used to obtain the $B_s^0 \to D_s^-\mu^+\nu_\mu$ yield the only systematic effect considered is the bias present in the fit.

The systematic uncertainty originating from the variation in the corrected mass template shape is investigated by generating the template shape with different corrections and weights applied. The $B_s^0 \to K^-\mu^+\nu_\mu$ template shape is plotted in Figure 6.12 reconstructed using all form factor hypotheses, with and without the addition of weights correcting Monte Carlo Simulation. A systematic uncertainty originating from the uncertainty in the corrected mass template shape is determined by repeating the signal using different possible template shapes. The systematic uncertainty assigned to variations in the template shape are summarised in Table 6.10. As discussed in Section 2.3.2 the uncertainty on the form factor shape is lowest at high q^2 resulting in a greater variation in the corrected mass template in the low q^2 bin; this is reflected in the calculated systematic uncertainty.

A systematic uncertainty on the$B^+ \to J/\psi K^+$yield due to an uncertainty on the parametrisation of the background shape is detailed in Section 6.3.2 and the systematic uncertainty on the $B_s^0 \to K^-\mu^+\nu_\mu$ yield is given in Table 6.10.

The systematic uncertainty on the $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ yield associated with a biased fit is quantified in Sections 6.3.5 and 6.2.3 by performing 1000 or 500 fits to pseudo-data. The systematic uncertainty is assigned by taking the mean of the pull distributions on the $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ yields of 1000 or 500 fits to pseudodata, and are listed in Table 6.10.

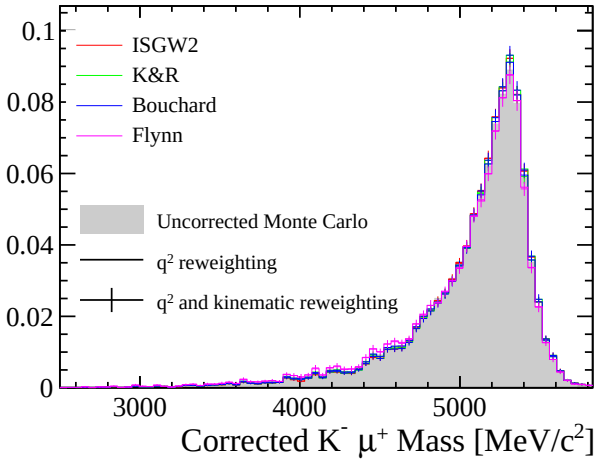


Figure 6.12 *The Corrected $K^-\mu^+$ mass distribution for simulated $B_s^0 \to K^-\mu^+\nu_\mu$ decays is plotted with different corrections applied. The shaded grey region consists of the uncorrected Monte Carlo with full selections applied. The lines display Monte Carlo with form factor corrections applied and the points represent Monte Carlo with form factor and kinematic corrections applied.*

$\sigma_{\text{syst.}}[\%]$	No q^2 Sel.	$q^2 < 7$ GeV$^2/c^4$	$q^2 > 7$ GeV$^2/c^4$
Template variation	1.36	3.64	0.87
$B^+ \to J/\psi K^+$ reconstruction	2.07	0.61	3.79
Fit Bias, $B_s^0 \to K^-\mu^+\nu_\mu$	1.22	4.55	8.09
Fit Bias, $B_s^0 \to D_s^-\mu^+\nu_\mu$	0.59		

Table 6.10 *Systematic uncertainties on the $B_s^0 \to K^-\mu^+\nu_\mu$ yield due uncertainties on the template shape, $B^+ \to J/\psi K^+$ reconstruction and biases within the fit. The systematic uncertainty on the $B_s^0 \to D_s^-\mu^+\nu_\mu$ yield due to biases in the fit is included.*

6.5 Relative Efficiency Determinations and corrections

6.5.1 Generator Efficiency

A pre-selection is applied to Monte Carlo events before the simulation of particle interactions with the detector. These selections are called *generator cuts* as they are applied immediately after the generation of the decay. A selection is made on the polar angle, θ_{flight},

$$0.01 < \theta_{\text{flight}} < 0.4, \tag{6.17}$$

	$\varepsilon_{\text{Gen.}}$ [%]	$\sigma_{\text{stat.}}$ [%]	$\sigma_{\text{FF.}}$ [%]	$\sigma_{\text{comb.}}$ [%]
$B_s^0 \to D_s^- \mu^+ \nu_\mu$	17.87	0.08		
$B_s^0 \to K^- \mu^+ \nu_\mu$	20.51	0.08	0.23	0.24
$B_s^0 \to K^- \mu^+ \nu_\mu$ $q^2<7$ GeV^2/c^4	19.67	0.12	0.03	0.12
$B_s^0 \to K^- \mu^+ \nu_\mu$ $q^2>7$ GeV^2/c^4	20.96	0.11	0.16	0.19

Table 6.11 *Generator efficiencies for $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ in different q^2 regions. The uncertainties originate from Monte Carlo statistics, form factor parametrisation and are summed in quadrature. For $B_s^0 \to D_s^- \mu^+ \nu_\mu$ there is negligible variation in Form Factors between parametrisations and the form factor uncertainty is ignored.*

on some or all of the final state particles. When simulating $B_s^0 \to K^- \mu^+ \nu_\mu$ events, generator cuts are applied to all charged final state particles, and when simulating $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ events the cuts are applied to the muon and daughters of the D_s^+. For all other backgrounds the generator cuts are applied on all charged final state tracks. As this analysis measures $|V_{ub}|$ with respect to a given q^2 selection the generator efficiency must be determined for that region rather than the whole sample. To measure the generator efficiencies for $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ small Monte Carlo samples of 250,000 events are generated before generator cuts and detector simulation. The Generator efficiencies are plotted against q^2 in Figure 6.13 with the simulated q^2 distributions overlaid before and after the selection. The generator efficiencies are quoted in Table 6.11. The efficiencies for $B_s^0 \to K^- \mu^+ \nu_\mu$ vary with q^2 and a systematic error is assigned to the calculated efficiency taking into account the variations in efficiency due to variations in form factor parametrisations. The variation in efficiency is probed by reweighting the Monte Carlo to be consistent with the three form factor parametrisation from Lattice QCD and sum rules, and the systematic is taken as half the difference between the highest and lowest values of calculated efficiency.

6.5.2 Particle Identification

Particles are identified by combining information from the calorimeters, muon system and the ring-imaging Cherenkov (RICH) detectors providing excellent charged particle separation and rejection. The simulation does not accurately model the efficiency of selecting events using particle identification likelihoods and a data driven method is needed to correctly calculate the particle identification, PID, efficiencies. The PIDCalib package [99] calculates the efficiency of applying a

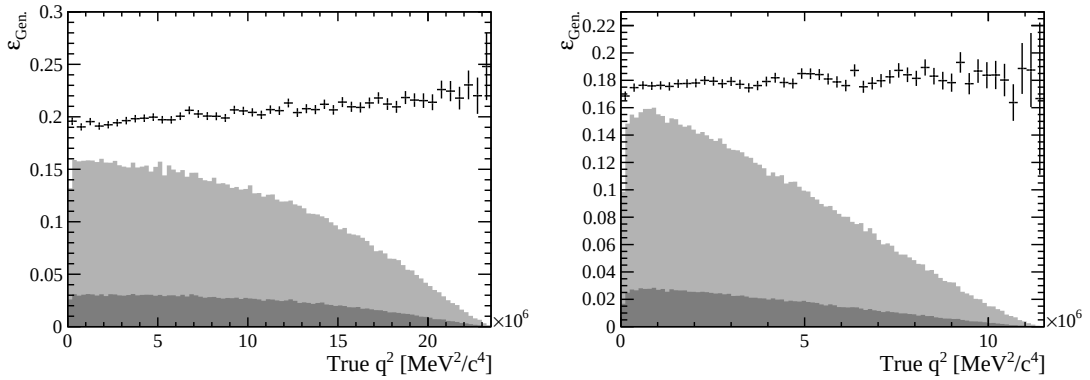


Figure 6.13 *The Generator efficiencies plotted against the true q^2 for $B_s^0 \to K^-\mu^+\nu_\mu$ (left) and $B_s^0 \to D_s^-\mu^+\nu_\mu$ (right). The q^2 distributions for the signal Monte Carlo samples are plotted in grey before and after the selections are applied.*

PID selection on an arbitrary dataset using a tag and probe method to determine the true efficiencies of a selection. The calibration decays used to calculate the PID efficiencies are listed in Table 6.12. The PID selections applied to data and simulation are listed in Table 6.13. To minimise systematic effects, tight PID selections are only applied to the opposite sign kaon and muon while for $B_s^0 \to D_s^-\mu^+\nu_\mu$, very soft selections are applied to the opposite sign π^-K^+ pair. Consequently the efficiency of PID selections will be similar for both the signal and normalisation decays and systematic effects are reduced when calculating corrections to the ratio of efficiencies.

Decay	Tag	Probe
$D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$	*soft* π^+	K^-
$D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$	*soft* π^+	π^+
Detached $J/\psi \to \mu^+\mu^-$	$\mu^\pm$	$\mu^\mp$
$\Lambda \to p\pi^-$	π^-	p

Table 6.12 *The decays used to calibrate PID efficiencies. The low momentum (soft) tag π^+ originates from the D^{*+} decay allowing the flavour of the D^0 to be unambiguously identified.*

μ^+	$\mathcal{L}_{\mu/\pi} > 3$	and $\mathcal{L}_{\mu/p} > 0$	and $\mathcal{L}_{\mu/K} > 0$
K^-	$\mathcal{L}_{K/\pi} > 5$	and $\mathcal{L}_{K/p} > 5$	and $\mathcal{L}_{K/\mu} > 5$
K^+	$\mathcal{L}_{K/\pi} > -2$		
π^-	$\mathcal{L}_{K/\pi} < 20$		

Table 6.13 *The PID likelihood selections applied to all particles. Selections are aligned between $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ minimising systematics when taking the ratio of efficiencies.*

The PID efficiencies are calculated using a *fit and count* method, A fit is

performed to the invariant mass distribution of the parent particle and *sWeights* are calculated. The efficiency is taken as the ratio of the sum of the *sWeights* before and after the PID selection. The PID efficiency varies with the kinematics of the track under consideration and the conditions of the underlying event, consequently differences in kinematics between the calibration sample and the signal sample could result in systematic differences in PID efficiency. To minimise systematic errors a lookup table binned in momentum, pseudorapidity and track multiplicity in the underlying events is generated with each entry containing the PID efficiency for that region of data. When choosing a binning scheme for the lookup table one must choose a binning scheme with a trade off between variance and bias. With a low number of bins the statistics in each bin will be high ensuring a precise measurement of the efficiency however the intra-bin variation in efficiency will be higher resulting in a biased measurement. With a large number of bins the intra-bin variations in efficiency will be minimised however the statistics in each bin will be lower resulting in an efficiency measurement with greater variance. Additional systematic effects are introduced via the *sWeight* procedure used to determine the yields of the calibration sample before and after PID selections in each bin. Variations in the shape of the signal peak or background distributions introduced by the application of a selection may result in the value obtained by summing the weights to differ from the yield under the signal peak resulting in unphysical values of the efficiency. When applying a loose PID selection it is not unusual to see quoted efficiencies greater than one, purely as a consequence of biases due to the *sWeight*ing procedure. A more rigorous approach would be to perform many fits of the invariant mass distribution before and after the selection and take a ratio of the yields obtained from the two fits, although this approach requires significant human input to ensure the quality of all the fits and is not feasible.

To minimise systematic effects from the intra bin variations in efficiency and the *sWeight* background subtraction a MC/Data driven correction is used instead of the pure data driven correction. The data driven correction returns a *true* efficiency value for a given PID cut, the MC/Data driven correction returns the ratio of PID efficiencies obtained from data and Monte Carlo. This ratio is used to correct the PID efficiency in the simulation.

To determine the PID efficiencies in Monte Carlo, samples are generated corresponding to the decays $D^{*+} \to D^0\pi^+$ and $J/\psi \to \mu^+\mu^-$ with a detached secondary vertex. Differences in the kinematics between the simulated samples

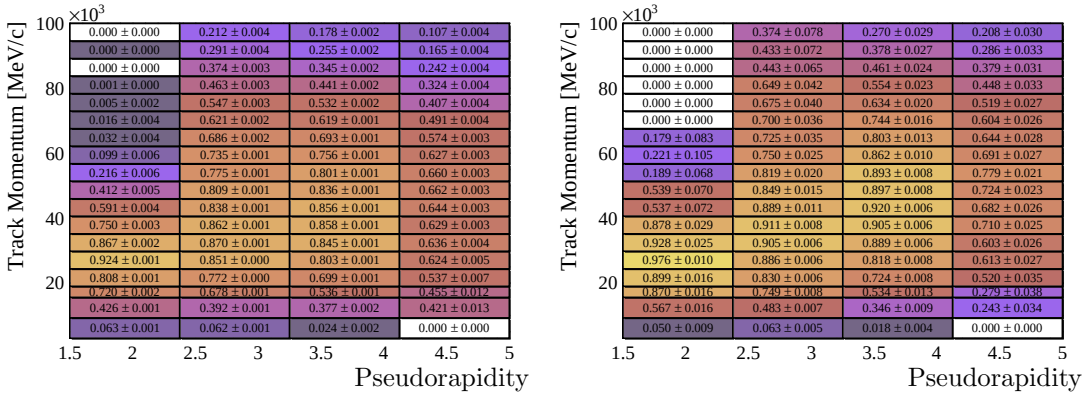


Figure 6.14 *A two dimensional projection of the PID efficiency lookup table for kaons determined from data (left) and Monte Carlo (right) $D^{*+} \to D^0\pi^+$ decays.*

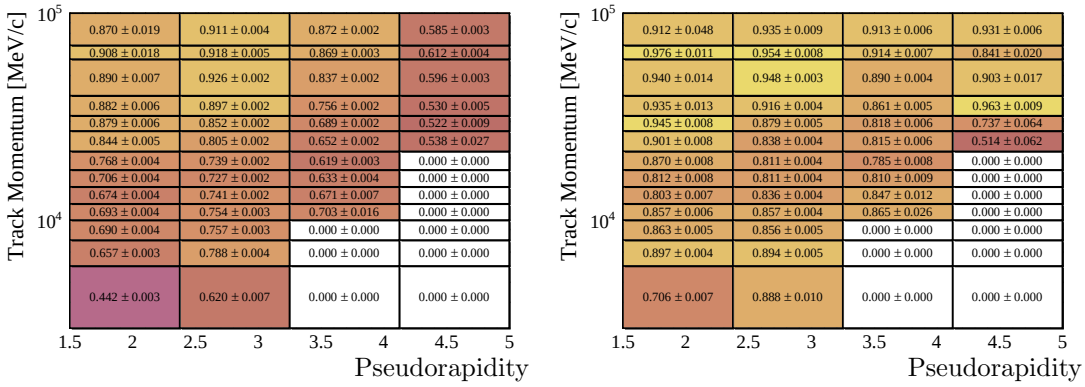


Figure 6.15 *A two dimensional projection of the PID efficiency lookup table for muons determined from data (left) and Monte Carlo (right) $J/\psi \to \mu^+\mu^-$ decays.*

and data are corrected using the GBReweighter package [90]. The target sample for the reweighting is the *sWeighted* data and the source sample is the Monte Carlo. As the target data set is weighted, the reweighting procedure has the effect of simultaneously correcting the kinematics and applying a weight mimicking the effects of the *sWeights* to the simulated sample. The only remaining discrepancy between the simulation and data are the mismodelled PID distributions. The efficiency of a PID selection in data is determined by taking the ratio of the sum of *sWeights* before and after a selection, and the efficiency in Monte Carlo is determined by taking the ratio of the sum of correcting weights before and after a selection. When comparing kinematically equal Monte Carlo and Data the intra bin variations in PID efficiency will be equal. Consequently when taking the ratio of efficiencies systematic effects from intra bin variations in efficiency cancel. This method relies on the assumption that the intra bin correction factor is constant. The calculated PID efficiencies binned in pseudorapidity and momentum for the muon and kaon are plotted in Figures 6.15 and 6.14 respectively for both data (left) and Monte Carlo (right).

