

DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and Mathematics

of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

put forward by

Svende Annelies Braun

born in Giessen

Oral examination: June 16th, 2020

**Measurement of the CKM matrix
elements $|V_{ub}|/|V_{cb}|$ from semileptonic B_s decays**

Referees:

Prof. Dr. Stephanie Hansmann-Menzemer

Apl. Prof. Dr. Monica Dunford

Abstract

This thesis presents the determination of the CKM matrix elements $|V_{ub}|/|V_{cb}|$ from measuring the ratio of branching fractions of the B_s^0 meson decays $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ with respect to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. The analysis is based on proton-proton collisions corresponding to an integrated luminosity of 2 fb^{-1} produced at a center of mass energy of $\sqrt{s} = 8 \text{ TeV}$ provided by the Large Hadron Collider and collected by the LHCb experiment in 2012. The branching fraction ratio is measured to be

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} = (6.079 \pm 0.201 \pm 0.476) \times 10^{-3}$$

where the first uncertainty is statistical and the second one systematic. This experimental measured ratio is combined with theoretical form factor predictions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ which results in the determination of the CKM matrix elements $|V_{ub}|/|V_{cb}| = 0.114 \pm 0.005 \pm 0.013$ where the first uncertainty is the combined experimental uncertainty and the second one is from theoretical predictions. This measurement provides an important constraint to global fits to the unitary triangle of the CKM sector of the SM. The branching fraction ratio together with theoretical predictions and external input can also be used to extract the total branching fraction for the signal decay, which is measured to be $\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = (1.31 \pm 0.14) \times 10^{-4}$ for the first time.

Zusammenfassung

In dieser Arbeit wird die Messung der CKM-Matrix Elemente $|V_{ub}|/|V_{cb}|$ aus dem Quotient des Verzweungsverhältnis von den B_s^0 -Mesonen Zerfällen $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ im Vergleich zu $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ vorgestellt. Die Analyse basiert auf einem Proton-Proton Kollisions-Datensatz der einer integrierten Luminosität von 2 fb^{-1} entspricht und bei einer Schwerpunktsenergie von $\sqrt{s} = 8 \text{ TeV}$ vom Large Hadron Collider produziert und vom LHCb Experiment in 2012 aufgenommen wurde. Der Quotient des Verzweungsverhältnis wird zu

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} = (6.079 \pm 0.201 \pm 0.476) \times 10^{-3}$$

bestimmt, wobei die erste Unsicherheit statistisch und die zweite systematisch ist. Dieser experimentell gemessene Quotient wird mit theoretisch vorhergesagten Form Faktoren kombiniert, was zur Bestimmung der CKM Matrix Elemente $|V_{ub}|/|V_{cb}| = 0.114 \pm 0.005 \pm 0.013$ führt, hierbei ist die erste Unsicherheit die kombinierte experimentelle Unsicherheit und die zweite kommt aus theoretischen vorhersagen. Diese Messung gibt eine wichtige Beschränkung für globale Fits des Unitaritäts Dreiecks vom CKM Sektor im Standardmodell. Der Quotient des Verzweungsverhältnisses kann auch verwendet werden um das totale Verzweungsverhältnisses für den Signal Kanal zu bestimmen, dieser wird zum ersten Mal zu $\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = (1.31 \pm 0.14) \times 10^{-4}$ gemessen.

Contents

1	Introduction	11
2	Theoretical background	15
2.1	The Standard Model	15
2.2	Flavour in the SM	19
2.2.1	CKM matrix parametrisations	20
2.2.2	The unitarity triangle	22
2.2.3	Constraining the CKM matrix	24
2.3	Semileptonic B -meson decays	25
2.4	Lattice QCD	29
2.5	Theoretical form factors predictions	30
2.5.1	Form factor calculations for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	31
2.5.2	Form factor calculations for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	33
2.5.3	Form factor results	34
3	The LHCb Experiment	37
3.1	The Large Hadron Collider	37
3.2	The LHCb Experiment	37
3.2.1	Tracking system	38
3.2.2	Particle identification	42
3.2.3	Trigger system and data flow	46
3.2.4	Simulation and data samples	48
4	The Analysis strategy	51
5	Methods and kinematics	55
5.1	Corrected mass	55
5.2	Neutrino reconstruction	58
5.2.1	Linear regression and q^2	60
5.3	$sPlot$ technique	61
5.4	Boosted Decision Tree	61
6	Analysis selection	65
6.1	Prerequisites	65
6.2	Trigger selection	66
6.3	Preselection	67
6.4	Offline selection	68
6.4.1	Background vetoes for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	70
6.4.2	Additional selection cuts for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	71
6.4.3	Charged track isolation	73
6.4.4	Corrected mass uncertainty	75
6.4.5	Summary of the offline selection	77
6.5	Selection BDT	77
6.6	Selection on data	81

7	Control channel	85
8	Fit to the normalisation channel	89
8.1	Beeston-Barlow method	89
8.2	HistFactory and Beeston-Barlow light method	90
8.3	Normalisation fit model	91
8.4	Background subtraction	93
8.5	Fit results	93
9	Fit to the signal channel	97
9.1	Signal fit model	97
9.2	Combined MC templates	98
9.3	Combinatorial background	100
9.4	Misidentified background	102
9.5	Fit results	106
10	Relative efficiencies and corrections	111
10.1	Generator efficiency	111
10.2	Selection efficiency	113
10.3	Trigger efficiency	115
10.4	Particle identification	118
10.5	Tracking correction	121
10.6	Migration in between q^2 bins	122
10.7	Final corrected relative efficiency	123
10.8	Systematic uncertainties	125
11	Results	131
11.1	Measurement of the branching fraction ratio	131
11.2	Measurement of CKM elements $ V_{ub} / V_{cb} $	132
11.3	Extraction of $\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)$	135
11.4	Implications of the measurement	136
11.5	Discussion and outlook	136
12	Conclusions	139
A	Form factor comparisons	142
A.1	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ publications	142
A.1.1	HPQCD parameters	142
A.1.2	RBC-UKQCD parameters	145
A.1.3	Fermilab/MILC parameters	147
A.1.4	LCSR parameters	149
A.2	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ publications	151
A.2.1	HPQCD parameters	151
A.2.2	Fermilab/MILC parameters	153
A.2.3	HPQCD 2019 parameters	154
B	Validation of selection BDT	157

C	Validation of reweighting BDT	160
D	Validation of event mixing	162
E	TisTos efficiency checks	165
	References	172

1 Introduction

Particle physics has aims to understand our universe at the level of its fundamental constituents, the elementary particles. These, together with their fundamental interactions, are embedded into the Standard Model (SM) of particle physics. The SM is a successful theory describing most of the physical phenomena observed so far, from the lowest energy scale up to high energy physics of the order of several TeV center-of-mass energy at the Large Hadron Collider (LHC) at CERN. The particle content of the SM is complete with the recent discovery of the Higgs boson [1,2], but it also has clear shortcomings e.g. the gravitational interaction is not included.

In the SM the fundamental particles quarks and leptons come in three generations as doublets of up and down-type quarks or charged leptons and their corresponding neutrinos. The only difference between those generations are the masses of their particles generated by their couplings to the Higgs field. The quark (lepton) masses vary by five (three) orders of magnitude. This hierarchical structure across generations is still to be understood as well as the origin of exactly three generations, which are other shortcomings of the SM.

The only way for quarks and leptons to change their flavour within the SM is through the charged weak interaction mediated by the W^\pm boson, a process first discovered through the radioactive beta decay of the neutron $n \rightarrow pe^-\bar{\nu}_e$, governed by the weak $d \rightarrow u$ transition from the neutron (udd) to the proton (uud). The weak force couples charged leptons to their neutrinos, respectively. The coupling is universal for all three generations of leptons. In contrast in the quark sector couplings across generations are possible. Their couplings are flavour dependent and specified by the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3, 4]. This matrix is almost diagonal with the smallest and least well known matrix element $|V_{ub}|$ of the order of $\mathcal{O}(0.001)$ and a fractional uncertainty of around 9%.

The CKM matrix is unitary, which provides a crucial test of the SM. It can be parametrised by three real mixing angles and one complex phase. The latter allowing for CP violation in the SM. However, the observed matter-antimatter asymmetry in the present universe is nine orders of magnitude smaller than what can be explained by CP violation in the quark sector. It is therefore important to determine the parameters of the CKM matrix, such as the magnitudes of the matrix elements, to test for the unitarity of the CKM matrix and to precisely settle the amount of CP violation in the quark sector. Specifically when performing global fits to over-constrain the four parameters of the CKM matrix the large uncertainty on $|V_{ub}|$ is one of the limiting factors and thus a reduced uncertainty will lead to a better global precision of unitarity tests of the CKM matrix. Any deviations from unitarity would be an indication of new physics beyond the SM.

The CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ can be determined from inclusive and exclusive semileptonic B meson decays as well as from purely leptonic decays of B hadrons. In exclusive measurements a specific meson is reconstructed. For inclusive decays, $B \rightarrow X_{u/c}\ell^-\bar{\nu}_\ell$, a sum over all possible hadronic final states is

performed. Such measurements are performed at e^+e^- colliders by BaBar and Belle experiments using B^0 and B^+ mesons as well as by LEP experiments. A discrepancy between inclusive and exclusive measurements of approximately three standard deviations is observed, a long-standing puzzle in flavour physics.

A different approach is used to extract the ratio $|V_{ub}|/|V_{cb}|$ for the first time at a hadron collider by LHCb using exclusive semileptonic Λ_b -baryon decays [5]. This thesis uses a similar approach to extract the ratio $|V_{ub}|/|V_{cb}|$ from exclusive semileptonic B_s^0 decays. The decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ is used which suffers from much higher background contamination compared to Λ_b decay and the main complication of this analysis comes from the development of tight selection cuts which suppress this background as much as possible. The branching fraction ratio of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ is measured with respect to the normalisation decay $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$, where the D_s^- is reconstructed from $K^+ K^- \pi^-$. The measurement is made using pp collision events produced by the Large Hadron Collider (LHC) which are collected by the LHCb experiment in 2012. This ratio of branching fractions can be combined with theoretical input from lattice QCD (LQCD) [6] and light-cone QCD sum rules (LCSR) [7] to precisely determine the ratio of CKM elements $|V_{ub}|/|V_{cb}|$ providing an important constraint for global fits testing the CKM unitarity. In addition a measurement of the branching ratio of the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay is provided for the first time.

This thesis is structured as follows. An overview about the theoretical framework and a motivation of the measurement is given in Chapter 2. Chapter 3 presents a general description of the experimental setup and data taking conditions. The measurement strategy is briefly outlined in Chapter 4. The methods and tools used to perform the analysis such as the corrected mass and neutrino reconstruction are discussed in Chapter 5. The selection of data for this analysis is given in Chapter 6, and details of the control channel which is used as a high statistic channel to correct for simulation and data differences are provided in Chapter 7. The fit to the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ is performed in Chapter 8 and that of the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decay in Chapter 9. The relative efficiencies and correction factors to extract the branching fraction ratio are calculated in Chapter 10 together with the systematic uncertainties. The final results are presented and discussed in Chapter 11 which leads to the conclusion of this thesis in Chapter 12.