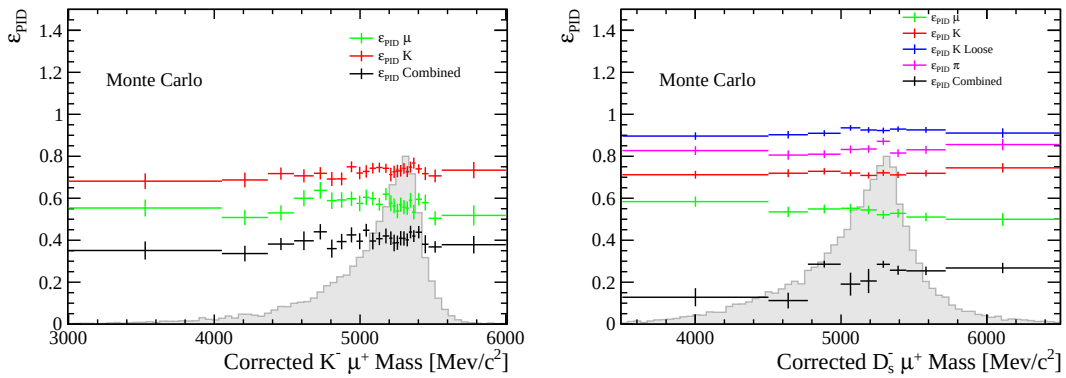


Figure 6.16 *The efficiencies of PID selections are plotted against the B_s^0 corrected mass for $B_s^0 \to K^-\mu^+\nu_\mu$ (left) and $B_s^0 \to D_s^-\mu^+\nu_\mu$ (right). The signal corrected mass distribution is shaded in light grey.*

When determining efficiencies form Monte Carlo each track from each event is weighted by the correction factor obtained from the lookup table. The corrected Monte Carlo yield is taken as the sum of the correction weights. Systematic uncertainties are quantified by performing 1000 pseudo-experiments, each time varying the contents of the lookup tables within the obtained errors. The PID corrections for each q^2 bin used in the fits is given in Table 6.14, and the PID efficiencies for Monte Carlo and data are plotted against the corrected mass in Figure 6.16.

	$B_s^0 \to K^- \mu^+ \nu_\mu$	$B_s^0 \to D_s^- \mu^+ \nu_\mu$	Ratio
No Sel.	0.855 ± 0.004	0.823 ± 0.014	1.039 ± 0.015
${q^2}_{K^-\mu^+} < 7$ GeV^2/c^4	0.850 ± 0.006		1.033 ± 0.015
${q^2}_{K^-\mu^+} > 7$ GeV^2/c^4	0.863 ± 0.002		1.048 ± 0.016

Table 6.14 *PID correction factors averaged over all tracks and all events applied to Monte Carlo. Corrected efficiencies are obtained by multiplying the Monte Carlo efficiency by the correction factor. Due to correlations between the uncertainties in the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ and $B_s^0 \to K^- \mu^+ \nu_\mu$ channels the uncertainty on the ratio is smaller than that obtained from a naive propagation of uncertainties.*

6.5.3 Tracking Correction

It is of vital importance that the efficiency of reconstructing tracks is well understood when performing a cross section or branching fraction measurement. The track reconstruction efficiency is is over 95% and is determined from Monte Carlo. A data driven correction is applied to the simulation using clean $J/\psi \to \mu^+\mu^-$ decays. The tracking reconstruction efficiency is measured using a tag and probe method, the tag muon is fully reconstructed as well identified muon and the probe track is partially reconstructed without information from at least one subdetector which is being probed. The tracking efficiency is determined by counting the amount of fully reconstructed tracks correspond to the partially reconstructed probe track. Performing the tag and probe analysis on both simulation and data yields a discrepancy of approximately 2% [102].

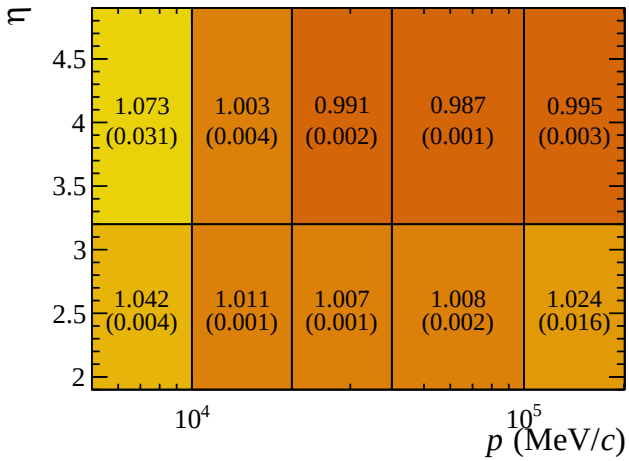


Figure 6.17 *The look-up table used to correct the tracking efficiency of charged tracks, binned in momentum and psedorapidity.*

A lookup table of the ratio of tracking efficiencies of data and Monte carlo

$K^-\mu^+$ q^2 Sel.	$B_s^0 \to K^-\mu^+\nu_\mu$	$B_s^0 \to D_s^-\mu^+\nu_\mu$	Ratio
No Sel.	1.007 ± 0.001	1.018 ± 0.005	0.990 ± 0.004
${q^2}_{K^-\mu^+} < 7$ GeV^2/c^4	1.006 ± 0.001	1.018 ± 0.005	0.989 ± 0.004
${q^2}_{K^-\mu^+} > 7$ GeV^2/c^4	1.010 ± 0.002	1.018 ± 0.005	0.992 ± 0.004

Table 6.15 *Tracking efficiency corrections applied to Monte Carlo events.*

is provided by the LHCb collaboration. The two dimensional table binned in momentum and pseudorapidity is visualised in Figure 6.17. The tracking efficiency corrections are applied as a weight on each track as determined from the lookup table and efficiencies are corrected by taking the product of the weights for each track. As $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ contain two and four charged particles in their final states the uncertainties partially cancel when taking the ratio of the efficiencies.

The uncertainties on the overall correction factor are determined by performing 1000 pseudo-experiments, each time the efficiencies in the lookup table are varied within their uncertainties. The tracking corrections to the efficiency calculations are summarised in Table 6.15 for each of the q^2 bins used in the fits.

6.5.4 $B^+ \to J/\psi\, K^+$ corrections

The decays $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ are partially reconstructed due to the missing neutrino and have broad distributions making it difficult or impossible to isolate a pure signal sample in data. In order to validate the efficiencies of a selection and ensure that biases between data and simulation are corrected, the decay $B^+ \to J/\psi\, K^+$ is used as a proxy for the signal decay. When reconstructed using only one muon the B^+ decay is kinematically very similar to the $B_s^0 \to K^-\mu^+\nu_\mu$ decay, allowing the efficiencies of selections on kinematic variables to be validated. When fully reconstructed the efficiencies of selecting B^+ decays is similar to the signal B_s^0 decay as there are no additional tracks which can be associated with the secondary vertex.

Efficiency corrections are calculated for the corrected mass uncertainty cut and the BDT response variables. The efficiency of a selection is calculated for $B^+ \to J/\psi\, K^+$ by performing a fit to the invariant mass distribution of the $\mu^-\mu^+K^+$ triad before and after a selection. The correction factor is the ratio of the efficiency in data and Monte Carlo and the efficiency for $B_s^0 \to K^-\mu^+\nu_\mu$ obtained from Monte Carlo is scaled by the correction factor. An uncertainty

	$K^-\mu^+$	$K^-\mu^+$ $q^2 > 7\text{GeV}^2/\text{c}^4$	$D_s^-\mu^+$
$\sigma m_{\text{Corr.}}$	1.02 ± 0.02	1.03 ± 0.02	
Isolation BDT	0.99 ± 0.03	1.00 ± 0.01	0.989 ± 0.014
Charged Track BDT	0.96 ± 0.03	0.96 ± 0.03	
Same Sign BDT	1.00 ± 0.04	0.95 ± 0.04	

Table 6.16 *Correction factors applied to Monte Carlo determined from simulated and real decays of* $B^+ \to J/\psi\, K^+$.

obtained from the correction factor is applied as a systematic correction. The corrections are listed in Table 6.16 and plots displaying *sPlot* background subtracted $B^+ \to J/\psi\, K^+$ data alongside Monte Carlo are given in Figures 6.18 - 6.20. The q^2 of $B^+ \to J/\psi\, K^+$ peaks at $m^2_{J/\psi}$ resulting in very few events being reconstructed in the low q^2 bin. The corrections applied in the low q^2 are set equal to those in the high q^2 bin.

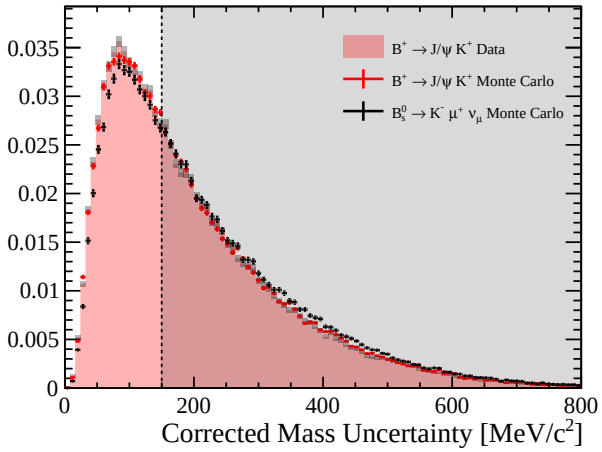


Figure 6.18 *The corrected mass uncertainty for* $B^+ \to J/\psi\, K^+$ *(red) and* $B_s^0 \to K^-\mu^+\nu_\mu$ *(black). A* sPlot *background subtraction is performed on the data. The correction factor is taken as the ratio of* $B^+ \to J/\psi\, K^+$ *data and Monte Carlo decays passing the selection. Rejected events are highlighted in the shaded region. Events rejected by the selection are in the shaded region.*

6.5.5 q^2 Migration

Having selected a neutrino solution using the linear regression method detailed in Section 5.1.2 a selection is made. The resolution on the reconstructed q^2 will result in some migration of events across the selection boundary with some events rejected that should have been selected and vice versa. The distribution of the

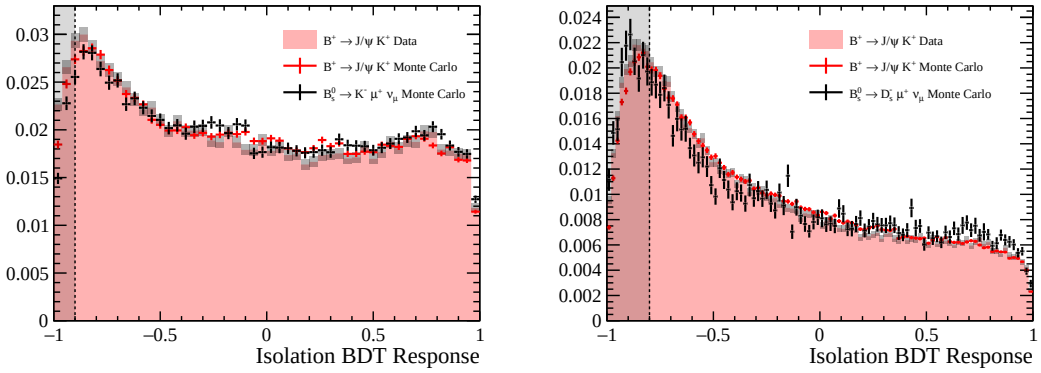


Figure 6.19 *The response of the isolation BDT for $B_s^0 \to K^-\mu^+\nu_\mu$ (left) and $B_s^0 \to D_s^-\mu^+\nu_\mu$ (right) is plotted against the $B^+ \to J/\psi K^+$ calibration samples.*

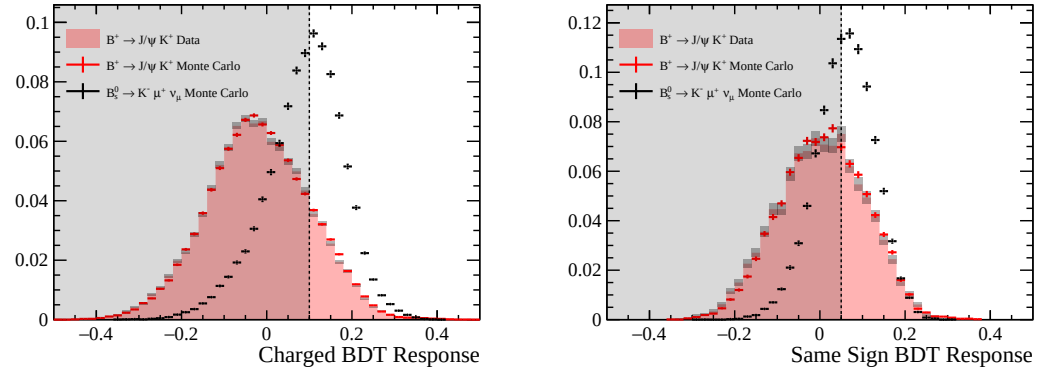


Figure 6.20 *The response of the BDTs rejecting charged (left) and same sign (right) backgrounds for $B_s^0 \to K^-\mu^+\nu_\mu$ are plotted against the $B^+ \to J/\psi K^+$ calibration samples.*

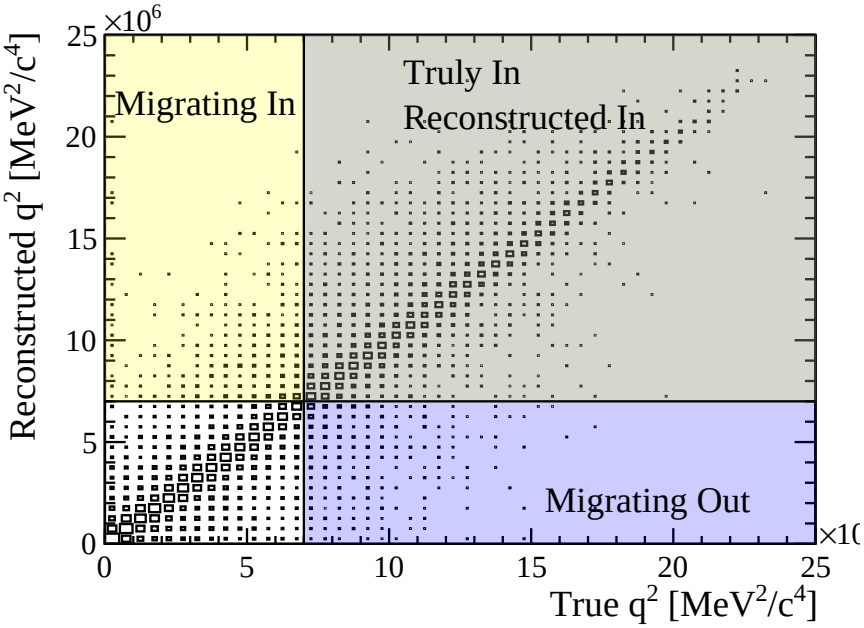


Figure 6.21 *The reconstructed q^2 solution selected using the regression model is plotted against the true q^2 for simulated $B_s^0 \to K^-\mu^+\nu_\mu$ events. The regions of inward and outward migration are shaded when requiring $q^2 > 7$ GeV^2/c^4.*

true q^2 from Monte Carlo is plotted against the reconstructed q^2 in Figure 6.21, the region containing events migrating either in or out of the high q^2 region are illustrated. Inward migration is defined by the events with a true q^2 outside the region of interest but are reconstructed inside due to the resolution. Outward migration is defined by the events which are truly in the region of interest but are reconstructed out. A correction factor is calculated from simulated Monte Carlo events by taking the ratio of events truly in the high q^2 with events reconstructed in the q^2 region. The Monte Carlo is reweighted to be consistent with form factor predictions from Lattice QCD and light cone sum rules, and the percentages of events migrating in and out are listed in Table 6.17. As the correction factor is dependant on the form factor modelling, a systematic uncertainty is assigned to the correction factor, taken as the standard deviation of the correction factors for all form factor predictions. The correction factor is taken as the mean value for all form factor models. The migration corrections and systematic uncertainties are found to be:

$$\begin{aligned} \text{Corr. Mig.}_{q^2<7\text{GeV}^2/\text{c}^4} &= 1.002 \pm 0.008 \\ \text{Corr. Mig.}_{q^2>7\text{GeV}^2/\text{c}^4} &= 0.996 \pm 0.009 \end{aligned} \tag{6.18}$$

Model	Migration in [%]	Migration out [%]	Correction
ISGW2	4.15	4.00	0.996
K&R	3.89	4.03	0.994
Bouchard	3.97	3.73	0.994
Flynn	3.49	4.34	1.02

Table 6.17 *Correction factors to the efficiency for migration in and out of the high q^2 region due to resolution on the reconstructed q^2 using the choice closest to the regression value. Results obtained from simulated $B_s^0 \to K^-\mu^+\nu_\mu$ events after a full selection is applied. The events have been reweighted to be consistent with predictions from Lattice QCD and LCSR. The Migration in is defined as the percentage of events with true q^2 below 7 $\mathrm{GeV}^2/\mathrm{c}^4$ and reconstructed q^2 above 7 $\mathrm{GeV}^2/\mathrm{c}^4$.*

No q^2 sel.	10.2 %
$q^2 < 7$ $\mathrm{GeV}^2/\mathrm{c}^4$	19.9 %
$q^2 > 7$ $\mathrm{GeV}^2/\mathrm{c}^4$	3.41 %

Table 6.18 *The relative uncertainty on the $B_s^0 \to K^-\mu^+\nu_\mu$ selection efficiency originating from a lack of knowledge on the q^2 distribution.*

6.5.6 Final Corrected Relative Efficiency

The efficiencies and corrections used to determine the full selection efficiency of $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ are listed in Table 6.19. The uncertainties on the corrections are taken as systematic uncertainties when calculating the final ratio of branching fractions. The corrected efficiency is plotted against the true q^2 in Figure 6.22 for $B_s^0 \to K^-\mu^+\nu_\mu$ (left) and $B_s^0 \to D_s^-\mu^+\nu_\mu$ (right). The unfortunately large bias on the efficiency of the signal mode combined with a lack of knowledge on the shape of true q^2 distribution results in the assignment of a systematic uncertainty on the final corrected efficiency. The systematic uncertainty on the final corrected efficiency originating from an uncertainty on the knowledge of the q^2 distribution is determined by calculating the corrected $B_s^0 \to K^-\mu^+\nu_\mu$ efficiency under each of the four form factor models and taking the standard deviation. As seen in Figure 2.4 the dominant factor contributing to the true q^2 distribution at low q^2 is the form factor parametrisation, while at high q^2 the dominant contribution comes from phase space. The systematic uncertainties originating from a lack of knowledge on the true q^2 distribution will therefore be greater at low q^2, and the systematic uncertainties are listed in Table 6.18. The full summary of systematic uncertainties is given in Table 6.21.

The final corrected ratio of efficiencies for $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ are

listed in Table 6.20.

6.6 Determination of $\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)$ and $|V_{ub}|/|V_{cb}|$

The ratio of branching fractions of $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ is determined by taking the ratio of signal yields at production, obtained by dividing the fit yields, N, by the selection efficiency, ε, and charm branching fraction,

$$\frac{\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} = \frac{N_{B_s^0 \to K^-\mu^+\nu_\mu}}{N_{B_s^0 \to (D_s^- \to K^+K^-\pi^-)\mu^+\nu_\mu}} \cdot \frac{\varepsilon_{B_s^0 \to (D_s^- \to K^+K^-\pi^-)\mu^+\nu_\mu}}{\varepsilon_{B_s^0 \to K^-\mu^+\nu_\mu}} \cdot \mathcal{B}(D_s^- \to K^+K^-\pi^-). \tag{6.19}$$

The ratio of branching fractions is found to be,

$$\frac{\mathcal{B}(B_s^0 \to K^-\mu^+\nu_\mu)}{\mathcal{B}(B_s^0 \to D_s^-\mu^+\nu_\mu)} = (3.59 \pm 0.34 \pm 0.51) \times 10^{-3}, \tag{6.20}$$

where the first uncertainty is statistical and the second is systematic. The uncertainty is systematics limited with the dominant uncertainty originating from a selection biased in q^2.

By performing a branching fraction measurement with a restricted q^2 of the $\mu^+\nu_\mu$ pair and combining the result with relative form factors, R_{FF}, from lattice

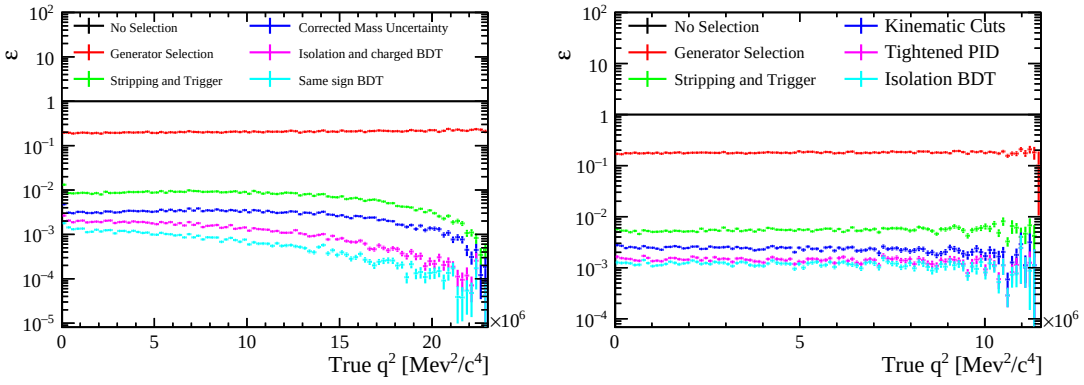


Figure 6.22 *The corrected Efficiencies for successive selections on $B_s^0 \to K^-\mu^+\nu_\mu$ (left) and $B_s^0 \to D_s^-\mu^+\nu_\mu$ (right) candidates are plotted against the true q^2.*

QCD [30, 31, 35, 36] and light-cone sum rules [32] the ratio of CKM elements $|V_{ub}|/|V_{cb}|$ is determined as in equations 4.1 and 4.2. The relative form factors are calculated as,

$$R_{FF} = \frac{\int_{q^2{}_{min}}^{q^2{}_{max}} \frac{1}{|V_{cb}|^2} \frac{d\Gamma}{dq^2}_{B_s^0 \to D_s^- \mu^+ \nu_\mu} dq^2}{\int_{q^2{}_{min}}^{q^2{}_{max}} \frac{1}{|V_{ub}|^2} \frac{d\Gamma}{dq^2}_{B_s^0 \to K^- \mu^+ \nu_\mu} dq^2}. \tag{6.21}$$

Using $B_s^0 \to K^- \mu^+ \nu_\mu$ form factors obtained from light-cone sum rules the ratio of CKM elements is found to be

$$\frac{|V_{ub}|}{|V_{cb}|} = \left(\frac{\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)_{|q^2 < 7 \ \mathrm{GeV}^2/c^4}}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)} \cdot R_{FF}^{LCSR} \right)^{1/2}$$
$$= 0.0625 \pm 0.0092(\mathrm{exp.}) \pm 0.0039(\mathrm{th.}), \tag{6.22}$$

where the first uncertainty is experimental and the second uncertainty is theoretical.

Using form factors obtained from lattice QCD the ratio of CKM elements is found to be

$$\frac{|V_{ub}|}{|V_{cb}|} = \left(\frac{\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)_{|q^2 > 7 \ \mathrm{GeV}^2/c^4}}{\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu} \cdot R_{FF}^{LQCD} \right)^{1/2}$$
$$= 0.0688 \pm 0.0061(\mathrm{exp.}) \pm 0.0086(\mathrm{th.}), \tag{6.23}$$

where the first uncertainty is experimental and the second uncertainty is theoretical. For the decay $B_s^0 \to D_s^- \mu^+ \nu_\mu$ the main sources of uncertainty originate from the statistical uncertainty in the fit and the systematic uncertainty on the calculation of the PID correction factors. For the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ the dominant uncertainty in the low q^2 bin originates from the biased efficiency in q^2 and the dominant uncertainty in the high bin originates from the systematic uncertainty in the form factor calculations. The statistical uncertainty in the fit used to extract the $B_s^0 \to K^- \mu^+ \nu_\mu$ yields is limited by the small size of the background Monte Carlo samples.

These results represent the first experimental measurement of the branching fraction $B_s^0 \to K^- \mu^+ \nu_\mu$ and ratio of $|V_{ub}|/|V_{cb}|$ using this decay. The determined values of $|V_{ub}|/|V_{cb}|$ are plotted in Figure 6.23 using lattice QCD (solid black line) and light-cone sum rules (dashed black line) alongside the inclusive and exclusive

averages of $|V_{ub}|$ and $|V_{cb}|$, the previous determination of $|V_{ub}|/|V_{cb}|$ performed by LHCb using the decay $\Lambda_b^0 \to p\mu^-\overline{\nu}_\mu$ is plotted in pink.

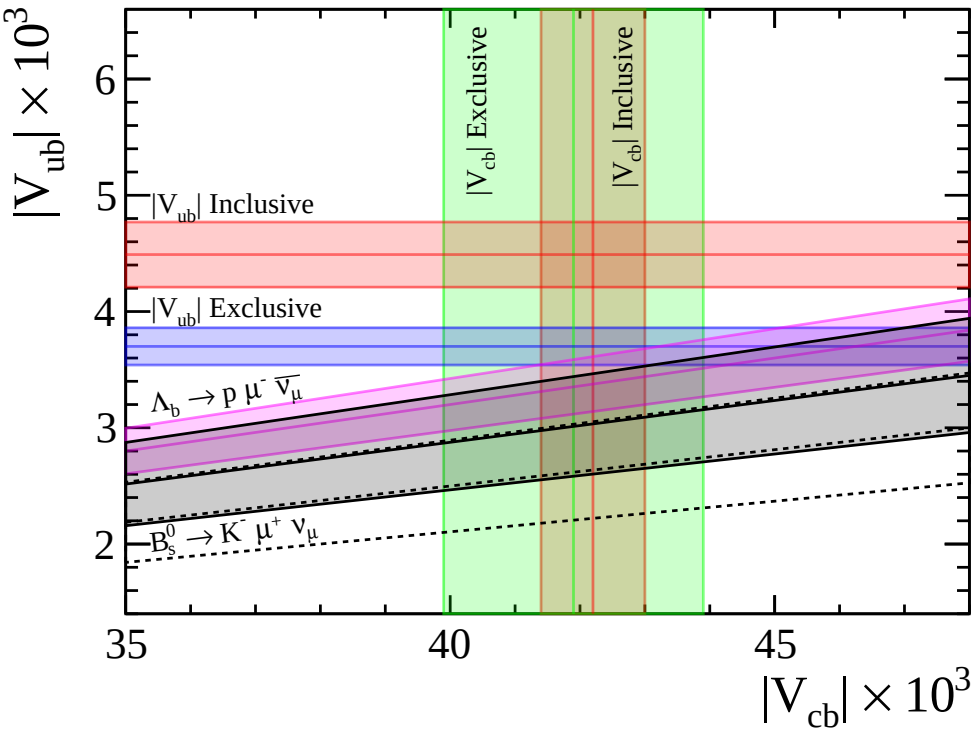


Figure 6.23 *The values for $|V_{ub}|/|V_{cb}|$ obtained using LQCD (solid line) and LCSR (dashed line) are plotted alongside the inclusive and exclusive $|V_{ub}|$ and $|V_{cb}|$ PDG averages. The previous LHCb measurement obtained using the decay $\Lambda_b^0 \to p\mu^-\overline{\nu}_\mu$ is plotted in pink.*