Besides the analysis presented in this thesis, the author contributed to two other projects at the beginning of her PhD. One was to work on track monitoring algorithms for the start-up of Run II at the LHC. During her first one and a half years she worked on the first Run II data analysis to extract the J/ψ cross section at 13 TeV, which was the first published LHCb Run II result [8] aiming to test for its production mechanism and to validate the optimised software trigger concept introduced in Run II.

The content of this thesis is part of an official LHCb analysis that is expected to be published within this year. This analysis is documented in detail in an internal note [9] which is in working group review at the moment. For such

an complex analysis it is common that a team of several people contribute as a collaborative effort and not every part of it can be done by a single person. This is also the case for the analysis presented here which was developed together with other LHCb collaborators. The author of this thesis contributed to the selection of the signal channel $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$, was responsible for the simulation of Monte Carlo samples, determined the efficiencies and developed the signal fit. It should be noted that the selection of the signal and the fit evolved during the review process such that the official analysis note deviates from the thesis and such this thesis presents an autonomous study. Also the results presented here are obtained by the author alone, those are not yet approved by the collaboration. In addition there is an already published PhD thesis from one of the LHCb collaborators performed on the same analysis which is published in this Reference [10] and based on an earlier stage of the analysis work including different selection cuts.

2 Theoretical background

This chapter introduces the basic theoretical concepts needed to extract $|V_{ub}|/|V_{cb}|$ from semileptonic B_s^0 decays. It starts with a brief overview of the Standard Model of particle physics and continues with a detailed description of Cabibbo-Kobayashi-Maskawa mechanism. Then a general introduction to semileptonic B -meson decays is given which puts the decay of interest into perspective. It ends with a description of the underlying theoretical parametrisations of so-called form factors from both Lattice QCD and Light-Cone Sum Rules, where the most recent predictions for the decay rates of interest $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ are discussed.

2.1 The Standard Model

In this section a short summary of the Standard Model of particles is given which is meant only as an overview of the fundamental particles and forces it describes, a more detailed description can be found for example in Reference [11].

The Standard Model (SM) of particle physics is a theoretical framework, putting together all elementary particles and their fundamental interactions, namely the strong, weak and electromagnetic interaction in a framework of a renormalisable quantum field theory. Only the gravitational force is excluded in the SM, but compared to the other forces it is very small when dealing with elementary particles and can therefore be neglected.

The **elementary particles** which make up the visible matter in our universe are called fermions with an intrinsic angular momentum or spin of $1/2$. These are described in terms of fields and can be further divided into two categories: *quarks* and *leptons*. Quarks interact via the strong interaction whereas leptons do not interact strongly. They both come in three generations with identical quantum numbers and increasing mass as shown in Table 1. Each quark generation consists of a so-called up- (u, c, t) and down-type (d, s, b) quarks, with the elementary electric charge of $+\frac{2}{3}e$ and $-\frac{1}{3}e$, respectively, as well as one charged lepton with charge $-1e$ and a corresponding neutral neutrino of different flavours (e, μ, τ).

The SM also describes the **fundamental interactions** between the previously introduced elementary particles. These interactions arise from the gauge invariance of the SM Lagrangian, which is invariant under local gauge transformations of the symmetry group

$$SU(3)_C \times SU(2)_L \times U(1)_Y, \quad (1)$$

where the subscripts C, L and Y stand for colour, left-handed chirality and weak hypercharge, respectively. Those are the conserved charges of the respective interaction coming from the different symmetries according to Noether's theorem [13]. The gauge group also uniquely determines the number of gauge bosons as the generators of the group, which are spin 1 particles. Those carry charges based on the underlying symmetry and as such only couple to alike particles. Due to that they mediate the three fundamental interactions, which are the electromagnetic,

Generation	Quarks			Leptons		
	Type	Mass	El. Charge	Type	Mass	El. Charge
1st	u	$2.2 \text{ MeV}/c^2$	$+2/3$	e	$511.0 \text{ keV}/c^2$	-1
	d	$4.7 \text{ MeV}/c^2$	$-1/3$	ν_e	$<2 \text{ eV}/c^2$	0
2nd	c	$1.275 \text{ GeV}/c^2$	$+2/3$	μ	$106 \text{ MeV}/c^2$	-1
	s	$95 \text{ MeV}/c^2$	$-1/3$	ν_e	$<0.19 \text{ eV}/c^2$	0
3rd	t	$173.5 \text{ GeV}/c^2$	$+2/3$	e	$1.777 \text{ GeV}/c^2$	-1
	b	$4.18 \text{ GeV}/c^2$	$-1/3$	ν_e	$<18.2 \text{ eV}/c^2$	0

Table 1: The fermionic content of the SM. The masses are taken from [12] and the electric charges are given in units of the elementary charge e .

Interaction	Particle	Spin	Mass	El. Charge
electromagnetic	photon (γ)	1	0	0
weak	W^\pm	1	$80.4 \text{ GeV}/c^2$	± 1
	Z^0	1	$91.2 \text{ GeV}/c^2$	0
strong	gluon (g)	1	0	0
	Higgs	0	$125.18 \text{ GeV}/c^2$	0

Table 2: The bosonic content of the SM. The masses are taken from [12] and the electric charges are given in units of the elementary charge e .

the weak and the strong force.

Quantum Chromo Dynamics (QCD) describes the strong interaction which is generated by the $SU(3)_C$ group and mediated by 8 massless gluons carrying so-called colour charges. These colours are charges of the strong force and come in three types red, green and blue as well as their corresponding anti-colours. Besides gluons also quarks carry colour charges, these are therefore the only two fundamental particles taking part in the strong interaction. Gluons couple to quarks via their colour charges as well as to themselves, this self-interaction leads to a short range of the strong interaction and makes it very hard to calculate. Two remarkable properties of QCD are asymptotic freedom and confinement. The first describes the running of the strong coupling constant α_s , which is a function of the transferred four-momentum squared q^2 between the interacting particles and as such dependent on the probed energy scale. The coupling constant increases with decreasing energy such that perturbative calculations are not possible in this region, only at high momentum scales above a certain scale of $\Lambda_{QCD} \sim 200 \text{ MeV}$. There the distance between coloured objects is small and quarks behave as quasi-free particles without interactions. Confinement comes from the property of the QCD potential, where the strong force between two coloured objects at large distances does not decrease but rather stays constant. As a result neither quark or anti-quark exist as isolated particles, they rather always form colour-neutral objects which are called hadrons. Traditionally, those can be either categorized as mesons made of a pair of quark and anti-quark or baryons as a bound system of three quarks or three anti-quarks. However QCD also allows for more exotic

combinations like multi quark states made off four or five quarks where the latter has been recently observed by LHCb as a pentaquark candidate [14].

The *Electroweak interaction* unifies the electromagnetic and the weak interaction and is generated by the $SU(2)_L \times U(1)_Y$ symmetry group. Its corresponding gauge bosons are three bosons of weak isospin from $SU(2)$, $W_\mu^{1,2,3}$, and one boson of weak hypercharge coming from $U(1)$, B_μ . These bosons are massless and the charges to which they couple are the weak isospin and the hypercharge. Fermions only carry weak isospin of 1/2 if they are left-handed, right-handed fermions do not have a weak isospin and therefore do not couple to the $SU(2)_L$ part of the electroweak interaction. Due to that $W_\mu^{1,2,3}$ bosons couple only to left-handed particles and right-handed anti-particles. In contrast to that the B_μ boson couples to all fermions of the SM via the weak hypercharge.

The spontaneous symmetry breaking of the electroweak gauge symmetry is realised by the Higgs mechanism [15, 16] in the SM. This mechanism generates fermion masses as well as it mixes the massless gauge bosons of the electroweak interaction into three massive bosons W^+ , W^- and Z^0 of the weak interaction and the massless photon γ coupling to the electric charge. Since W^+ and W^- are a superposition of two W_μ^i bosons, they inherit their chirality conditions and couple only to left-handed particles and right-handed anti-particles. As the neutral massive gauge boson Z^0 is a linear combination of the $U(1)$ gauge boson B_μ , it couples to all particles independent of chirality. The large masses of the W^\pm and Z^0 boson of around $80 \text{ GeV}/c^2$ and $91 \text{ GeV}/c^2$, respectively, set a limit on the range of the weak interaction of 10^{-18} m , whereas there is no limit for the electromagnetic force due to the massless photon. The Higgs mechanism also predicts a massive spin-0 particle, the so-called Higgs boson, which was discovered in 2012 by ATLAS and CMS collaborations [1, 2] and such completes the bosonic content of the SM as summarised in Table 2.

Beyond the Standard Model

Even though the SM is very successful in describing particle physics, it has some clear shortcomings and is therefore not considered as the final theory in this field. It is incomplete in the sense that it can not explain the fundamental physical concept of gravity as well as the observed *neutrino oscillations* [17, 18], which indicates that neutrinos are massive particles and is not included in the SM. Neutrino oscillations can be described in terms of the mixing matrix in the lepton sector, the so-called Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [19, 20], which is similar to the CKM matrix for quarks but the origin of neutrino masses remains an open question. Heavy right-handed sterile neutrinos have been postulated as a solution, but they haven't been observed so far [21]. Those would not participate in the SM interaction and could also be a possible dark matter candidate.

Also experimental cosmological observations from weak gravitational lensing [22] or modified galactic rotation curves [23, 24] have shown that the SM explains only

5% of the energy density of the universe. The remaining part consist of 68% of so-called *dark matter* and 27% *dark energy*. The SM does not provide fundamental particles which are candidates for either of those. Many extensions of the SM include suitable candidates which are stable on cosmological time scales and have very weak couplings to SM particles, the most prominent ones are so-called weakly interacting massive particles (WIMPS) from supersymmetric (SUSY) models [25]. Another shortcoming of the SM is the observed *matter-antimatter asymmetry* of the order of $\sim 10^{-11} - 10^{-10}$ [26, 27] in the universe today can not be explained by the small CP violating phase included in the SM and requires additional sources from new physics contributions. The CP violating phase will be explained in more detail in the next section and CP transformations are simultaneous transformations of the charge conjugation (C) and parity (P). The former changes the sign of all particle charges, while the latter transformation flips the signs of the spatial coordinates of the system ($\vec{x} \rightarrow -\vec{x}$). Many models extending the SM exist which predict new sources of CP violation, therefore precisely measuring CP violation in the SM is also a powerful tool to search for new physics.

On the theoretical side an open question of the SM is the so-called *hierarchy problem*. Within the SM no symmetry protects the mass of the Higgs boson to its measured value such that large quantum corrections from virtual particles can lead to much higher masses. The experimentally observed Higgs boson mass can only be accommodated with a fine tuning that cancels these quantum corrections. This level of fine-tuning is considered as unnatural by many theorists and several models were proposed in order to solve this issue: SUSY [28], little Higgs [29] or grand unified theories (GUT) [30].

An open question is also why there are exactly *three generations* of fermions and what determines the large spread of fermion masses across many orders of magnitudes. Therefore theorists aim to develop a more fundamental theory that contains mechanism or symmetries which explain these observations with fewer input parameters.

Due to all of these shortcomings experimental particle physics is searching for fundamental particles or interactions which are not part of the SM, so-called New Physics contributions. There are different approaches to search for particles beyond the SM.

In *direct searches* they are searched for in the production of high energy collisions at particles colliders to explore the energy frontier or they are detected via their interaction with ordinary baryonic matter in sensitive detectors. The latter method is explored in dark matter experiments such as the XENON1T experiment [31] as an example which uses large detector volumes below the surface and try to detect possible rare interaction of dark matter particles with the target nuclei. The energy frontier is explored by the ATLAS [32] and CMS [33] experiments at CERN which are looking for signatures of heavy stable particles which are produced in high energy proton-proton collisions. Here the maximum mass of a new particle to be discovered is limited by the available center-of-mass energy of the *pp* collision, which is currently 14 TeV.

A complementary approach is the *indirect search* for new particles which is

followed by the LHCb experiment. Here indirect means that those new particles might be too heavy to be produced as real particles but they can enter in quantum loops as virtual particles. These quantum corrections to SM processes are smaller the heavier the particles involved are, which is why the key ingredient for this type of searches is the precision. Those corrections can then either modify existing SM processes or allow processes which are forbidden in the SM. Therefore in order to perform an indirect search one needs precise theoretical prediction of the observable of interest in the context of the SM to be able to compare the measurement to it as well as an observable that gets sizeable contributions from these quantum loops which can be precisely measured with the experiment. Flavour physics uses indirect searches to be able to probe higher new physics scales by increasing precision of the measurements.

2.2 Flavour in the SM

In the SM quarks and leptons acquire mass through the Yukawa interactions with the Higgs field after the spontaneous symmetry breaking. The leptonic part of the Lagrangian will be neglected since it is not relevant for the processes discussed in this thesis. The Yukawa term for quark fields can be written as:

$$\mathcal{L}_{Yukawa}^{quarks} = -\frac{v}{\sqrt{2}}(\bar{d}_L Y_d d_R + \bar{u}_L Y_u u_R) + h.c., \quad (2)$$

where left-handed quarks \bar{d}_L are coupled to right-handed ones d_R by the Yukawa matrices Y_u, Y_d as the coupling constants. Here v represents the vacuum expectation value of the Higgs potential after the electroweak symmetry breaking and $q_{L,R}$ represents the weak-eigenstates. The Yukawa matrices are 3x3 complex matrices with non-zero diagonal elements. Therefore the weak eigenstates q are different from the physical mass eigenstates of the quarks q' . The quark masses m_q can be found through their coupling with the Higgs field $m_q = Y_q \frac{v}{\sqrt{2}}$, given in terms of weak eigenstates. In order to write mass terms for the quarks from the mass eigenstates, the Yukawa matrices need to be diagonalised by a unitary transformation $V_{A,q}$ which is determined by:

$$M_u = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} = \text{diag}(m_u, m_c, m_t) = \frac{v}{\sqrt{2}} V_{L,u} Y_u V_{R,u}^\dagger \quad (3)$$

$$M_d = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} = \text{diag}(m_d, m_s, m_b) = \frac{v}{\sqrt{2}} V_{L,d} Y_d V_{R,d}^\dagger \quad (4)$$

The mass eigenstates can then be obtained from the weak eigenstates using the same unitarity matrix $q'_A = V_{A,q} q_A$ with $q = u, d, A = L, R$ and $V_{A,q} V_{A,q}^\dagger = 1$. With that one can rewrite the Yukawa term including the quark masses M_d, M_u and mass eigenstates u', d' as

$$\mathcal{L}_{Yukawa} = -\bar{d}'_L M_d d'_R + \bar{u}'_L M_u u'_R + h.c. \quad (5)$$

These transformations leave all parts of the SM Lagrangian unchanged, except for the term describing the charged current weak interaction which connects the up and down-type quarks. In the basis of weak eigenstates this is given as

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}(\bar{u}_L\gamma^\mu W_\mu^+ d_L + \bar{d}_L\gamma^\mu W_\mu^- u_L) \quad (6)$$

here g is the $SU(2)$ coupling constant. Rewriting it in terms of mass eigenstates leads to the following formula

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}(\bar{u}'_L\gamma^\mu W_\mu^+ \underbrace{V_{L,u}V_{L,d}^\dagger}_{V_{CKM}} d'_L + \bar{d}'_L\gamma^\mu W_\mu^- \underbrace{V_{L,d}V_{L,u}^\dagger}_{V_{CKM}^\dagger} u'_L). \quad (7)$$

There is a net effect of the basis change since the up-type and down-type Yukawa matrices cannot be diagonalised simultaneously by the same unitary transformation, $V_{A,d} \neq V_{A,u}$. Therefore the charged-current interaction gets a flavour structure which is encoded in the Cabibbo-Kobayashi-Maskawa (CKM) [34, 35] quark mixing matrix $V_{CKM} = V_{L,u}V_{L,d}^\dagger$. Each element of this matrix $(V_{CKM})_{ij}$ connects the left-handed up-type quark of the i -th generation with the left-handed down-type quark of the j -th generation, but it is labelled according to quark flavour instead to the generation index. Writing out all transition gives

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (8)$$

As a convention the weak eigenstates and the mass eigenstates for the up-type quarks are chosen to be equal, whereas the down-type quarks are rotated such that the weak eigenstates are a mixture of the mass eigenstates

$$u_i = u'_i, \quad d_i = (V_{CKM})_{ij}d'_j, \quad (9)$$

or explicitly

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (10)$$

Because of the non-diagonal structure of the CKM matrix the weak interaction allows for transitions between different quark generations through charged current interactions. Figure 1 shows the Feynman diagram for a $b \rightarrow u$ transition as an example. This transition matrix element is proportional to the corresponding CKM matrix element V_{ub} , whereas the corresponding anti-quark transition $\bar{b} \rightarrow \bar{u}$ is proportional to its complex conjugate element V_{ub}^* .

2.2.1 CKM matrix parametrisations

The CKM matrix is a 3x3 unitary matrix and it consists of 9 complex elements, satisfying $V_{ij} \neq V_{ij}^*$, therefore four free parameters remain. These are three rotation

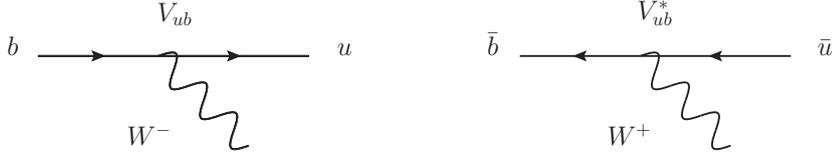


Figure 1: Feynman diagram of a $b \rightarrow u$ transition and a $\bar{b} \rightarrow \bar{u}$ transition.

angles and one phase, the latter one is responsible for CP violation in the SM. There are different parametrizations for the CKM matrix existing.

One of the standard choices uses the three Euler angles θ_{23} , θ_{13} , θ_{12} and one Phase δ [36]:

$$\begin{aligned}
 V_{CKM} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}s_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (11)
 \end{aligned}$$

where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$ and δ is the phase responsible for CP violation in the quark sector.

Alternatively due to its hierarchical structure with smaller contributions away from the diagonal, the so-called Wolfenstein parametrisation [37] is mainly used as an approximation up to $\mathcal{O}(\lambda^4)$ terms:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad (12)$$

with $\lambda = 0.225$ as the expansion parameter and A , ρ and η are real parameters between 0.1-1 which do not change the order of the magnitude of the CKM elements. The two parametrisations are related by

$$\begin{aligned}
 s_{12} = \lambda &= \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, & s_{23} = A\lambda^2 &= \lambda \frac{|V_{cb}|}{|V_{us}|}, \\
 s_{13}e^{i\delta} = V_{ub}^* &= A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1 - \lambda^2[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]}}, \quad (13)
 \end{aligned}$$

which ensures that

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \quad (14)$$

is phase convention independent and the CKM matrix is unitary to all orders of λ . From the observed structure of the CKM matrix, which will be discussed in Section 2.2.3, one can see that quark transitions within the same generation are preferred since diagonal elements are of order 1 whereas further away from the diagonal the transitions are more suppressed. Between the first two generations the transition is

suppressed by λ , between the second and third by λ^2 , whereas between the first and third generation by λ^3 . This parametrization also shows that the complex components up to order λ^3 terms are the CKM matrix elements V_{ub} and V_{td} , allowing for CP violation in the SM.

2.2.2 The unitarity triangle

The unitarity of the CKM matrix

$$V_{CKM}^\dagger V_{CKM} = V_{CKM} V_{CKM}^\dagger = \mathbb{1} \quad (15)$$

$$\Leftrightarrow \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} V_{ud}^* & V_{cd}^* & V_{td}^* \\ V_{us}^* & V_{cs}^* & V_{ts}^* \\ V_{ub}^* & V_{cb}^* & V_{tb}^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (16)$$

leads to a set of 9 equations. Three of them are unitary relations

$$V_{ud}V_{ud}^* + V_{us}V_{us}^* + V_{ub}V_{ub}^* = 1 \quad (17)$$

$$V_{cd}V_{cd}^* + V_{cs}V_{cs}^* + V_{cb}V_{cb}^* = 1 \quad (18)$$

$$V_{td}V_{td}^* + V_{ts}V_{ts}^* + V_{tb}V_{tb}^* = 1. \quad (19)$$

These express the so-called *weak universality* since the squared sum of the coupling strengths of the u -quark to d , s and b -quarks is equal to the overall charged coupling of the c -quark and the t -quark. Since all of them add up to 1, there is no probability remaining to couple to 4-th down-type quark. Those relations need to be tested experimentally.

The remaining 12 relations are known as *orthogonality conditions*, where the following 6 linear independent equations remain since the other six are only the complex conjugate version:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (db) \quad (20)$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \quad (sb) \quad (21)$$

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad (ds) \quad (22)$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 \quad (ut) \quad (23)$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \quad (ct) \quad (24)$$

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \quad (uc), \quad (25)$$

$$(26)$$

where the symbols (db) indicates the row and column whose orthogonality condition is used. As they are sums of three complex numbers that must yield zero they can be displayed as a triangle in the complex plane as shown in Figure 2. All of these unitarity triangles have the same area, the so-called Jarlskog invariant [39] $J/2$ which is a measure of CP violation of the SM. It is defined as

$$Im[V_{ij}V_{kl}V_{il}^*V_{kj}^*] = J \sum_{mn} \epsilon_{ikm}\epsilon_{jln}, \quad (27)$$

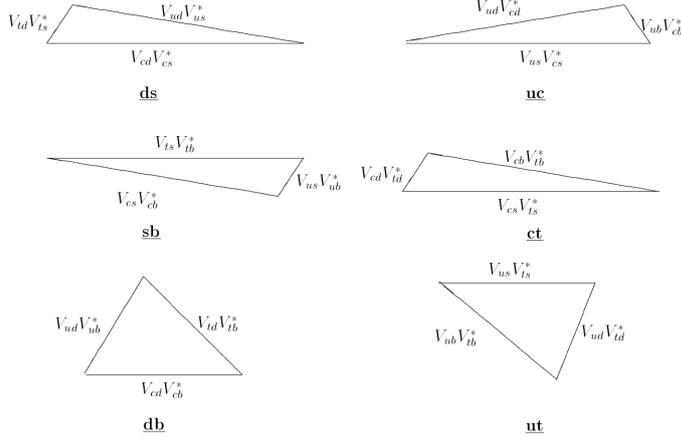


Figure 2: A schematic of the six different unitarity triangles of the quark sector, the magnitudes are not to scale. Figure taken from [38].