Source	Efficiency [%] $B_s^0 \to D_s^- \mu^+ \nu_\mu$	$B_s^0 \to K^- \mu^+ \nu_\mu$		
			$q^2 < 7\text{GeV}^2/\text{c}^4$	$q^2 > 7\text{GeV}^2/\text{c}^4$
Generator	17.87 ± 0.08	20.5 ± 0.2	19.7 ± 0.1	21.0 ± 0.2
Selection	0.62	0.422	0.504	0.217
Source	Correction			
Tracking	1.018 ± 0.004	1.007 ± 0.001	1.006 ± 0.001	1.010 ± 0.002
PID.	0.823 ± 0.085	0.855 ± 0.017	0.850 ± 0.025	0.863 ± 0.006
$\sigma_{m_{\text{corr.}}}$		1.02 ± 0.02	0.91 ± 0.14	1.026 ± 0.002
Isolation	0.989 ± 0.014	0.993 ± 0.033	0.995 ± 0.013	0.995 ± 0.013
Charged BDT		0.966 ± 0.034	0.959 ± 0.029	0.959 ± 0.029
Same sign BDT		0.995 ± 0.034	$0.948\pm$	0.948 ± 0.041
q^2 migration			1.002 ± 0.008	0.996 ± 0.009
	Corrected Efficiency [%]			
	0.109 ± 0.011	0.084 ± 0.011	0.082 ± 0.021	0.0456 ± 0.0032

Table 6.19 *Summary of efficiencies and corrections entering into the combined efficiency for the $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ modes.*

	ε_{rel}
No q^2 sel.	0.671 ± 0.056
$q^2 < 7\text{GeV}^2/\text{c}^4$	0.682 ± 0.115
$q^2 > 7\text{GeV}^2/\text{c}^4$	0.356 ± 0.027

Table 6.20 *Final corrected efficiency ratio, $\varepsilon_{B_s^0 \to K^- \mu^+ \nu_\mu} / \varepsilon_{B_s^0 \to D_s^- \mu^+ \nu_\mu}$, for the signal and normalisation channels within each region of q^2.*

Uncertainty [%]	$B_s^0 \to D_s^- \mu^+ \nu_\mu$	$B_s^0 \to K^- \mu^+ \nu_\mu$		
		No q^2 sel.	$q^2 < 7$	$q^2 > 7$
$\mathcal{B}(D_s^- \to K^- K^+ \pi^-)$	3.3			
Form factor uncertainty	3.1		9.7	22.45
Tracking	0.41	0.15	0.15	0.16
Particle Identification	10.2	2.0	3.0	0.74
m_{corr} error		2.0	2.0	2.0
Isolation	1.4	3.3	1.3	1.3
Charged BDT		3.5	3.0	3.0
Same Sign BDT		4.9	4.3	4.3
q^2 migration			0.85	0.90
ε generator	0.08	0.24	0.12	0.19
ε error from FF.		10.2	19.9	3.4
$B^+ \to J/\psi K^+$ reco.		2.1	0.61	3.8
Fit Bias	0.59	1.2	4.6	8.1
m_{corr} template		1.4	3.6	0.87

Table 6.21 *Systematic uncertainties on the evaluated yields at production for $B_s^0 \to D_s^- \mu^+ \nu_\mu$ and $B_s^0 \to K^- \mu^+ \nu_\mu$. When taking the ratio of branching fractions some of the systematic uncertainties will partially cancel, and when calculating the ratio of $|V_{ub}|/|V_{cb}|$ the uncertainties will be approximately halved.*

Chapter 7

Implications

7.1 Inclusive and exclusive determinations of $|V_{ub}|/|V_{cb}|$

The global average values of $|V_{ub}|/|V_{cb}|$ determined by the PDG [37] from exclusive and inclusive decays are

$$\begin{aligned} |V_{ub}|/|V_{cb}| &= 0.107 \pm 0.007 \quad \text{(inclusive)}, \\ |V_{ub}|/|V_{cb}| &= 0.088 \pm 0.006 \quad \text{(exclusive)}. \end{aligned} \tag{7.1}$$

The values of $|V_{ub}|/|V_{cb}|$ obtained in this thesis using semileptonic decays of the B_s^0 meson in combination with form factor predictions from lattice QCD and light cone sum rules are

$$\begin{aligned} |V_{ub}|/|V_{cb}| &= 0.072 \pm 0.010 \quad \text{(LQCD)}, \\ |V_{ub}|/|V_{cb}| &= 0.062 \pm 0.010 \quad \text{(LCSR)}. \end{aligned} \tag{7.2}$$

The results obtained in this thesis are consistent with the exclusive averages for $|V_{ub}|/|V_{cb}|$ and are slightly lower, the values differ by 1.4σ and 2.2σ when comparing the results obtained from LQCD and LCSR respectively.

The results obtained in this thesis are significantly lower than the average of inclusive $|V_{ub}|/|V_{cb}|$ measurements, the values differ by 3.0σ and 3.9σ when comparing the LQCD and LCSR results respectively. These results confirm the tension between inclusive and exclusive measurements of $|V_{ub}|$.

7.2 Outlook for $|V_{ub}|/|V_{cb}|$ from $B_s^0 \to K^-\mu^+\nu_\mu$ decays and LHCb

From an experimental point of view, the uncertainty on $|V_{ub}|/|V_{cb}|$ is dominated by the uncertainties originating from the fit extracting the $B_s^0 \to K^-\mu^+\nu_\mu$ yield. In turn the errors in the fit are dominated by the limited Monte Carlo statistics in the samples modelling inclusive $b \to c$ backgrounds, with additional large uncertainties arising from biases in the fit and constraints on background yields.

From a theoretical point of view the uncertainty on $|V_{ub}|/|V_{cb}|$ is dominated by finite volume and chiral extrapolation systematic uncertainties in the lattice calculation. Additionally the form factor predictions for $B_s^0 \to K^-\mu^+\nu_\mu$ disagree dramatically at low q^2 with the consensus from the theoretical community being that the systematic uncertainties from the lattice are underestimated at low q^2.

The analysis was performed using a dataset with a total integrated luminosity of 2 fb^{-1} collected during the year 2012 with a centre of mass energy, $\sqrt{s} =$ 8 TeV. This represents a small fraction of the total data collected by the LHCb experiment and as the systematic uncertainties are dominated by Monte Carlo statistics and theoretical uncertainties this small amount of data was of ample size.

A measurement of the branching fraction of $B_s^0 \to K^-\mu^+\nu_\mu$ in bins of q^2 should be considered, as there are large uncertainties from a theoretical perspective on the q^2 distribution of this decay, and an experimental determination of the form factors would provide the theoretical community with valuable constraints. A binned measurement would need to employ the full LHCb dataset and due to a limited resolution on the reconstructed q^2 would require a careful unfolding of the q^2 distribution, or a folding of theoretical predictions. A binned measurement would require vast amounts of simulated Monte Carlo events to correctly model various backgrounds. The amount of additional simulated data required is an order of magnitude greater than currently possessed and presents a significant challenge in the computation required for production and disk space for storage. Recent developments in simulation know as ReDecay [103], where the simulation reuses the underlying event and regenerates the candidate of interest rather than simulating a full new event, and RICHless reconstruction, where the simulations is run without modelling the Cherenkov radiation and RICH detectors provide a means of generating large amounts of simulated data with significant reductions

in compute time.

Chapter 8

Conclusions

A measurement of the ratio of CKM matrix elements $|V_{ub}|/|V_{cb}|$ provides a direct constraint on global fits to the unitary triangles and provides an important constraint when performing global fits to the unitarity of the CKM matrix. A long standing discrepancy between inclusive and exclusive measurements of $|V_{ub}|$ has puzzled both experimentalists and theorists alike, and it is unknown if this difference is due to an unknown problem with the experimental measurements, an unaccounted for systematic in the theoretical calculations of the form factors, or most excitingly the result of unexplained physics beyond the standard model. A number of proposals have been presented to explain this discrepancy including the leptoquark [104], a hypothetical particle with a simultaneous coupling to leptons and quarks, and the addition of a heavy right handed $W^{\pm}$ boson [105].

An experimental measurement of the differential branching fraction of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ provides a vital constraint for the theoretical community. Current predictions of the $B_s^0 \to K^- \mu^+ \nu_\mu$ decay rate differ by an order of magnitude at low q^2 and an experimental measurement provides a vital constraint for theoretical models.

Two measurements of $|V_{ub}|/|V_{cb}|$ were performed using data collected from the LHCb experiment, measurements of the ratios of the branching fractions $\mathcal{B}(B_s^0 \to K^- \mu^+ \nu_\mu)/\mathcal{B}(B_s^0 \to D_s^- \mu^+ \nu_\mu)$ restricted to high and low regions of q^2 were combined with form factor calculations obtained from lattice QCD and light cone sum rules. This measurement included a first observation of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$. The measurement was performed using the decay products resulting from the pp collisions with a centre of mass energy of $\sqrt{s} = 8$ TeV.

The data sample collected by the LHCb experiment during the year 2012 and used for this measurement has an integrated luminosity of 2 fb^{-1} and represents a small fraction of the total dataset collected by LHCb. The measurements of $|V_{ub}|/|V_{cb}|$ obtained in this thesis, $|V_{ub}|/|V_{cb}| = 0.072 \pm 0.010$ (LQCD) and $|V_{ub}|/|V_{cb}| = 0.062 \pm 0.010$ (LCSR) are consistent with exclusive averages calculated by the PDG and are significantly lower than the inclusive averages increasing the tension between inclusive and exclusive measurements of $|V_{ub}|$

The measurement of $|V_{ub}|/|V_{cb}|$ presented in this thesis represents a *proof of concept* analysis demonstrating the feasibility of a measurement of the differential branching fraction of the decay $B_s^0 \to K^-\mu^+\nu_\mu$. Despite using less than a quarter of the full dataset available for analysis the dominant limiting factor came from the modelling backgrounds using simulated Monte Carlo. Recent developments in the simulation of Monte Carlo events will significantly reduce these limiting factors and allow for more refined measurements of this decay. The differential branching fraction measured in this thesis using two bins in q^2 demonstrates the feasibility of performing an analysis with multiple bins. These results will be highly valuable to the theoretical community and will allow for the modelling of $B_s^0 \to K^-\mu^+\nu_\mu$ form factors to be constrained, this will provide additional constraints to the theories of lattice QCD and light cone sum rules.

Appendix A

Form Factor Comparisons

This chapter summarises the results of $B_s^0 \to K^- \mu^+ \nu_\mu$ and $B_s^0 \to D_s^- \mu^+ \nu_\mu$ form factor calculations and compares plots presented in the published papers with those generated by the analysis software using results taken from the papers. This is to ensure that results taken from theory have been reproduced accurately and that there are no errors from copying tables of numbers.

This chapter contains

- Names and references of publication used
- Fitted parameters of the z-expansions
- Selected reproductions of plots verifying analysis software

A.1 Publications Used

$B_s^0 \to K^- \mu^+ \nu_\mu$

Title	Authors	arXiv
$B_s \to Kl\nu$ form factors from lattice QCD	C.M. Bouchard, G.Peter Lepage, Christopher Monahan, Heechang Na, Junko Shigemitsu	arXiv:1406.2279v2 [31]
$B \to \pi l\nu$ and $B_s \to Kl\nu$ form factors and $\lvert V_{ub}\rvert$ from $2+1$-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks	J.M. Flynn, T. Izubuchi, T. Kawanai, C. Lehner, A. Soni, R.S. Van de Water, O. Witzel	arXiv:1501.05373v3 [30]
$B_s \to Kl\nu_l$ and $B_{(s)} \to \pi(K)l^+l^-$ decays at large recoil and CKM matrix elements	Alexander Khodjamirian, Aleksey V. Rusov	arXiv:1703.04765v2 [32]

Table A.1 *Details of the papers providing form factor results for* $B_s^0 \to K^- \mu^+ \nu_\mu$

$B_s^0 \to D_s^- \mu^+ \nu_\mu$

Title	Authors	arXiv
$B_s \to D_s l\nu$ form factors and the fragmentation fraction ratio f_s/f_d.	Christopher J. Monahan, Heechang Na, Chris M. Bouchard, G. Peter Lepage, Junko Shigemitsu	arXiv:1703.09728v1 [36]
$B_s \to D_s/B \to D$ semileptonic form-factor ratios and their application to BR($B_s^0 \to \mu^+\mu^-$)	Jon A. Bailey, A. Bazavov, C. Bernard, C.M. Bouchard, C. DeTar, Daping Du, A.X. El-Khadra, J. Foley, E.D. Freeland, E. Gamiz, Steven Gottlieb, U.M. Heller, Jongjeong Kim, A.S. Kronfeld, J. Laiho, L. Levkova, P.B. Mackenzie, Y. Meurice, E. Neil, M.B. Oktay, Si-Wei Qiu, J.N. Simone, R. Sugar, D. Toussaint, R.S. Van de Water, Ran Zhou	arXiv:1202.6346v2 [35]

Table A.2 *Details of the papers providing form factor results for* $B_s^0 \to D_s^- \mu^+ \nu_\mu$

A.2 z-expansion Fit Parameters

$B_s^0 \to K^- \mu^+ \nu_\mu$

Bouchard et.al.

	$b_1^{(0)}$	$b_2^{(0)}$	$b_3^{(0)}$	$b_1^{(+)}$	$b_2^{(+)}$	$b_3^{(+)}$
Value	0.31500	0.9450	2.3910	0.368000	-0.7500	2.7200
Error	0.12900	1.3050	4.6710	0.021400	0.1930	1.4580
$b_1^{(0)}$	0.01676	0.1462	0.4453	0.001165	0.0214	0.1434
$b_2^{(0)}$	0.14620	1.7020	5.8520	0.009481	0.2255	1.5390
$b_3^{(0)}$	0.44530	5.8520	21.810	0.029630	0.7472	5.3250
$b_1^{(+)}$	0.00117	0.0095	0.0296	0.000458	0.0012	-0.0013
$b_2^{(+)}$	0.02140	0.2255	0.7472	0.001157	0.0372	0.1858
$b_3^{(+)}$	0.14340	1.5390	5.3250	-0.001309	0.1858	2.1240

Table A.3 *Extrapolated coefficients of a HPChPT z expansion for the $B_s^0 \to K^- \mu^+ \nu_\mu$ form factors with the associated covariance matrix. Results taken from [31].*

Flynn et.al.