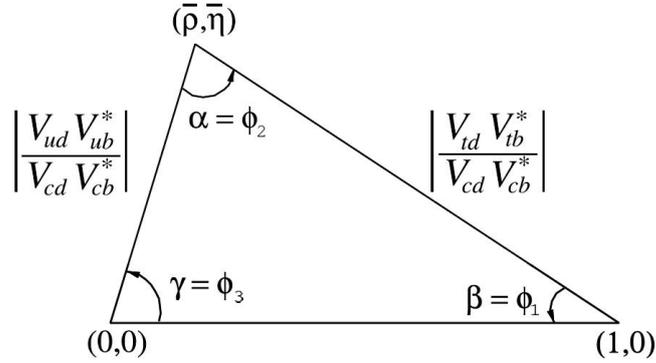


Figure 3: A sketch of the unitarity triangle [12].

in terms of CKM parametrisations this corresponds to

$$J = c_{12}c_{23}c_{13}^2 s_{12}s_{23}s_{13} \sin \delta \approx \lambda^6 A^2 \eta. \quad (28)$$

Only two out of the 6 unitarity triangles have all three sides with similar lengths of the order of λ^3 , these are marked as (db) and (ut) triangles in Figure 2. These two triangles are relevant for B-decays. For historic reasons due to measurements performed by the B-factories Belle and BaBar the (db) triangle is known as *the unitarity triangle*, the analogous triangle of the B_s^0 system is the (sb) triangle. By dividing the three sides of the (db) triangle by the best known CKM elements $V_{cd}V_{cb}^*$ yields to the famous unitarity triangle shown in Figure 3 and the relation:

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \quad (29)$$

In this triangle one side has a unit length and points along the real axis, the apex is located by definition at $(\bar{\rho}, \bar{\eta})$ which was defined already before in Equation 14. They can also be expressed in terms of Wolfenstein parameters ρ and η as: $\bar{\rho} = \rho(1 - 1/2\lambda^2) + \mathcal{O}(\lambda^4)$, $\bar{\eta} = \eta(1 - 1/2\lambda^2) + \mathcal{O}(\lambda^4)$.

The angles of the unitarity triangle are defined as:

$$\alpha \equiv \arg \left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right], \quad \beta \equiv \arg \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right], \quad \gamma \equiv \arg \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right]. \quad (30)$$

where β and γ are to first order the phases of V_{td} and V_{ub} , respectively, using the Wolfenstein parametrisation. Those are measured from CP violating observables, such as CP asymmetries of B decays and are briefly explained in the next section.

2.2.3 Constraining the CKM matrix

One goal of flavour physics is to measure experimentally the four free parameters of the unitarity triangle very precisely in order to verify the Standard Model description. This is done by overconstraining the triangle from complementary measurements which are sensitive to the magnitude and phases of the CKM matrix elements and can reveal effects beyond the SM if a disagreement between the angles of the triangle and the lengths of the sides is found.

In general the CKM matrix elements are fundamental parameters of the SM with no theory predictions, therefore the precise determination of these parameters is very important. Magnitudes of CKM matrix elements are usually measured with semileptonic or fully leptonic decays including charged current quark transitions, such as the nuclear β decay to extract V_{ud} for example. The measured decay rates of the respective flavour changing transition is proportional to the coupling strength $|V_{ij}|^2$. Previous measurements lead to an almost diagonal hierarchical structure depicted already by the Wolfenstein parametrization [12]:

$$V_{CKM} = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359_{-0.00011}^{+0.00010} & 0.04214 \pm 0.00076 \\ 0.00896_{-0.00023}^{+0.00024} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix}. \quad (31)$$

The smallest CKM matrix element is $|V_{ub}|$, it has also the biggest relative uncertainty and is therefore the least well-known element. Its uncertainty is the dominant uncertainty on the length of the triangle side opposite to angle β . Therefore it is important to measure this CKM matrix element with greater precision.

The matrix elements $|V_{tb}|$ and $|V_{td}|$ are needed to determine the length opposite to the angle γ , but their combination $|V_{td}V_{tb}^*|$ can be extracted more precisely by measuring the mass difference of B^0 meson eigenstates $\Delta m_d \propto |V_{td}V_{tb}^*|^2$. This quantity is important for $\bar{B}^0 - B^0$ oscillations. Due to significant reduced theory uncertainties the ratio of B^0 and B_s^0 mass differences $\Delta m_d/\Delta m_s \propto |V_{td}V_{tb}^*|^2/|V_{ts}V_{tb}^*|$ is often used as a better constraint. The dominating uncertainty extracting the magnitude of CKM matrix elements comes from theoretical uncertainties on hadronic matrix elements including the QCD nature, which will be explained in more detail in the next section.

The phases of CKM matrix elements can be determined by measurements of CP violating observables, especially important are measurements which can

extract the angles of the unitarity triangle. The angle $\beta = \arg(-V_{cd}V_{cb}^*/V_{td}V_{tb}^*)$ can be determined from time dependent decay rates of B^0 and \bar{B}^0 to the same final state f , here the B^0/\bar{B}^0 meson can either directly decay to the final state f or first oscillate $B^0 - \bar{B}^0$ into its antiparticle and then decay to f . Since the two paths have a phase difference (weak and strong phase difference), interference can occur and results in a time dependent asymmetry. If f is an eigenstate and amplitudes with one CKM phase dominate the decay, the time-dependent CP symmetry is given by:

$$A_{CP} = \frac{\Gamma(\bar{B}^0(t) \rightarrow f) - \Gamma(B^0(t) \rightarrow f)}{\Gamma(\bar{B}^0(t) \rightarrow f) + \Gamma(B^0(t) \rightarrow f)} = \eta_f \sin(2\beta) \sin(\Delta m_d t), \quad (32)$$

where η_f is the CP eigenvalue of f . $\sin(2\beta)$ can be measured from $b \rightarrow c\bar{c}s$ transitions into the same final state f which is a CP eigenstate, such as $B^0 \rightarrow J/\psi K_{S/L}^0$, $b \rightarrow c\bar{c}d$ transitions like $B^0 \rightarrow J/\psi \pi^0$ and $b \rightarrow c\bar{u}d$ transitions as $B^0 \rightarrow \bar{D}^0 h^0$ [12].

The angle $\alpha = \arg(-V_{td}V_{tb}^*/V_{ud}V_{ub}^*)$ is measured using time-dependent CP asymmetries from $b \rightarrow u\bar{u}d$ decay dominated modes such as $B \rightarrow \pi\pi$, $\rho\pi$ and $\rho\rho$ decays. Unlike α and β , the angle $\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)$ does not depend on CKM elements involving the top quark and can therefore directly be measured in tree-level decays. For example it can be measured from the interference of the $B^- \rightarrow D^0 K^-$ ($b \rightarrow c\bar{u}s$) and $B^- \rightarrow \bar{D}^0 K^-$ ($b \rightarrow u\bar{c}s$) decays, where D^0 and \bar{D}^0 decay into the same final state.

Additional constraints on the CKM sector come from CP violation measurement in the kaon system such as $K^0 - \bar{K}^0$ mixing and the CP violating phase $\beta_s = \arg(-V_{ts}V_{tb}^*/V_{cs}V_{cb}^*)$ from $B_s^0 \rightarrow J/\psi\phi$ decays.

The parameters of the CKM matrix $\bar{\rho}, \bar{\eta}, A$ and λ are most precisely determined using a global fit to all of the available measurements and imposing SM constraints such as the 3 generation unitarity $V_{ud}^2 + V_{us}^2 + V_{ub}^2 = 1$ for different rows. There are two collaborations combining experimental data with theory predictions for hadronic matrix elements using different approaches: the CKMfitter [40] uses frequentist statistics and the UTfit [41] collaboration using the Bayesian approach. Both of them provide similar results and the constraints implied by unitarity significantly reduce the allowed range for some of the CKM elements [12]. Figure 4 compares the apex of the unitarity triangle from the global fits obtained by both collaborations. Their results of the global fits are in excellent agreement with each other and also with SM predictions, no significant deviation of the SM CKM picture is found up to now. As an example the sensitivity of the different measurements entering the global fit is shown in Figure 5, where angle measurements are shown on the left and constraints from all other measurements on the right.

2.3 Semileptonic B -meson decays

As emphasized in the previous chapter it is important to determine the unitarity triangle with great precision. Especially the precise determination of $|V_{ub}|$ and

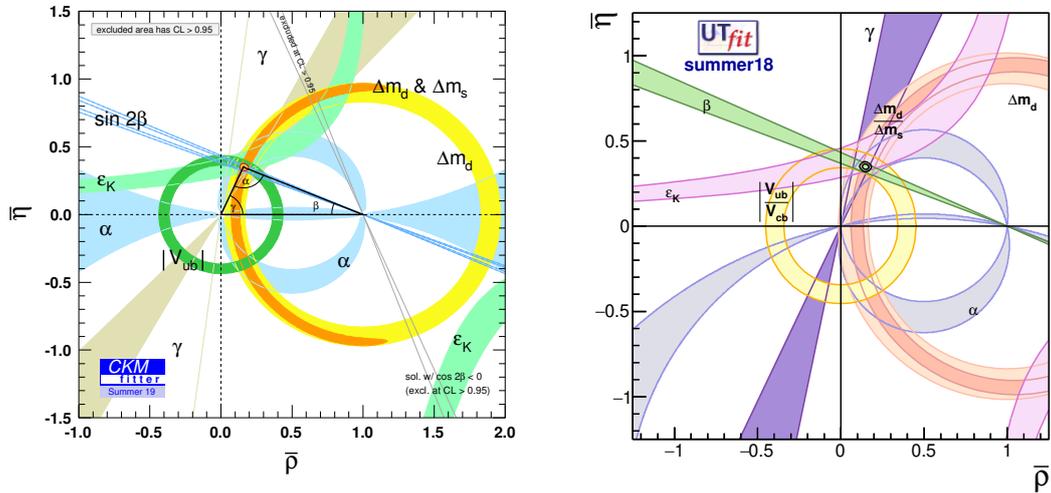


Figure 4: Global fits to the CKM parameters $\bar{\rho}$, $\bar{\eta}$, A and λ from the CKMfitter [40] and UTfit [41] collaborations as of Spring 2020.

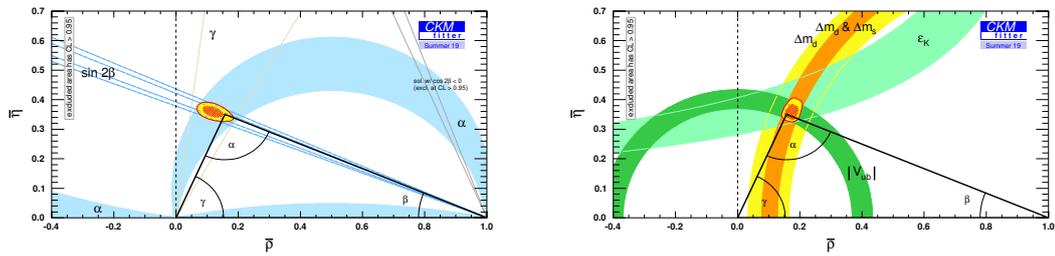


Figure 5: Constraints to the $\bar{\rho}$, $\bar{\eta}$ -plane, from the CKMfitter [40] collaboration including only angle measurements on the left and excluding angle measurements on the right, as of Spring 2020.

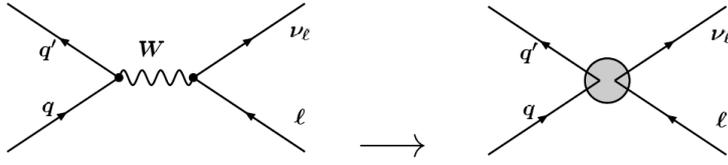


Figure 6: Effective semileptonic quark transitions.

$|V_{cb}|$ are central tests of the CKM sector of the SM, as the length of the side of the unitary triangle opposite to the angle β is directly proportional to the ratio $|V_{ub}|/|V_{cb}|$. Therefore extracting this quantity precisely is of great importance in the heavy-flavour physics program and also the aim of this thesis.

Semileptonic $b \rightarrow u\mu\nu$ and $b \rightarrow c\mu\nu$ transitions are well suited to study the CKM matrix elements as they have large branching fractions of $\approx 10\%$ of all B decays and provide therefore large yields. Experimentally reconstructing the neutrino in the final state is challenging as it can not be directly detected such that reconstructing the decay kinematics is more challenging. Theoretically semileptonic decays are much simpler to calculate than fully hadronic processes since the leptons in the final state do not interact strongly, such that the decay can be factorized into a hadronic and leptonic part. In general there are different strategies to extract the CKM matrix elements: either from inclusive, exclusive semi-leptonic or purely leptonic B-decays. Here the focus lies on exclusive semi-leptonic decays as they are used for this thesis, its theoretical description will be further discussed in the following.

Theoretical overview: Weak decays of hadrons

In the SM charged current semi-leptonic decays at quark level are mediated by an exchange of a W^\pm -boson between the quark and lepton-current. Since the W -boson is much heavier than the b -quark, one can integrate out the W boson and consider it to be infinitely heavy:

$$\langle 0|T[W_\mu(x)W_\nu^*(0)]|0\rangle \sim \frac{1}{M_W^2}\delta^4(x). \quad (33)$$

This leads to an effective 4-point fermion interaction as depicted in Figure 6, with a semi-leptonic local effective Hamiltonian for the decay of interest of the form:

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}}V_{ub}(\bar{u}\gamma_\mu(1 - \gamma^5)d)(\bar{\mu}\gamma^\mu(1 - \gamma^5)\nu_\mu) \quad (34)$$

with $G_F = \frac{g^2}{4\sqrt{2}M_W^2}$ is the Fermi coupling constant which is obtained from the comparison to the full SM calculations. This approximation is correct up to order m_b^2/M_W^2 and one can see how the leptonic and hadronic part of a decay factorizes. Using this the matrix element for a semileptonic decay of a b-hadron to an exclusive

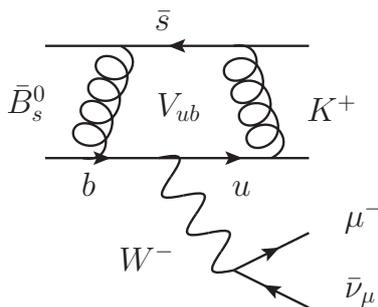


Figure 7: Feynman diagram for the semi-leptonic decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$.

final state $X_q^- \mu^+ \nu_{mu}$ can be written as

$$\begin{aligned} \mathcal{M} &= -i \langle X_q(p') \mu^+(k') \nu_\mu(k) | \mathcal{H}_{\text{eff}} | B(p) \rangle \\ &= -i \frac{G_F}{\sqrt{2}} V_{ub} H_\mu L^\mu, \end{aligned} \quad (35)$$

where p and k denote the 4-momentum. The leptonic current is given as

$$L^\mu = \bar{\mu} \gamma^\mu (1 - \gamma_5) \nu_\mu, \quad (36)$$

and the hadronic matrix element

$$H_\mu = \langle X_q(p') | \bar{u} \gamma^\mu (1 - \gamma_5) b | B(p) \rangle. \quad (37)$$

The latter is not trivial to calculate due to the strong dynamics describing the bound states of QCD, which at hadronic energy scales of a few MeV to GeV are non-perturbative and can therefore not be calculated analytically. The hadronic matrix element can be written in term of vector and axial vector Lorentz symmetries and such decomposed into so-called *form factors* which incorporate QCD effects. In general form factors describe shape corrections to the approximation that the scattering object is not point-like. Those form factors can be determined using non-perturbative methods such as QCD sum rules and Lattice QCD, where the latter is further explained in the next section.

The Feynman diagram for the decay of interest in this thesis is shown in Figure 7. Since $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ is a pseudoscalar meson transition $B_s^0 (J^P = 0^-) \rightarrow K^- (J^P = 0^-)$, the axial-vector component of the matrix element H_μ is zero due to the Lorentz structure of the initial and final state, which is conserved in QCD. Therefore only the vector component of H^μ is remaining which can be written in terms of scalar and vector form factors f_0 and f_+ :

$$\langle K^-(p') | \bar{u} \gamma^\mu b | B_s^0(p) \rangle = f_+(q^2) \left(p^\mu + p'^\mu - \frac{M_{B_s^0}^2 - M_K^2}{q^2} q^\mu \right) + f_0(q^2) \frac{M_{B_s^0}^2 - M_K^2}{q^2} q^\mu, \quad (38)$$

where p^μ , p'^μ are the B_s^0 and K 4-momenta, $M_{B_s^0}$ and M_K are the corresponding meson masses and $q^\mu = p^\mu - p'^\mu$ is the momentum transfer to the lepton pair.

This expression above also applies for the form factors of the corresponding $b \rightarrow c$ transition for the decay $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ when substituting the u with the c -quark and the kaon with the D_s^- meson mass. The determination of the form factors, $f_+(q^2)$ and $f_0(q^2)$ are discussed in Section 2.4.

2.4 Lattice QCD

As explained before the form factors incorporating the QCD effects of the semileptonic decay have to be calculated in order to be able to extract the CKM matrix elements of interest. Theoretical input from non-perturbative methods is needed, such as QCD sum rules [42] or Lattice QCD (LQCD) which are used for form factor predictions.

LQCD is a non-perturbative method where the QCD action

$$S_{QCD} = \int \mathcal{L}_{QCD} d^4x \quad (39)$$

can be calculated numerically by discretising space-time on a lattice [6, 43, 44]. When considering a quantum mechanical path $x(t)$ of a particle at the time t with the boundary positions $x(0)$ and $x(t_f)$, the path integral formalism of quantum mechanics gives each path a probability proportional to $\exp(-\int dt \mathcal{L})$. The expectation value of an operator combination is given by the correlation function

$$\langle \mathcal{O}(x(t_1)x(t_2)) \rangle = \frac{\int \mathcal{D}x(t) \mathcal{O}(x(t_1)x(t_2)) e^{-\int \mathcal{L} dt}}{\int \mathcal{D}x(t) e^{-\int \mathcal{L} dt}} \quad (40)$$

where $\mathcal{D}x(t)$ is an integral over all possible paths $x(t)$.

This expectation value can be solved numerically using a one dimensional lattice in time with spacing a . This needs Hybrid Monte Carlo methods [45] to generate large ensembles of lattice configurations N_{conf} , where each configuration corresponds to a different path along the lattice traversed by the particle. The probability of finding a given configuration within the ensemble is proportional to $\exp(-\int dt \mathcal{L})$. Then the computation of the correlation function in Equation 39 simplifies to an average of the correlation function for each of the configurations

$$\langle \mathcal{O}(x(t_1)x(t_2)) \rangle = \frac{1}{N_{conf}} \sum_{n=1}^{N_{conf}} \mathcal{O}(x(t_1)x(t_2)). \quad (41)$$

Its corresponding statistical uncertainty can be calculated using the standard deviation of the sample which is proportional to $1/\sqrt{N_{conf}}$. This concept of using a lattice to numerically calculate the correlation function using the path integral can be extended to QCD, where the quark field can be represented at point in a 4-dimensional lattice. Using this approach of the lattice calculation leads to several sources of systematic errors, which have to be evaluated:

- **Continuum limit:** the computation of the action on the lattice is only exact in the limit of $a \rightarrow 0$, so for an extrapolation of the lattice spacing to the

continuum. This is computationally not possible such that lattice results are usually computed for a range of lattice spacings.

- **Finite volume:** Lattice QCD calculations cover only a finite volume in space and time where the action is integrated over, while in quantum mechanics an integral over an infinite volume of space-time is done. This results in a shift of physical quantities computed on the lattice with respect to the true value.
- **Chiral extrapolation:** The predicted pion mass varies between different lattice configurations such that an extrapolation of the results to the physical pion mass is required. This extrapolation is done based on models or effective field theories.
- **Operator matching:** Operators which are defined in terms of the lattice scheme must be matched to their continuum version under the appropriate renormalisation. This requires non-perturbative techniques which come with their own systematic uncertainties.
- **Quark masses:** LQCD simulations use light quark masses above the true masses and heavy quarks below their true value to be IR and UV safe, otherwise it would require tiny lattice spacings smaller than the inverse mass of the heavy quark, so below $0.04 fm$ for the b -quark and very large lattices. Therefore extrapolations to the true values are needed.

In addition it is very difficult to simulate objects with large momenta on the lattice. Therefore form factors for decays of heavy mesons cannot be computed in regions where the outgoing meson has large momentum in the rest frame of the decaying meson [46]. Rather lattice form factor simulations are restricted to regions of maximum momentum transfer to the leptons.