	$b_{(+)}^0$	$b_{(+)}^1$	$b_{(+)}^2$	$b_{(0)}^0$	$b_{(0)}^1$	$b_{(0)}^2$
Value	0.338	-1.161	-0.458	0.210	-0.169	-1.235
Error	0.024	0.192	1.009	0.024	0.202	0.880
$b_{(+)}^0$	1.000	0.255	0.146	0.873	0.603	0.423
$b_{(+)}^1$	0.255	1.000	0.823	0.311	0.954	0.770
$b_{(+)}^2$	0.146	0.823	1.000	0.346	1.060	0.901
$b_{(0)}^0$	0.873	0.311	0.346	1.000	0.556	0.479
$b_{(0)}^1$	0.603	0.954	1.060	0.556	1.000	0.965
$b_{(0)}^2$	0.423	0.770	0.901	0.479	0.965	1.000

Table A.4 *Central values, errors, and correlation matrix for the BCL z-parametrisations of f_+ and f_0 for $B_s^0 \to K^- \mu^+ \nu_\mu$. Results taken from [30].*

Khodjamirian and Rusov

	$f_{BP}(0)$	$b_{1(BP)}$	Correlation
f_+	0.336(23)	-2.53(1.17)	0.79
f_0	0.320(19)	-1.08(1.53)	0.74

Table A.5 *Central values, errors, and correlations for the BCL z-parametrisations of f_+ and f_0 for $B_s^0 \to K^- \mu^+ \nu_\mu$. Results taken from [32].*

$B_s^0 \to D_s^- \mu^+ \nu_\mu$

Monahan et.al.

	$a_0^{(0)}$	$a_1^{(0)}$	$a_2^{(0)}$	$a_0^{(+)}$	$a_1^{(+)}$	$a_2^{(+)}$
Value	0.663	-0.10	1.3	0.868	-3.35	0.6
Error	0.031	0.30	2.8	0.032	0.41	4.7
$a_0^{(0)}$	0.0009534	-0.00303547	-0.00542391	0.000594503	0.00158251	0.0160091
$a_1^{(0)}$	0.00303547	0.0903097	-0.101760	0.000446248	0.0236283	0.0456659
$a_2^{(0)}$	0.00542391	-0.101760	8.02283	0.00848079	0.104246	0.760797
$a_0^{(+)}$	0.000594503	0.000446248	0.00848079	0.00100761	-0.00423358	-0.0264511
$a_1^{(+)}$	0.00158251	0.0236283	0.104246	-0.00423358	0.165251	-0.617234
$a_2^{(+)}$	0.0160091	0.0456659	0.760797	-0.0264511	-0.617234	22.49292

Table A.6 *Central values, errors, and covariance matrix for the z-parametrisations of f_+ and f_0 for $B_s^0 \to D_s^- \mu^+ \nu_\mu$. Results taken from [36].*

Bailey et.al.

	$a_0^{(+)}$	$a_1^{(+)}$	$a_2^{(+)}$	$a_0^{(0)}$	$a_1^{(0)}$	$a_2^{(0)}$
Value	0.01191	-0.111	0.47	0.01081	-0.0662	0.18
Error	0.00006	0.002	0.05	0.00004	0.0002	0.06
$a_0^{(+)}$	1.0	-0.055	-0.002	0.593	0.254	0.014
$a_1^{(+)}$	-0.055	1.0	-0.318	-0.067	0.867	-0.180
$a_2^{(+)}$	-0.002	-0.318	1.0	-0.038	-0.307	0.974
$a_0^{(-)}$	0.593	-0.067	-0.038	1.000	-0.050	-0.054
$a_1^{(-)}$	0.254	0.867	-0.307	-0.050	1.000	-0.233
$a_2^{(-)}$	0.014	-0.180	0.974	-0.054	-0.233	1.000

Table A.7 *Central values, errors, and correlation matrix for the three term z-parametrisations of f_+ and f_0 for $B_s^0 \to D_s^- \mu^+ \nu_\mu$. Results taken from [35].*

A.3 Comparison Plots

$B_s^0 \to K^- \mu^+ \nu_\mu$

Bouchard et.al.

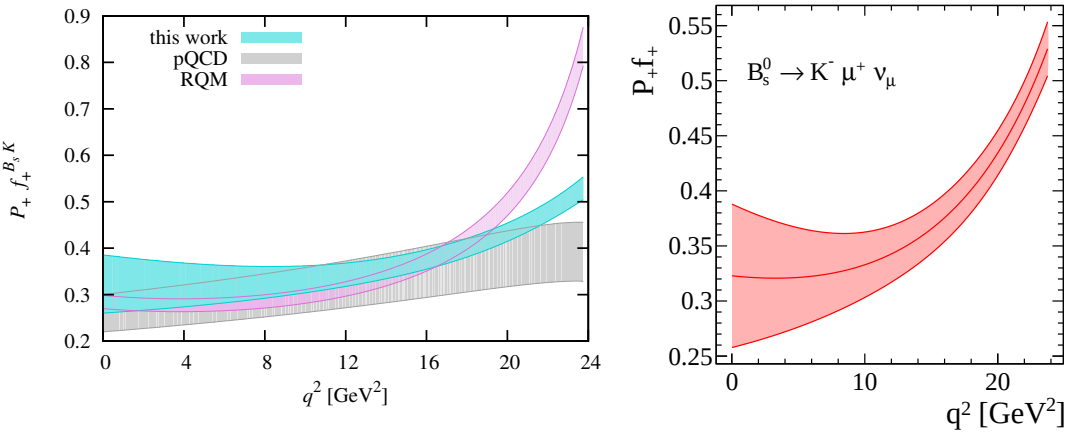


Figure A.1 *Form factors plotted against z. Image, left, taken from [31] and right, generated using fit parameters taken from [31]. The blue shaded section (left) should be compared to the red section (right).*

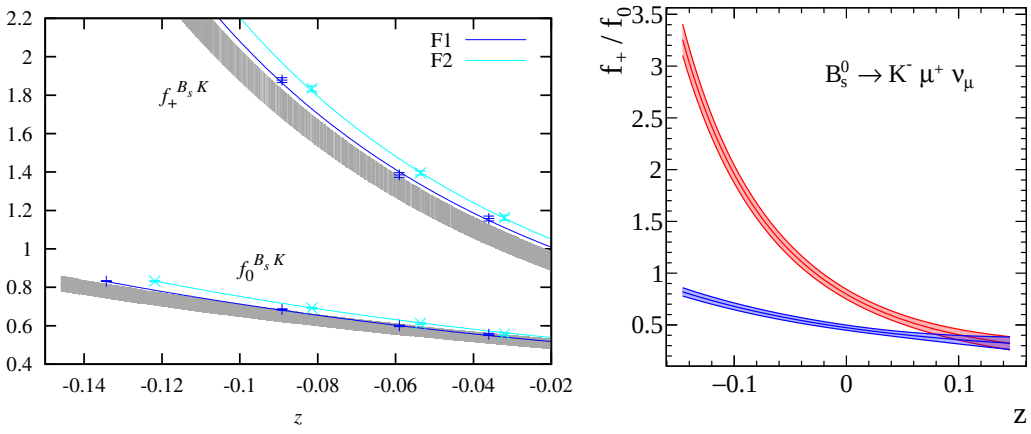


Figure A.2 *Form factors plotted against z. Image, left, taken from [31] and right, generated using fit parameters taken from [31].*

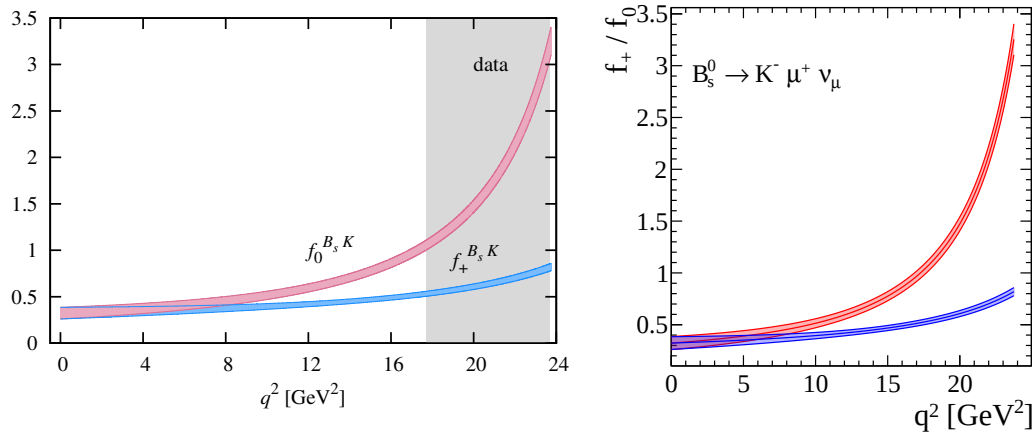


Figure A.3 *Form factors plotted against q^2. Image, left, taken from [31] and right, generated using fit parameters taken from [31].*

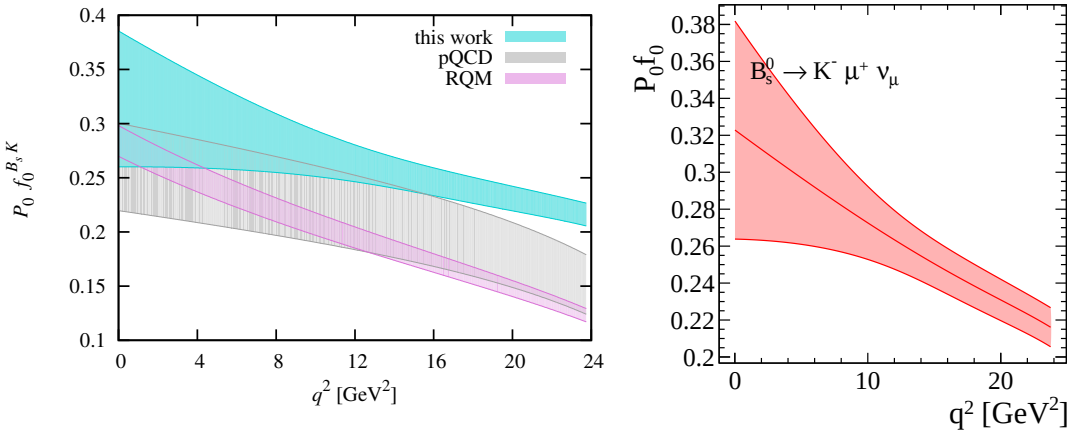


Figure A.4 *Form factors plotted against q^2. Image, left, taken from [31] and right, generated using fit parameters taken from [31]. The blue shaded section (left) should be compared to the red section (right).*

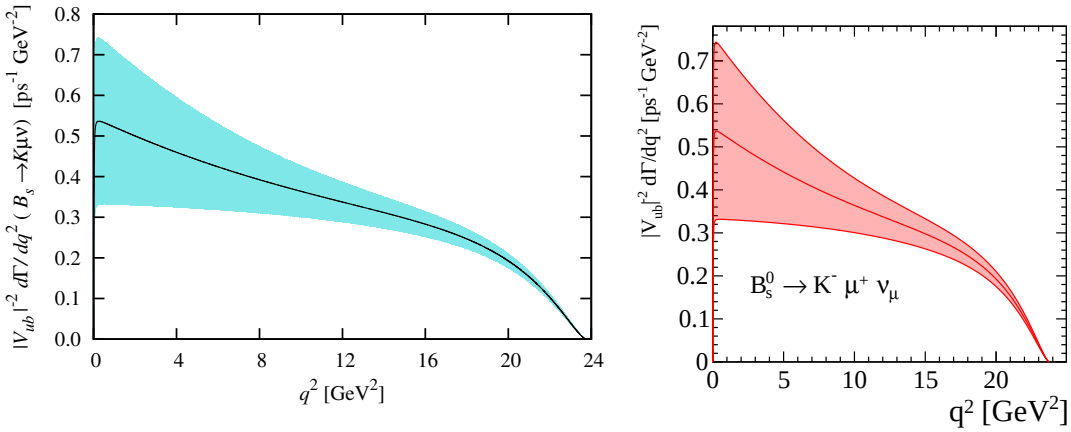


Figure A.5 *The differential $B_s^0 \to K^-\mu^+\nu_\mu$ decay rate plotted against q^2. Image, left, taken from [31] and right, generated using fit parameters taken from [31].*

Flynn et.al.

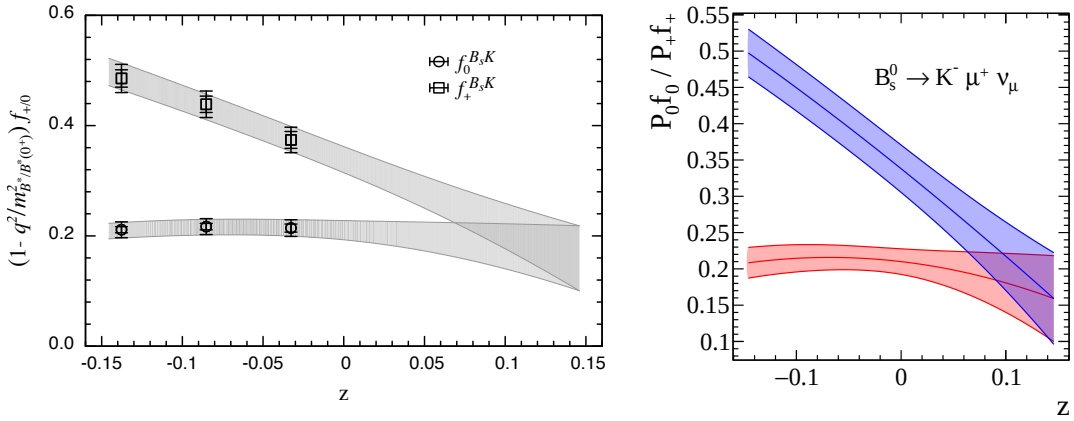


Figure A.6 *Form factors plotted against z. Image, left, taken from [30] and right, generated using fit parameters taken from [30].*

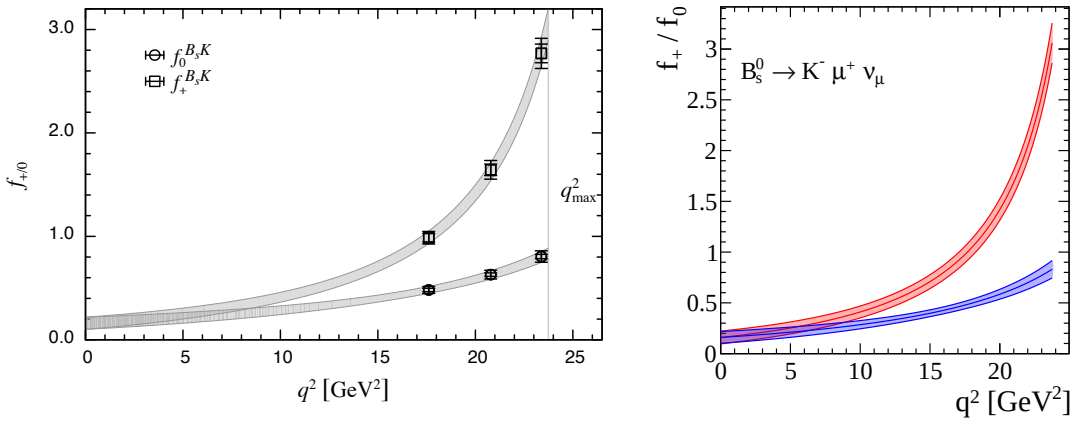


Figure A.7 *Form factors plotted against q^2. Image, left, taken from [30] and right, generated using fit parameters taken from [30].*