2.5 Theoretical form factors predictions

For the decay of interest $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ there are different calculations of form factors available from either LQCD and light-cone sum rules (LCSR). The two calculation methods provide predictions which are complementary in phase space: LQCD calculations are restricted to high values of q^2 whereas LCSR calculations are restricted to low values of q^2 . Therefore two different q^2 regions are used for the signal decay in this thesis. Requirements on unitarity and analyticity can then be used to extrapolate results to the full available q^2 region of the decay of interest leading to additional uncertainties. The decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ is normalised to the decay $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ for which form factor calculations from LQCD are available. The form factor calculations for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ needs no restriction to only the high momentum transfer region as they have been precisely determined in [47] for the first time for the full q^2 spectrum.

2.5.1 Form factor calculations for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$

From the hadronic matrix element defined in Equation 38 of Section 2.3 the differential decay rate for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ in the B_s^0 rest frame can be written as

$$\frac{d\Gamma(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{dq^2} = \frac{G_F^2 |V_{ub}|^2}{24\pi^3} \frac{(q^2 - m_\mu^2)^2 \sqrt{E_K^2 - M_K^2}}{q^4 M_{B_s^0}^2} \times \quad (42)$$

$$\left[\left(1 + \frac{m_\mu^2}{2q^2} \right) M_{B_s^0}^2 (E_K^2 - M_K^2) |f_+(q^2)|^2 + \frac{3m_\mu^2}{8q^2} (M_{B_s^0}^2 - M_K^2)^2 |f_0(q^2)|^2 \right],$$

where $E_K = \frac{M_{B_s^0}^2 + M_K^2 - q^2}{2M_{B_s^0}}$ is the energy of the kaon and q^2 is the momentum transfer and $|f_+(q^2)|$ and $|f_0(q^2)|$ as the vector and scalar form factors, respectively. The form factors are calculated on the lattice using 3-point correlation functions from simulated data at large q^2 , which is typically above 17 GeV². In order to extrapolate the form factors to the whole kinematically allowed region a model-independent parametrisation is used, the so-called *z-expansion*. It is based on the analyticity and unitarity of the form factors. There are different extrapolation parametrisations existing, most commonly the parametrisation from Bourrely, Caprini, and Lellouch [48] (BCL) is used, where the q^2 variable is mapped to a new variable z according to

$$z = (q^2, t_0) = \frac{\sqrt{1 - q^2/t_+} - \sqrt{1 - t_0/t_+}}{\sqrt{1 - q^2/t_+} + \sqrt{1 - t_0/t_+}}. \quad (43)$$

Here $t_\pm \equiv (M_{B_s^0} \pm M_K)^2$ and $t_0 = t_{opt} \equiv (M_{B_s^0} + M_K)(\sqrt{M_{B_s^0}} - \sqrt{M_K})^2$. This transformation maps the semileptonic kinematic region $0 < q^2 < (M_{B_s^0} - M_K)^2$ onto a unit circle in the complex z -plane. The form factors can then be expanded as a power series in z :

$$f_+(q^2) = \frac{1}{1 - q^2/M_+^2} \sum_{k=0}^{K-1} b_+^{(k)} \left[z^k - (-1)^{k-K} \frac{k}{K} z^K \right], \quad (44)$$

$$f_0(q^2) = \frac{1}{1 - q^2/M_0^2} \sum_{k=0}^{K-1} b_0^{(k)} z^k,$$

where different poles are included at around $M_+ = 5.325$ GeV and $M_0 = 5.68$ GeV. Different lattice collaborations use different truncation values of K to end the power series, usually values of 2 or 3 are chosen. The exact values of the coefficients $b_+^{(k)}$ and $b_0^{(k)}$ and their correlations for the different lattice results are specified in Appendix A.

The vector and scalar form factors are related through the kinematic constraint:

$$f_0(q^2 = 0) = f_+(q^2 = 0). \quad (45)$$

Four different calculations of form factors are used for the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decay. Lattice QCD predictions provide a precise determination of the form factors

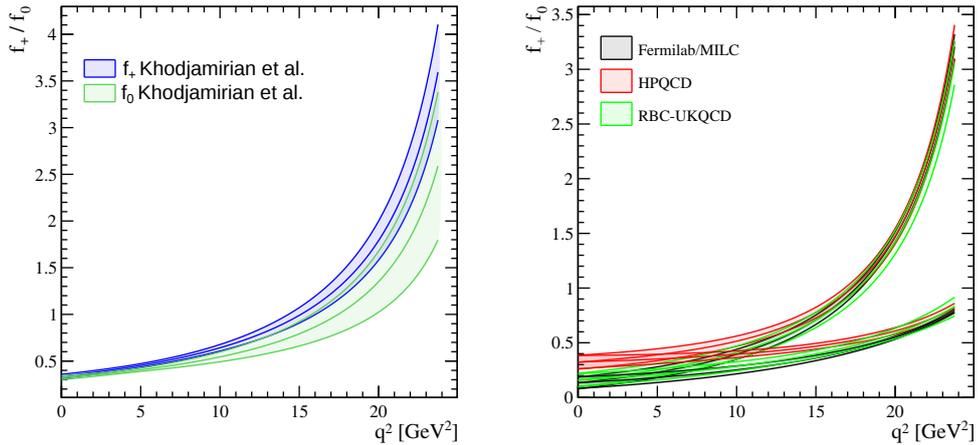


Figure 8: The form factor predictions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ calculated from LCSR [52] (left) and lattice QCD (right) from references [49–51]. In the right Figure the distribution of the scalar form factor f_0 is the lower component of the plot. Figure modified from [9].

at low recoil or high momentum transfer q^2 and are provided by the HPQCD [49], RBC-UKQCD [50] and Fermilab/MILC [51] collaboration. Calculations from light-cone sum rules are most precise at large hadronic recoil or low q^2 , they are provided by Khodjamirian et al. [52]. Their results are compared in the following.

The predicted vector and scalar form factors as a function of q^2 are shown in Figure 8, as the q^2 dependence of the form factors is parametrization independent which is not the case as a function of z since different groups use different thresholds for the z -expansion. The left plot shows the predictions from LCSR from Khodjamirian et al. and the right plot shows the comparison of the the different LQCD collaborations. In the latter one can see how the different predictions agree at high q^2 where lattice QCD is most precise, but they differ at low q^2 . Whereas the results from the RBC-UKQCD and Fermilab/MILC collaborations are consistent, the predicted form factors from HPQCD are higher at low q^2 which is better visible in Figure 9 made by the Fermilab/MILC collaboration. The Fermilab/MILC collaboration estimated the tension with respect to the the HPQCD result to be 2.3σ at $q^2 = 0$ [51] and noted that the HPQCD collaboration uses a so-called modified z -expansion which may effect the shape of the form factors. A more detailed discussion can be found in their paper or Appendix A.5 of the FLAG Review 2019 [53] which suggests underestimated systematic uncertainties for the modified expansion. Further lattice calculations might clarify this difference which are both expected in the near future from the Fermilab/MILC and RBC-UKQCD collaboration using improved lattice simulations.

The differential decay rate according to Equation 42 in the B_s^0 rest frame can then be calculated using the scalar and vector form factor predictions from LCSR and LQCD. The predicted decay rate is shown in Figure 10 as a function of q^2 . It differs significantly at low q^2 coming from the form factor calculations as

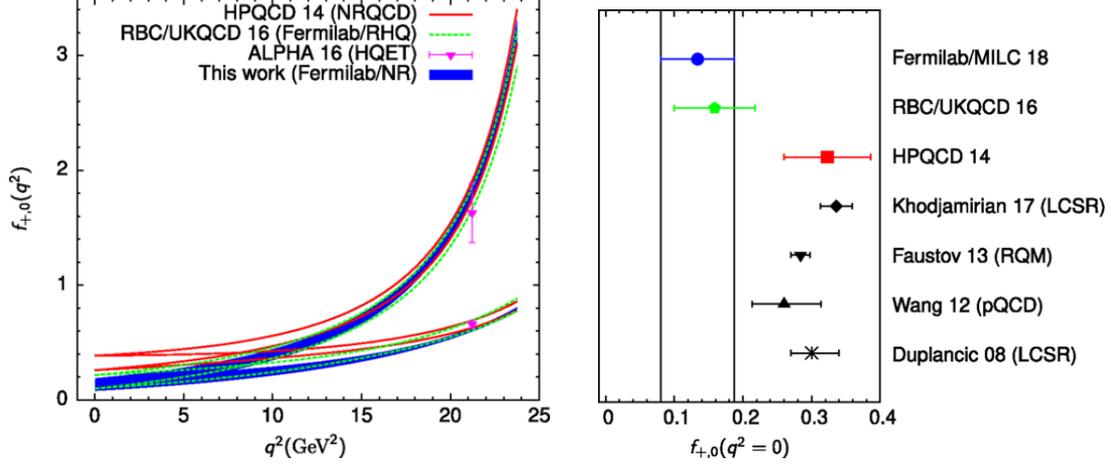


Figure 9: The comparison of form factor predictions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ calculated from lattice QCD for the full q^2 range (left) from references [49–51, 54] and a comparison at $q^2 = 0$ for predictions from LCSR [52, 55], Lattice QCD [49–51], NLO perturbative QCD [56] (pQCD) and relativistic quark model [57] (RQM) on the right. Figures are taken from [51].

explained above. Predictions from LCSR are much larger here which is consistent with predictions from HPQCD within their large errors but not with the other lattice results from the RBC-UKQCD and Fermilab/MILC collaborations. The results of the form factor calculations are given at the end of this section.

2.5.2 Form factor calculations for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$

The $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ differential decay rate is similar to Equation 42 and given by

$$\frac{d\Gamma(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{dq^2} = \frac{G_F^2 |V_{cb}|^2 (q^2 - m_\mu^2)^2 \sqrt{E_{D_s}^2 - M_{D_s}^2}}{24\pi^3 q^4 M_{B_s}^2} \times \left[\left(1 + \frac{m_\mu^2}{2q^2} \right) M_{B_s}^2 (E_{D_s}^2 - M_{D_s}^2) |f_+(q^2)|^2 + \frac{3m_\mu^2}{8q^2} (M_{B_s}^2 - M_{D_s}^2)^2 |f_0(q^2)|^2 \right], \quad (46)$$

where $E_{D_s} = \frac{M_{B_s}^2 + M_K^2 - q^2}{2M_{B_s}}$ is the energy of the D_s meson. There are also different versions existing in the literature [58, 59] which express it in terms of a different variable ω which can be related to q^2 by $\omega(q^2) = 1 + \frac{q_{\max}^2 - q^2}{2m_{B_s} m_{D_s}}$. Here the discussion follows the latest paper from McLean et al. [47], as it can be related to the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal channel very easily.

The form factor $f_+(q^2)$ and $f_0(q^2)$ are again expressed in terms of the new variable z according to Equation 43 where $t_\pm = (m_{B_s} \pm m_{D_s})^2$ is now defined in terms of the D_s meson and $t_0 = 0$. The form factors are parametrised using the BCL parametrisation introduced in Equation 44 with parameters $a_{+0}^{(k)}$ and the pole masses $M_+ = 6.329$ GeV and $M_0 = 6.704$ GeV.