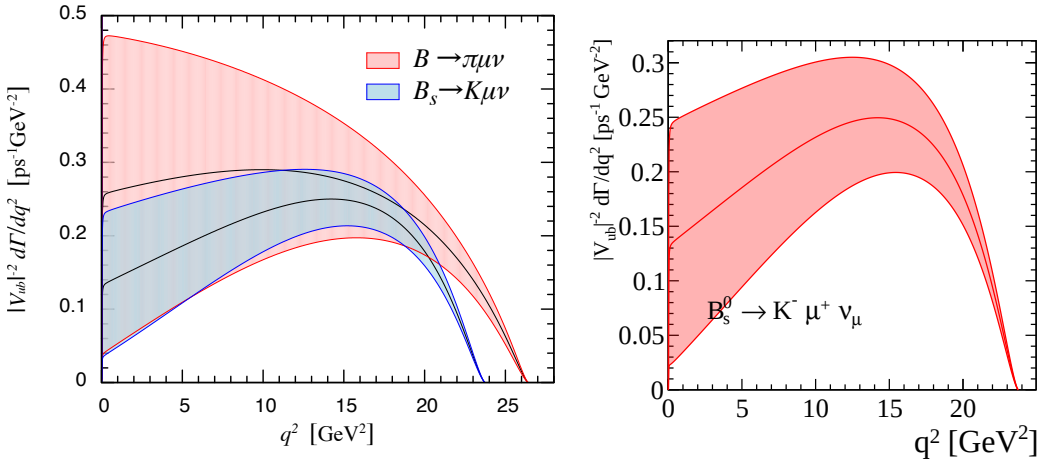


Figure A.8 *The differential $B_s^0 \to K^-\mu^+\nu_\mu$ decay rate plotted against q^2. Image, left, taken from [30] and right, generated using fit parameters taken from [30].*

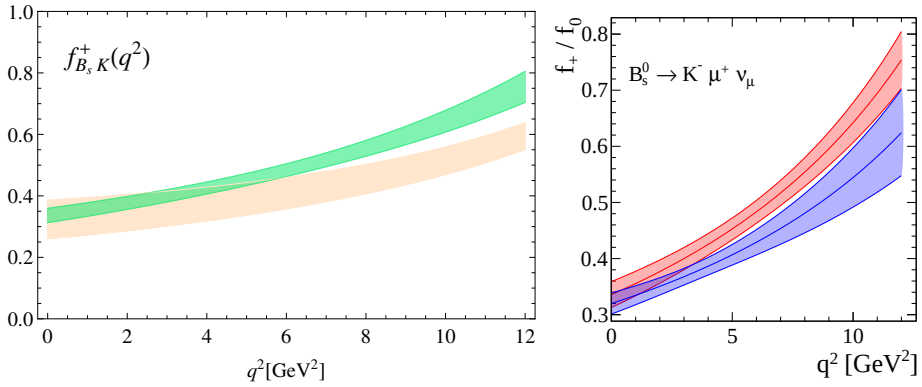


Figure A.9 *Form factors plotted against q^2. Image, left, taken from [32] and right, generated using fit parameters taken from [32]. The green shaded region (left) should be compared to the red shaded region (right).*

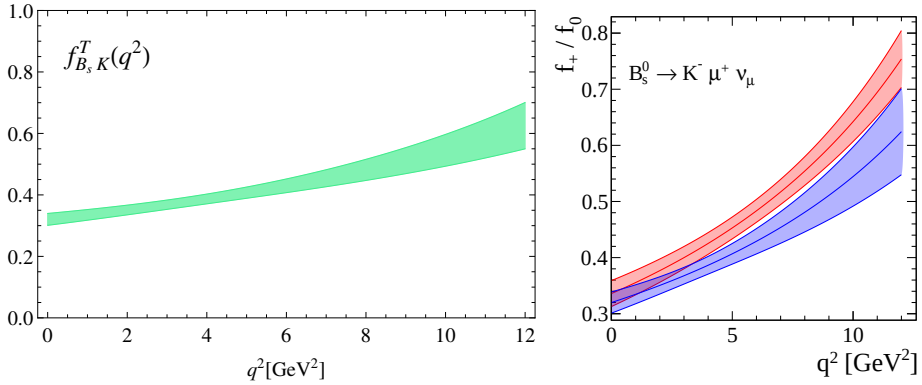


Figure A.10 *Form factors plotted against q^2. Image, left, taken from [32] and right, generated using fit parameters taken from [32]. The green shaded region (left) should be compared to the blue shaded region (right).*

$B_s^0 \to D_s^- \mu^+ \nu_\mu$

Monahan et.al.

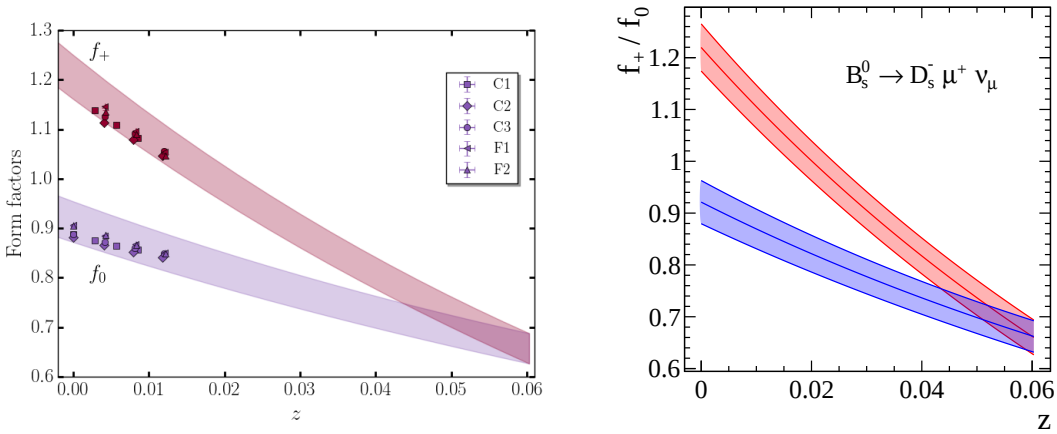


Figure A.11 *Form factors plotted against z. Image, left, taken from [36] and right, generated using fit parameters taken from [36].*

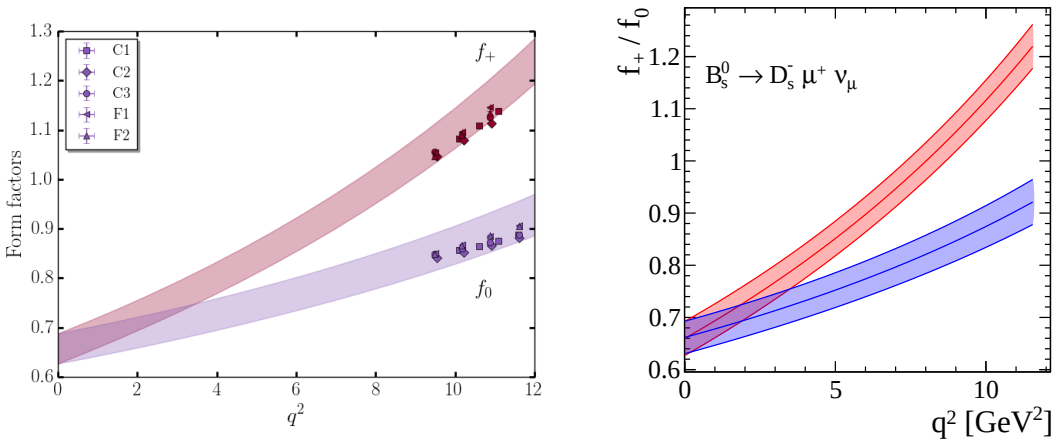


Figure A.12 *Form factors plotted against q^2. Image, left, taken from [36] and right, generated using fit parameters taken from [36].*

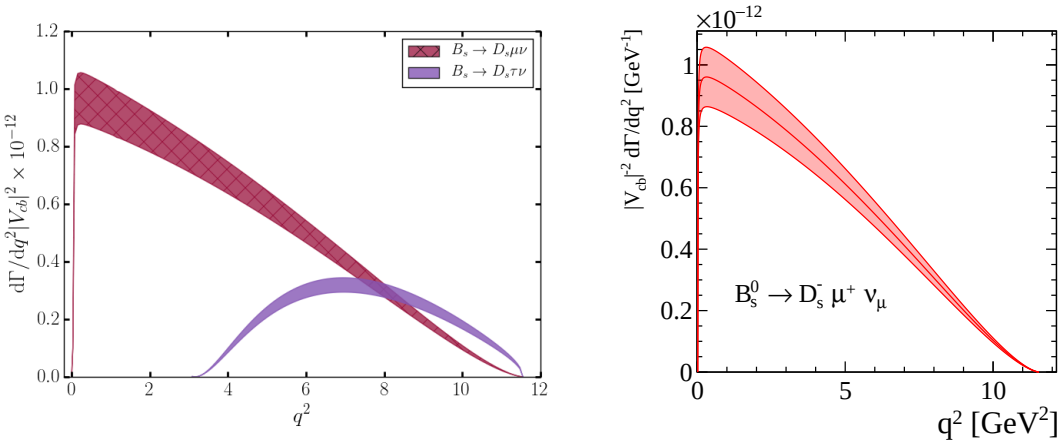


Figure A.13 *The differential $B_s^0 \to D_s^- \mu^+ \nu_\mu$ decay rate plotted against q^2. Image, left, taken from [36] and right, generated using fit parameters taken from [36].*

Bailey et.al.

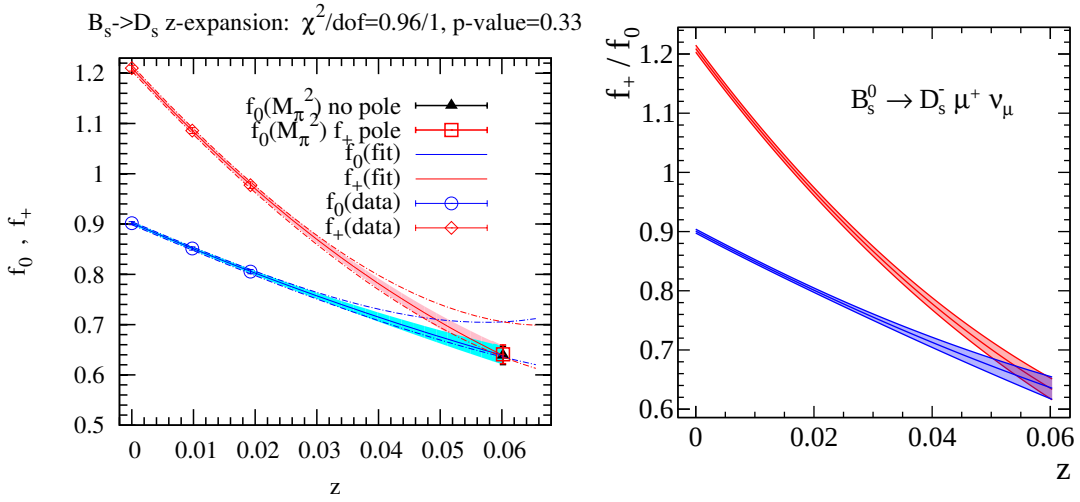


Figure A.14 *Form factors plotted against z. Image, left, taken from [35] and right, generated using fit parameters taken from [35].*

Appendix B

Validation of Combinatoric Modelling

This appendix contains additional plots validation the modelling of combinatoric samples. Figure B.1 contains kinemtic distributions of true combinatoric events from data (solid black line), modelled events (blue points) and modelled events after a kinematic reweighting (red points). All plots are restricted to the kinematic region $m_{K^-\mu^+} > 5400$ MeV/c^2.

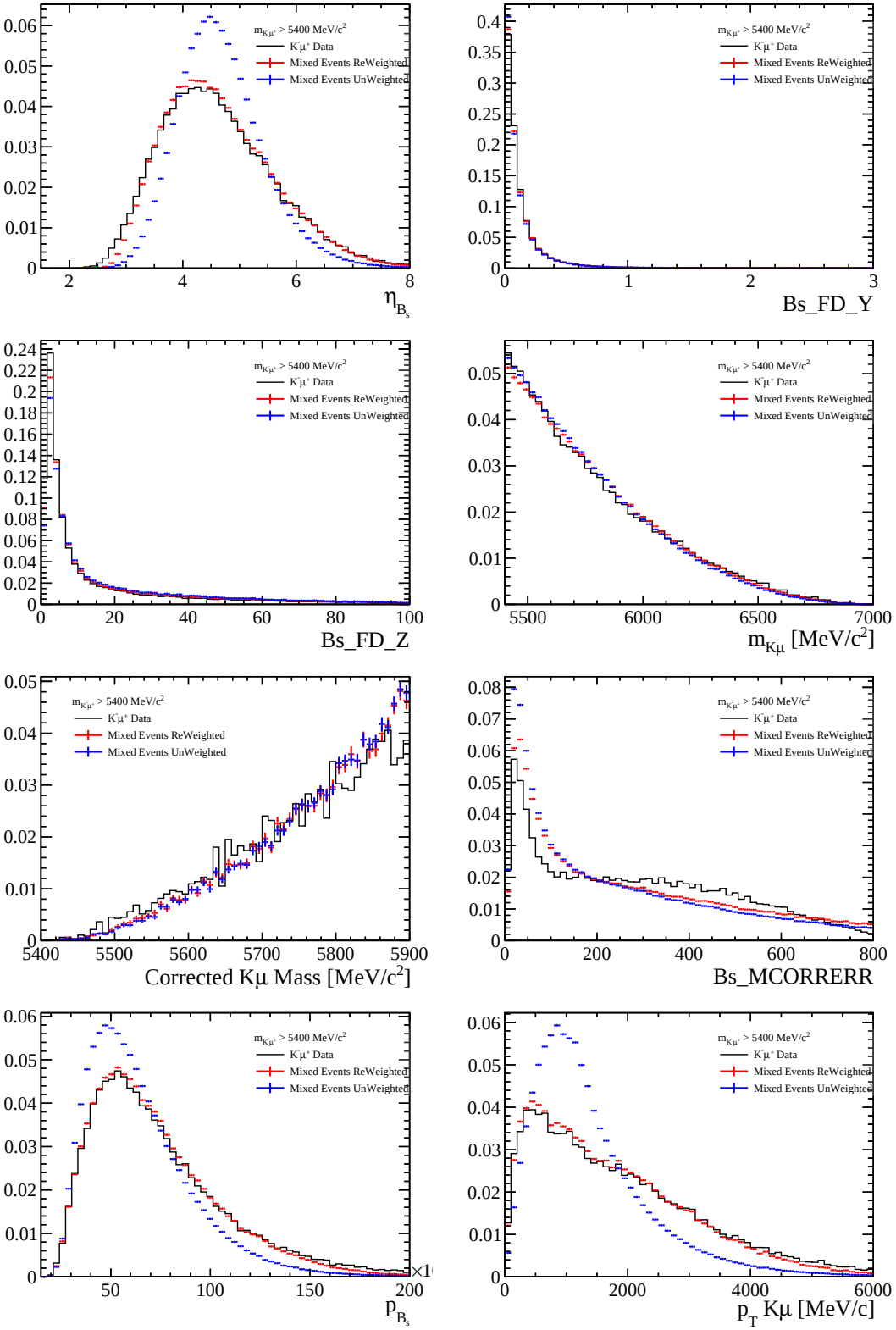


Figure B.1 *$K^-\mu^+$ candidates in data are plotted with simulated combinatorics before and after a kinematic correction*

Appendix C

sPlot Background subtraction Results

This appendix contains the fit results used as inputs to the *sPlot* background subtraction tabulated in Table C.1 and plotted in Figures C.1- C.3. One *sPlot* background subtraction is performed on the $D_s^- \to K^-K^+\pi^-$ invariant mass peak and two background subtractions are performed on the $B^+ \to J/\psi\, K^+$ invariant mass obtained by reconstructing the three body final state and by reconstructing a $K^-\mu^+$ final state with the additional muon found via isolation.

	$D_s^+ \to K^-K^+\pi^+$	$B^+ \to K^+\mu^+\mu^-$	$B^+ \to K^+\mu^-$ Iso(μ^+)
$\text{Yield}_{\text{Sig.}}$	$(683.0 \pm 3.5) \times 10^3$	$(125.6 \pm 0.4) \times 10^3$	$(22.8 \pm 0.4) \times 10^3$
$\text{Yield}_{\text{BG.}}$	$(902.3 \pm 3.5) \times 10^3$	$(9.3 \pm 0.3 \times 10^3$	$(10.3 \pm 0.4) \times 10^3$
μ [MeV/c^2]	1969.7 ± 1.0	5283.84 ± 0.06	5288.8 ± 0.2
σ_1 [MeV/c^2]	5.92 ± 0.09	15.7 ± 0.3	16.9 ± 0.5
σ_2 [MeV/c^2]	12.5 ± 0.5	25.3 ± 0.7	33.9 ± 0.46
f_1	0.198 ± 0.0098	0.68 ± 0.34	0.14 ± 0.09
f_2	1.30 ± 0.013	0.46 ± 0.26	0.053 ± 0.034
τ [MeV^{-1}c^2]	$(1.82 \pm 0.05) \times 10^{-3}$	$(-8.09 \pm 0.23) \times 10^{-3}$	$(-2.3 \pm 0.2) \times 10^{-3}$

Table C.1 *Fit results obtained from a maximum likelihood fit in order to obatin sWeights.*

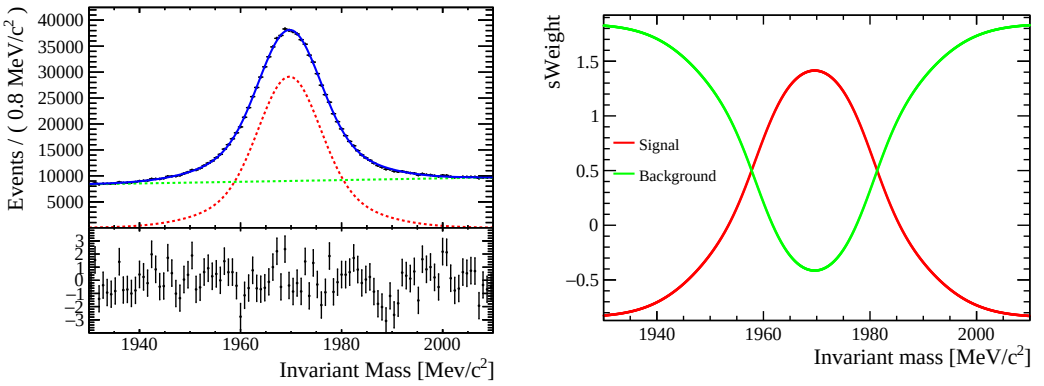


Figure C.1 *Fit to $K^-K^+\pi^+$ invariant mass spectrum and sWeights obtained from fit.*

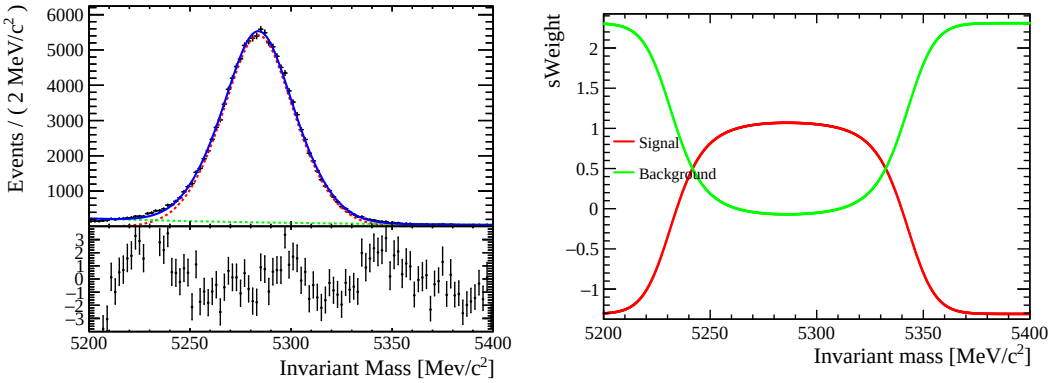


Figure C.2 *Fit to $K^-\mu^+\mu^-$ invariant mass spectrum and sWeights obtained from fit.*

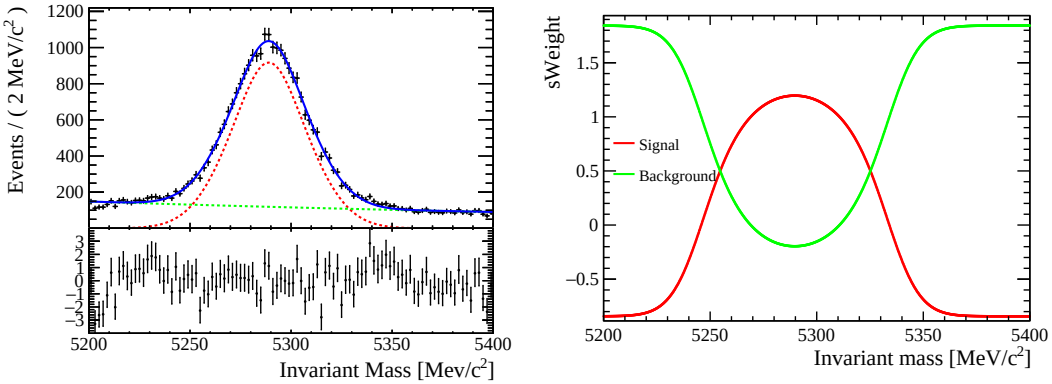


Figure C.3 *Fit to $K^-\mu^+\mu^-$ invariant mass spectrum and sWeights obtained from fit.*

Appendix D

Validation of BDT Reweighting

This appendix contains plots validating the use of a BDT to simultaneously correct multiple Monte Carlo distributions. A $k = 2$ k-factor cross validation method is used with data separated by magnet polarity, i.e. the MagUp data is used to correct MagDown data. The BDT response variables and correction weights for both polarities is plotted in Figure D.1 for the correcting of $B_s^0 \to K^-\mu^+\nu_\mu$ Monte Carlo using $B^+ \to J/\psi K^+$ decays. The BDT response and correction weights for $B_s^0 \to D_s^-\mu^+\nu_\mu$ are plotted in Figure D.2. A selection of kinematic distributions for Data, corrected and uncorrected Monte Carlo are plotted in Figure D.3 for both $B_s^0 \to K^-\mu^+\nu_\mu$ and $B_s^0 \to D_s^-\mu^+\nu_\mu$ demonstrating the effectiveness of this method.

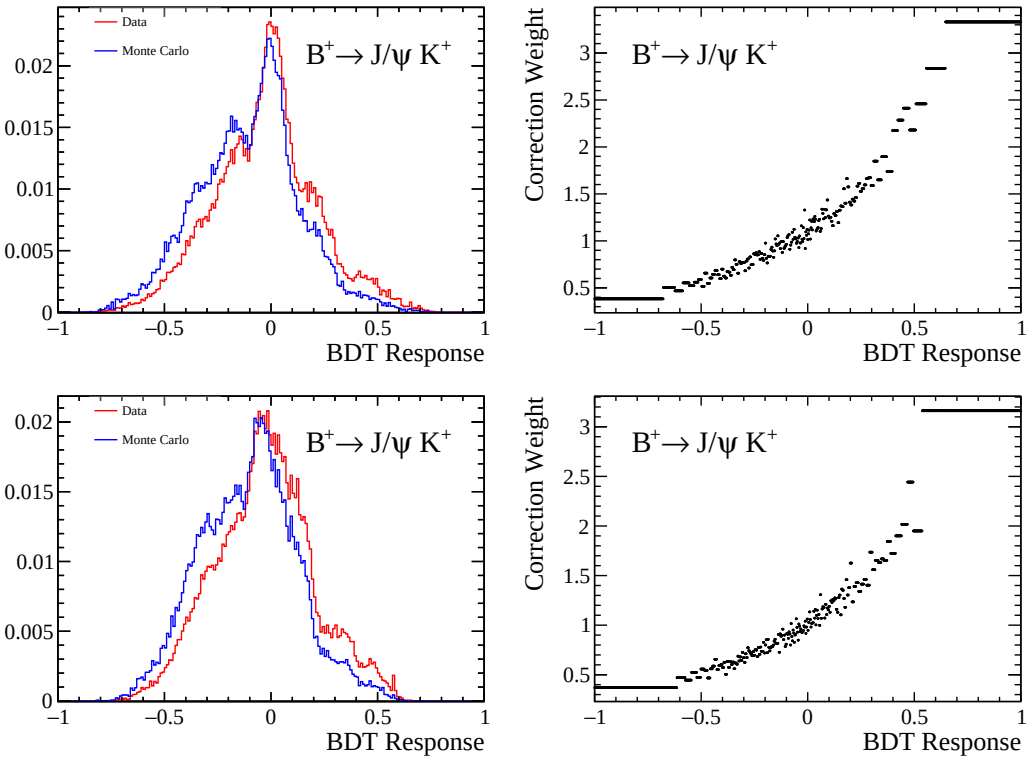


Figure D.1 *The BDT Response and weights when using $B^+ \to J/\psi K^+$ reconstructed as $B_s^0 \to K^-\mu^+$. Trained using MagUp and used to correct MagDown (top) and vice versa (bottom)*

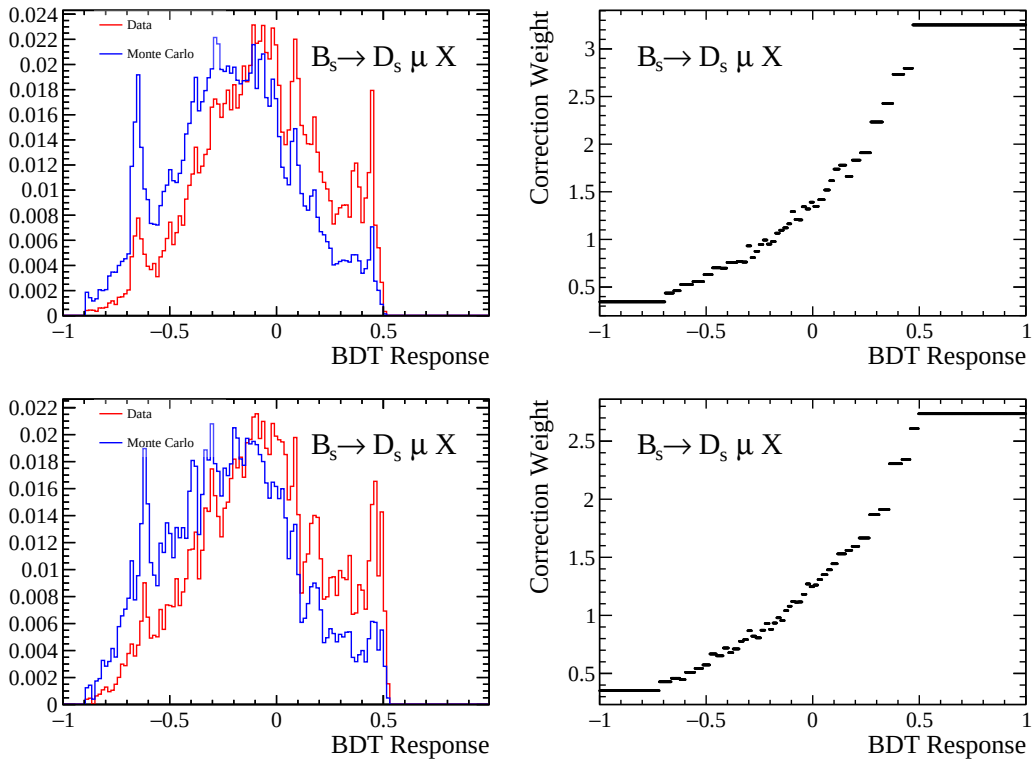


Figure D.2 *The BDT Response and weights when using* $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$. *Trained using MagUp and used to correct MagDown (top) and vice versa (bottom)*