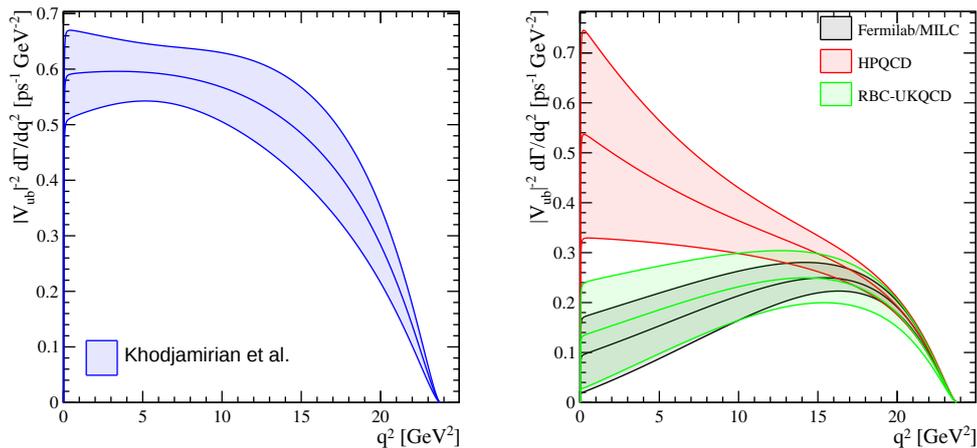


Figure 10: The predicted differential decay rates for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ calculated using QCD sum rules (left) and lattice QCD (right) from references [49–52]. Figure modified from [9].

Three different sets of lattice QCD form factor calculations are used for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$, one from the Fermilab/MILC collaboration [58] and two more recent papers from the HPQCD collaboration [47, 59]. The calculated form factors and differential decay rates are plotted in Figure 11. One can see that the updated measurement from HPQCD [47] is much more precise than the previous two predictions and in general the predictions agree much better than the ones from $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. The coefficients $a_{+,0}^{(k)}$ used for all models discussed in this section are given in Appendix A together with their correlation matrix.

2.5.3 Form factor results

The predicted decay widths for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ are calculated from the different form factor predictions [49–52] by integrating Equation 42 over q^2 . Different ranges of q^2 are used in this thesis: $q^2 < 7 \text{ GeV}^2$ (low q^2), $q^2 > 7 \text{ GeV}^2$ (high q^2) and the full available q^2 range which is $M_\mu^2 < q^2 < (M_{B_s^0} - M_K)^2$ up to $\approx 24 \text{ GeV}^2$. The choice to measure the decay in different bins of q^2 is motivated by the fact that the LQCD form factor predictions are evaluated at high q^2 ($q^2 \geq 17 \text{ GeV}^2$) and as such most precise in that region, whereas the LCSR predictions are valid at low or intermediate q^2 range at $0 < q^2 < 12 \text{ GeV}^2$. The calculated decay widths of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ are summarised in Table 3. They differ widely between the used predictions due to the large differences in the estimated form factors given by LQCD and LCSR as discussed before.

Table 3 also gives the estimated branching fractions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ from the different theory predictions which can be obtained from the decay width by multiplying with B_s^0 lifetime taken from the Heavy Flavor Averaging Group (HFLAV) [60] and the PDG average [12] of exclusive $|V_{ub}|^2$. Here the predictions vary between a large range of branching fractions from as low as 0.88×10^{-4} to 2.29×10^{-4} .

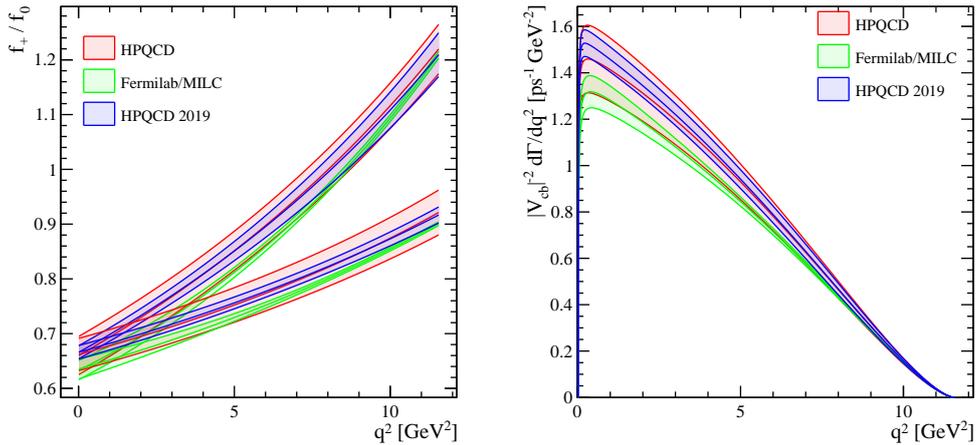


Figure 11: The form factor predictions (left) and differential decay rates (right) for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ calculated using lattice QCD from references [47, 58, 59]. In the left Figure the distribution of the scalar form factor f_0 is the lower component of the plot. Figure modified from [9].

	$\Gamma V_{ub} ^{-2}[ps^{-1}]$ full q^2	$\Gamma V_{ub} ^{-2}[ps^{-1}]$ low q^2	$\Gamma V_{ub} ^{-2}[ps^{-1}]$ high q^2	$\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)$ [10^{-4}]
RBC-UKQCD	4.54 ± 1.35	1.18 ± 0.67	3.37 ± 0.70	(0.94 ± 0.27)
HPQCD	7.75 ± 1.57	3.29 ± 0.99	4.47 ± 0.61	(1.60 ± 0.32)
Fermilab/MILC	4.26 ± 0.92	0.94 ± 0.48	3.32 ± 0.46	(0.88 ± 0.20)
LCSR	11.07 ± 1.13	4.14 ± 0.40	6.94 ± 1.02	(2.29 ± 0.23)

Table 3: The predicted decay widths and branching fractions of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ are calculated from the form factor predictions given in References [49–52] for the full q^2 region as well as the high and low q^2 bins.

The predicted decay widths for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ are calculated using the form factor predictions from References [47, 58, 59] and are compared in Table 4. Here the full available q^2 range is used which is smaller due to the heavier D_s mass, it covers the region of $M_\mu^2 < q^2 < (M_{B_s^0} - M_{D_s})^2$ up to around 12 GeV^2 . In addition the estimated branching fractions for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ are calculated from Lattice QCD predictions using the PDG average [12] of exclusive $|V_{cb}|^2$. In general the predicted decay width as well as branching fractions have a better agreement with each other.

These predictions can be compared to the recent LHCb paper [61] which measured the branching fraction of $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ to be

$$\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu) = (2.49 \pm 0.12(stat) \pm 0.14(syst) \pm 0.16(ext)) \times 10^{-2}. \quad (47)$$

This measurement is very compatible with the prediction from Reference [47] and consistent within errors with respect to Reference [59]. The older HPQCD prediction from Reference [58] predicts a significantly lower branching fraction.

	$\Gamma V_{cb} ^{-2}[ps^{-1}]$	$\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu) [10^{-2}]$
HPQCD	8.98 ± 0.73	2.38 ± 0.20
Fermilab/MILC	8.17 ± 0.24	2.15 ± 0.06
HPQCD 2019	9.16 ± 0.37	2.43 ± 0.09

Table 4: The predicted decay widths and branching fractions of $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ are determined using the form factor predictions given in References [47, 58, 59].

3 The LHCb Experiment

This chapter gives an overview of the LHCb experiment at the Large Hadron Collider. It is not meant as a review of all components of the experiment but rather to briefly introduce the components which are relevant for the analysis presented in this thesis. This chapter starts with a short description of the accelerator complex, then discusses the LHCb experiment and ends with the used data and simulated samples for this thesis.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is operated by the European Organization for Nuclear Research (CERN) and located near Geneva, Switzerland. It has a circumference of about 27 km and is the most powerful particle accelerator built up-to-now [62]. The LHC consists of two circular storage rings where mainly protons but also heavy ions are accelerated in opposite directions up to energies of 7 TeV. It serves only as the last step of a complex pre-accelerator system consisting of linear (LINACs) and circular accelerators (PS, SPS). They use as input ionised hydrogen atoms and increase the beam energy up to 450 GeV before the proton beams get injected to the LHC rings. The beams are accelerated by sixteen radiofrequency cavities, 1232 superconducting dipole magnets with a magnetic field of 8 Tesla keep them in their circular trajectory and quadrupole magnets help to focus the beam. Each proton beams consist of up to 2808 bunches of approximately 10^{11} protons which are spaced 25 or 50ns apart. Those bunches are brought to collision at four interaction points with a collision rate of up to 40 MHz, there the main detectors are located: ALICE [63], ATLAS [32], CMS [33] and LHCb [64] which are built to investigate these collisions.

The LHC started operating in 2010 with symmetric proton-proton collisions at a center-of-mass energy of 7 TeV and went up to 8 TeV in 2012. This run period between 2010-2012 is referred to as *Run I* and was followed by a long shutdown period (LS1) of two years, where dipole magnets were upgraded to increase the center-of-mass energy of the LHC. This allowed for a center-of-mass energy of 13 TeV in the data taking period from 2015-2018, also referred to as *Run II*. Currently a second long shutdown phase (LS2) is taking place in order to upgrade the LHC to the maximum design center-of mass energy of 14 TeV as well as to prepare the detectors to run at higher luminosities. The LHC will finally resume operation in 2021.

3.2 The LHCb Experiment

The *Large Hadron Collider beauty* (LHCb) is one of the four large detectors at the LHC, but has in comparison with ATLAS and CMS as symmetric multi-purpose detectors around the interaction point a unique detector geometry: it is a single-arm forward spectrometer covering an angle between 10 to 300 mrad with respect to the beam axis. The LHCb experiment is shown in left of Figure 12. This design was chosen to study heavy hadrons containing bottom or charm quarks which are

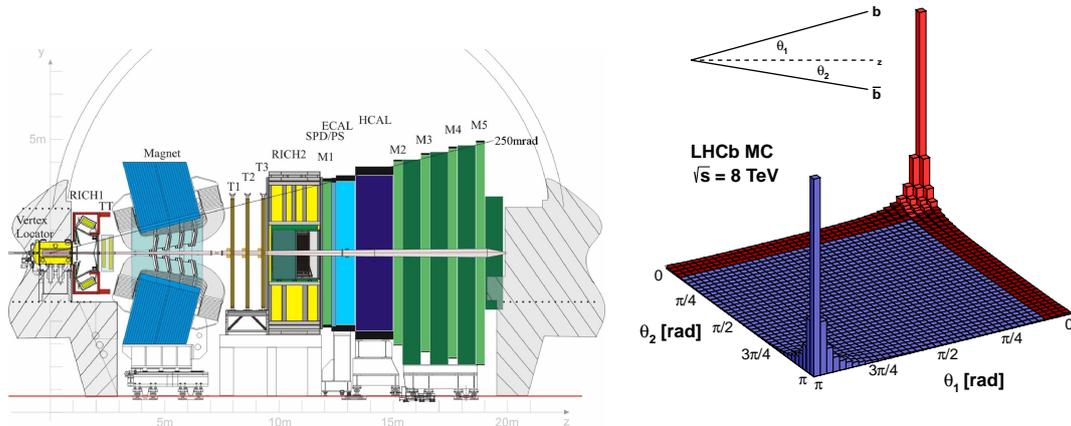


Figure 12: Schematic view of the LHCb detector [65] [66] on the left and the angular distribution between the produced $b\bar{b}$ pairs with respect to the beam axis for simulated proton-proton collisions [67] is shown on the right. Here the red marked region corresponds to the angular acceptance of the LHCb detector.

mainly produced as $q\bar{q}$ pairs in the forward region at the LHC as detailed in the right of Figure 12. Here forward means the direction along the beam pipe, which is denoted by the positive z -axis in the Figure. In this region heavy hadrons have a large boost which leads to a measurable distance from the interaction point before they decay, so that they can be reconstructed by the LHCb detector.