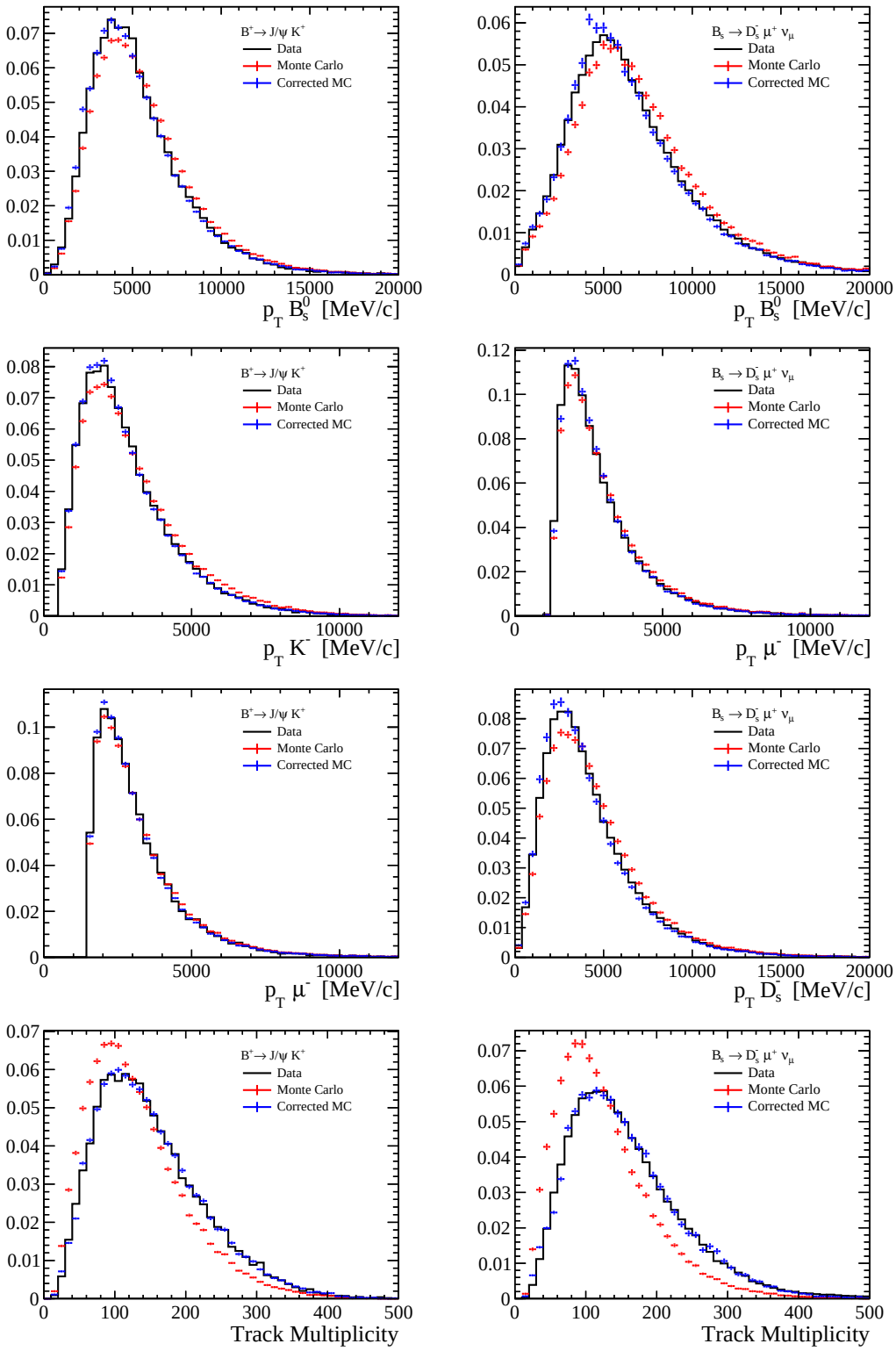


Figure D.3 *The kinematic distributions of all variables corrected using the BDT ReWeighter method.*

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[4] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156.

[5] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, [,160(1964)].

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585, [,162(1964)].

[7] T. W. B. Kibble, *Symmetry breaking in nonAbelian gauge theories*, Phys. Rev. **155** (1967) 1554, [,165(1967)].

[8] ATLAS, 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.

[9] CMS, 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.

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

[11] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[12] M. Gell-Mann and M. Lévy, *The axial vector current in beta decay*, Il Nuovo Cimento (1955-1965) **16** (1960) 705.

[13] LHCb, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv:1611.03076.

[14] R. Aaij *et al.*, *Measurement of the ckm angle from a combination of bdh analyses*, Physics Letters B **726** (2013), no. 1 151 .

[15] R. Aaij *et al.*, *Measurement of cp observables in bd()k and bd() decays*, Physics Letters B **777** (2018) 16 .

[16] R. Aaij *et al.*, *Measurement of cp observables in bdk and bd with two- and four-body d decays*, Physics Letters B **760** (2016) 117 .

[17] UTfit, M. Bona, *Latest results for the Unitary Triangle fit from the UTfit Collaboration*, PoS **CKM2016** (2017) 096.

[18] UTfit Collaboration, M. Bona, *updated results and plots available at*, `http://www.utfit.org/UTfit/`.

[19] UTfit, M. Bona *et al.*, *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Delta m(s) and BR(B -¿ tau nu(tau))*, JHEP **10** (2006) 081, `arXiv:hep-ph/0606167`.

[20] CKMfitter Group, M. Pivk, *CP violation and the CKM matrix: Impact of the asymmetric B factories*, in *Proceedings, 32nd International Conference on High Energy Physics (ICHEP 2004): Beijing, China, August 16-22, 2004*, pp. 909–912, 2004. `arXiv:hep-ph/0410173`.

[21] CKMfitter Group, J. C. et al. *updated results and plots available at*, `http://ckmfitter.in2p3.fr`.

[22] S. Fajfer, J. F. Kamenik, and I. Nišandžić, $b \to D^* \tau \overline{\nu}_\tau$ *sensitivity to new physics*, Phys. Rev. D **85** (2012) 094025.

[23] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (1974) 2445.

[24] H. J. Rothe, *Lattice gauge theories: An Introduction*, World Sci. Lect. Notes Phys. **43** (1992) 1, [World Sci. Lect. Notes Phys.82,1(2012)].

[25] C. Gattringer and C. B. Lang, *Quantum chromodynamics on the lattice*, Lect. Notes Phys. **788** (2010) 1.

[26] P. B. Mackenzie, *An Improved Hybrid Monte Carlo Method*, Phys. Lett. **B226** (1989) 369.

[27] C. A. Dominguez, *Introduction to QCD sum rules*, Mod. Phys. Lett. **A28** (2013) 1360002, `arXiv:1305.7047`.

[28] D. Bigi and P. Gambino, *Revisiting* $B \to D\ell\nu$, Phys. Rev. **D94** (2016), no. 9 094008, `arXiv:1606.08030`.

[29] C. Bourrely, I. Caprini, and L. Lellouch, *Model-independent description of B —¿ pi l nu decays and a determination of —V(ub)—*, Phys. Rev. **D79** (2009) 013008, `arXiv:0807.2722`, [Erratum: Phys. Rev.D82,099902(2010)].

[30] J. M. Flynn *et al.*, *$B \to \pi\ell\nu$ and $B_s \to K\ell\nu$ form factors and $|V_{ub}|$ from 2+1-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks*, Phys. Rev. **D91** (2015), no. 7 074510, arXiv:1501.05373.

[31] C. M. Bouchard *et al.*, *$B_s \to K\ell\nu$ form factors from lattice QCD*, Phys. Rev. **D90** (2014) 054506, arXiv:1406.2279.

[32] A. Khodjamirian and A. V. Rusov, *$B_s \to K\ell\nu_\ell$ and $B_{(s)} \to \pi(K)\ell^+\ell^-$ decays at large recoil and CKM matrix elements*, JHEP **08** (2017) 112, arXiv:1703.04765.

[33] C. M. Bouchard. Private Communication, 2018.

[34] C. Bourrely, L. Lellouch, and I. Caprini, *Model-independent description of $b \to \pi l\nu$ decays and a determination of $|V_{ub}|$*, Phys. Rev. D **79** (2009) 013008.

[35] J. A. Bailey *et al.*, *$B_s \to D_s/B \to D$ Semileptonic Form-Factor Ratios and Their Application to $BR(B_s^0 \to \mu^+\mu^-)$*, Phys. Rev. **D85** (2012) 114502, arXiv:1202.6346, [Erratum: Phys. Rev.D86,039904(2012)].

[36] C. J. Monahan *et al.*, *$B_s \to D_s\ell\nu$ Form Factors and the Fragmentation Fraction Ratio f_s/f_d*, Phys. Rev. **D95** (2017), no. 11 114506, arXiv:1703.09728.

[37] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.

[38] L. Rossi, *The LHC Superconducting Magnets*, .

[39] A. Morris, M. Needham, and F. Muheim, *Measurements of charmless B_s^0 meson decays at LHCb.*, Aug, 2017. Presented 29 Sep 2017.

[40] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross section in 7 and 13 tev pp collisions*, Phys. Rev. Lett. **118** (2017) 052002.

[41] A. Abashian *et al.*, *The Belle Detector*, Nucl. Instrum. Meth. **A479** (2002) 117.

[42] BaBar, B. Aubert *et al.*, *The BaBar detector*, Nucl. Instrum. Meth. **A479** (2002) 1, arXiv:hep-ex/0105044.

[43] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. **C75** (2015), no. 5 204, arXiv:1412.3989.

[44] A. Breskin and R. Voss, *The CERN Large Hadron Collider: Accelerator and Experiments*, CERN, Geneva, 2009.

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

[46] P. Koppenburg, *Simulation of the vertex trigger preprocessor: effects of noise on L1 performance*, Tech. Rep. LHCb-99-003, CERN, Geneva, Feb, 1999.

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

[48] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[49] M. Needham and D. Volyanskyy, *Updated geometry description for the LHCb Trigger Tracker*, Tech. Rep. LHCb-2006-032. CERN-LHCb-2006-032, CERN, Geneva, Jun, 2006.

[50] A. Perrin and K. Vervink, *The Inner Tracker detector description and its implementation in the XML database*, Tech. Rep. LHCb-2006-018. CERN-LHCb-2006-018. LPHE-Note-2006-06, CERN, Geneva, May, 2006.

[51] LHCb Collaboration, *LHCb outer tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.

[52] LHCb, *LHCb: RICH technical design report*, .

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

[54] LHCb, E. Picatoste Olloqui, *LHCb preshower(PS) and scintillating pad detector (SPD): Commissioning, calibration, and monitoring*, J. Phys. Conf. Ser. **160** (2009) 012046.

[55] O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, .

[56] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[57] LHCb HLT project, J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv:1310.8544.

[58] R. Cornat, J. Lecoq, and P. Perret, *Level-0 decision unit for LHCb*, Tech. Rep. LHCb-2003-065, CERN, Geneva, Sep, 2003.

[59] M. Williams *et al.*, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.

[60] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014.

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

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

[63] G. Corti *et al.*, *Software for the LHCb experiment*, IEEE Trans. Nucl. Sci. **53** (2006) 1323.

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

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

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

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

[68] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[69] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[70] The LHCb Collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 tev pp collisions*, Phys. Rev. D **85** (2012) 032008.

[71] LHCb, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, Nature Phys. **11** (2015) 743, arXiv:1504.01568.

[72] W. L. Sutcliffe and U. Egede, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, Sep, 2016. Presented 19 Jul 2016.

[73] S. Dambach, U. Langenegger, and A. Starodumov, *Neutrino reconstruction with topological information*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **569** (2006), no. 3 824 .

[74] S. Stone and L. Zhang, *Method of Studying Λ_b^0 decays with one missing particle*, Adv. High Energy Phys. **2014** (2014) 931257, arXiv:1402.4205.

[75] A. C. Rencher and W. F. Christensen, *Methods of multivariate analysis*, Wiley, 2012.

[76] D. A. Freedman, *Statistical Models: Theory and Practice*, Cambridge University Press, 2009.

[77] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, Wiley, 2013.

[78] G. Ciezarek, A. Lupato, M. Rotondo, and M. Vesterinen, *Reconstruction of semileptonically decaying beauty hadrons produced in high energy pp collisions*, JHEP **02** (2017) 021, arXiv:1611.08522.

[79] F. Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python*, ArXiv e-prints (2012) arXiv:1201.0490.

[80] K. Govorkova, *Study of photons and neutral pions reconstruction efficiency in the LHCb experiment*, Master's thesis, Moscow State U., 2015-12-22.

[81] D. Tadić and S. Žganec, *Covariant generalization of the isgur-scora-grinstein-wise quark model*, Phys. Rev. D **52** (1995) 6466.

[82] R. D. Kenway, *The Isgur-Wise function*, Nucl. Phys. Proc. Suppl. **34** (1994) 153, arXiv:hep-lat/9312021.

[83] A. Puig, *The LHCb trigger in 2011 and 2012*, Tech. Rep. LHCb-PUB-2014-046. CERN-LHCb-PUB-2014-046, CERN, Geneva, Nov, 2014.

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

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

[86] L. Breiman, *Classification and regression trees*, Chapman Hall/CRC, 2005.

[87] J. R. Quinlan, *Induction of decision trees*, Machine Learning **1** (1986) 81.

[88] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, Springer, 2017.

[89] R. Kohavi, *A study of cross-validation and bootstrap for accuracy estimation and model selection*, pp. 1137–1143, Morgan Kaufmann, 1995.

[90] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016), no. 1 012036, arXiv:1608.05806.

[91] LHCb, B. Adeva *et al.*, *Roadmap for selected key measurements of LHCb*, arXiv:0912.4179.

[92] L. Gavardi, M. Calvi, and J. Albrecht, *Search for lepton flavour violation in τ decays at the LHCb experiment*, Nov, 2013. Presented 28 Nov 2013.

[93] LHCb, R. Aaij *et al.*, *A precise measurement of the B^0 meson oscillation frequency*, Eur. Phys. J. **C76** (2016), no. 7 412, arXiv:1604.03475.

[94] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[95] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, J. Comput. Syst. Sci. **55** (1997), no. 1 119.

[96] R. Barlow and C. Beeston, *Fitting using finite monte carlo samples*, Computer Physics Communications **77** (1993), no. 2 219 .

[97] ROOT Collaboration, K. Cranmer *et al.*, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016, New York U., New York, Jan, 2012.

[98] J. E. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'*, PhD thesis, SLAC, 1982.

[99] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.

[100] LHCb, B. Storaci, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, .

[101] ATLAS, G. Aad *et al.*, *Determination of the ratio of b-quark fragmentation fractions f_s/f_d in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. Lett. **115** (2015), no. 26 262001, arXiv:1507.08925.

[102] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, arXiv:1408.1251.

[103] D. Muller and B. G. Siddi, *Fast simulation options in LHCb from ReDecay to fully parametrised*, Feb, 2017.

[104] M. Bauer and M. Neubert, *Minimal Leptoquark Explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_g$ Anomalies*, Phys. Rev. Lett. **116** (2016), no. 14 141802, arXiv:1511.01900.

[105] P. S. B. Dev, R. N. Mohapatra, and Y. Zhang, *Heavy right-handed neutrino dark matter in left-right models*, Mod. Phys. Lett. **A32** (2017) 1740007, arXiv:1610.05738.

[106] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, Comput. Phys. Commun. **210** (2017) 103, arXiv:1601.05437.

[107] D. Binosi, J. Collins, C. Kaufhold, and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams. Version 2.0 release notes*, Comput. Phys. Commun. **180** (2009) 1709, arXiv:0811.4113.