In order to do so, LHCb uses a large dipole magnet to deflect charged particles and measure their momentum. To reconstruct the decays of heavy flavoured hadrons, the primary pp interaction point and the decay vertex of unstable heavy particles has to be precisely measured, as well as an excellent momentum resolution and ability to identify the type of particle are needed. The following chapter discusses the technical aspects and sub-detectors which are necessary to study heavy hadrons in detail.

3.2.1 Tracking system

The tracking system of LHCb consists of different detectors: the Vertex Locator (VELO) [69,70] surrounding the primary pp interaction point, two tracking stations in front of (Trigger Tracker or TT) and three tracking stations (T1-T3) [71,72] behind the dipole magnet [73] as displayed before in Figure 12.

The **VELO** is reconstructing the primary vertex of pp collisions as well as secondary vertices from precise measurements of reconstructed particle trajectories (track) coordinates close to the interaction point. It consists of 21 stations along the beam direction, each of them made of two semi-circular shaped halves of silicon-strip detectors as shown in Figure 13. The R-type sensor provides a radial measurement whereas the ϕ -type measures the azimuthal angle, the z -position is obtained from the different sensor stations along the beam line. When charged

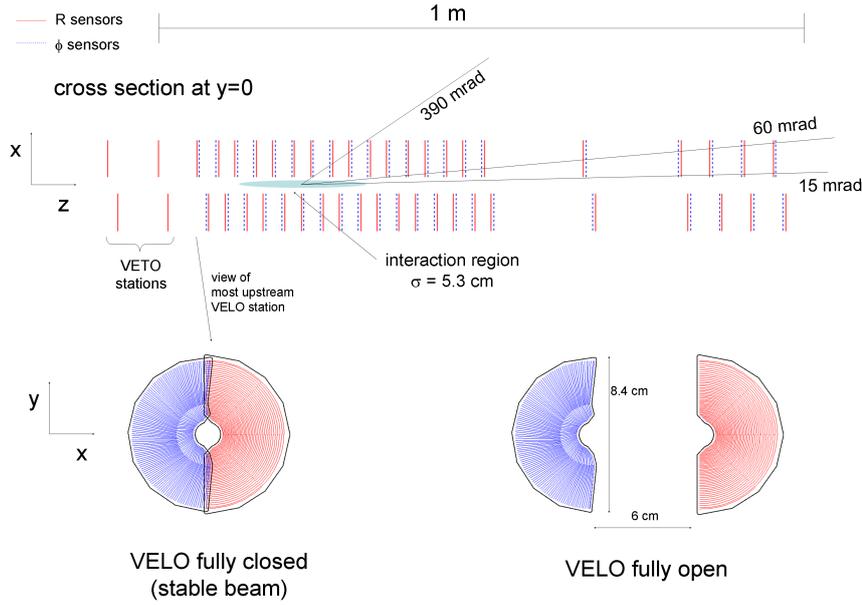


Figure 13: Schematic view of the VELO subdetector [68]. The arrangement of the different station is shown on top as well the two sensor types and their locations for the different positions is displayed in the bottom.

particles are passing through the semi-conductive silicon sensors they create electron-hole pairs which are measurable as pulses at the electrodes. The two detector halves are movable and have a minimal radial distance of 7 mm to the LHC beam only when stable beams are declared. During the injection phase of the LHC the VELO is moved out of its nominal position to protect the detector. As shown in Figure 13 there are additional sensors placed in the negative z -direction which are used to veto events with a very large number of produced tracks. The VELO performance depends on the number of reconstructed tracks and their corresponding momenta [70]. As a baseline for a vertex with 25 reconstructed tracks the primary pp interaction point, the so-called *primary vertex*, can be reconstructed with a resolution of $13 \mu\text{m}$ in the transverse plane and $71 \mu\text{m}$ along the z -axis. In order to reconstruct decay vertices the so-called *impact parameter* (IP) is used, it refers to the minimal distance of a charged track with respect to the PV. Therefore selecting tracks with a large IP identifies them to come from a secondary vertex. The IP can be reconstructed with a resolution of less than $35 \mu\text{m}$ for particles with a transverse momentum greater than $1 \text{ GeV}/c$. The angular acceptance of the detector is determined by the solid angle of at least three hits in the sensors $1.6 < |\eta| < 4.9$, which is given in terms of pseudo-rapidity $\eta = -\ln \tan \theta/2$, where θ is the spherical angle with respect to the beam axis.

The **magnet** is needed to measure the momenta of the charged particles via the curvature of a particle trajectory inside the magnetic field. In LHCb a dipole magnet is used which is aligned with the y -axis as shown in Figure 12

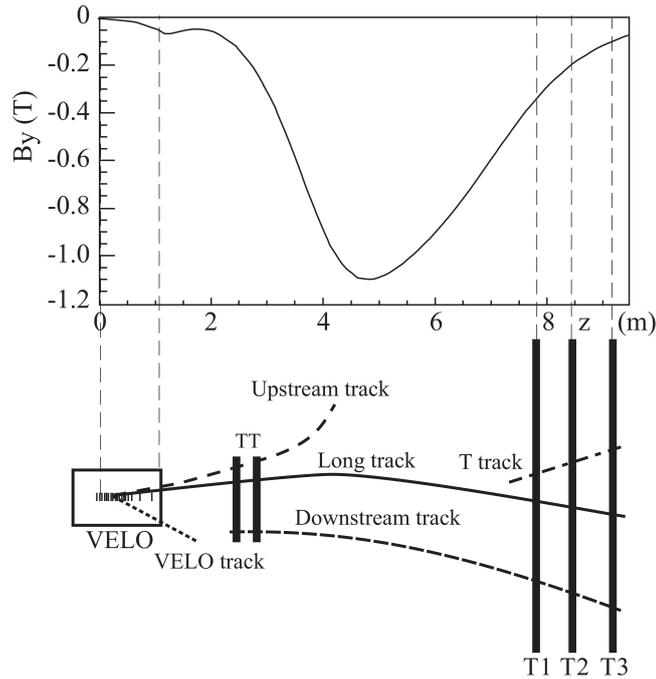


Figure 14: Schematic illustration of the main B -field component B_y as a function of the z position as the beam line. This is compared to the position of tracking detectors in the bottom together with the various track types: long, upstream, downstream, VELO and T tracks [64].

such that the magnetic field is along this axis. The magnet polarity is regularly reversed in order to control systematic effects from potential detector response asymmetries. Figure 14 shows the magnetic field of the dipole as a function of z . Particles traversing the whole detector feel an integrated magnetic field up to 4Tm. Because the magnetic field inside the VELO is negligible, no momentum measurement can be extracted from the VELO hits. Therefore additional tracking stations are located in front of and behind the magnet, as shown as well in Figure 14, those will be described in more detail in the following.

The tracking stations before the magnet, the **TT**, consist of four layers of silicon micro-strip detectors which are arranged in two stations TTa and TTb as displayed in Figure 15 on the left. The strips of the inner layers are tilted by $\pm 5^\circ$ in opposite direction around the z -axis to obtain a more precise position measurement along the y -axis, whereas the outer ones are aligned with the y -axis. It has a strip pitch of $183 \mu\text{m}$ which gives a spatial hit resolution in x of around $50 \mu\text{m}$. The TT is used to measure very low momentum tracks of about $2 \text{ GeV}/c$ which are bent out of the detector due to the magnetic field or tracks of long-living neutral hadrons decaying outside the VELO volume such as K_s^0 -mesons or Λ^0 -baryons. The TT detector covers the full angular acceptance and its measurements help improving the momentum resolution of the LHCb detector by further constraining the slope of the particle trajectories before the magnet.

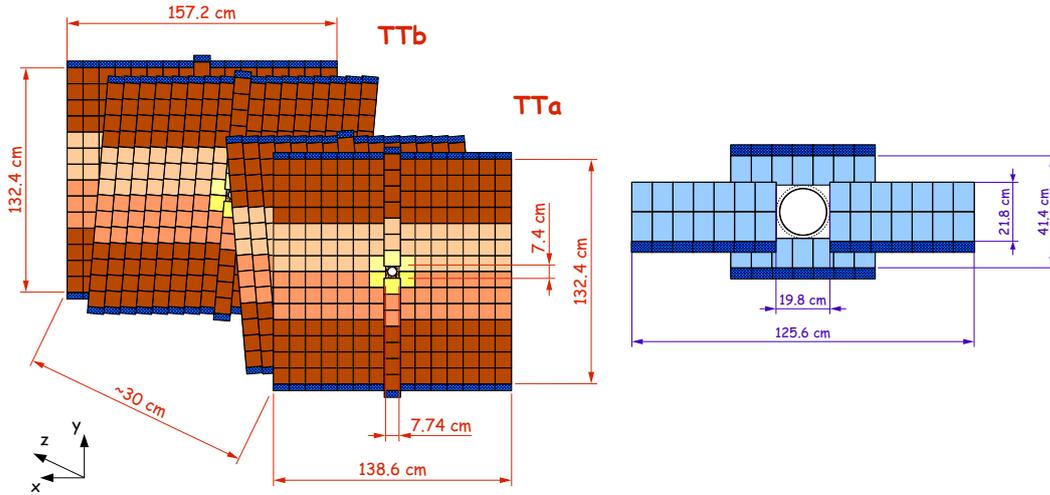


Figure 15: Schematic view of the four TT layers [74] on the left as well as one IT layer on the right [75].

The three main tracking stations **T1-T3** are installed directly after the magnet, where each station consists of four layers which are made of two different technologies. The inner detector area of 120 x 40 cm gets the highest particle flux and is also made of a silicon strip detector with a similar technology as the TT. It is the so-called *Inner Tracker* (IT) with the same high spatial resolution of 50 μm and is shown on the right of Figure 15. The outer region is covered by the so-called *Outer Tracker* (OT) with much smaller particle multiplicities and therefore coarser granularity. It is a gaseous detector made of straw tubes with 4.9 mm diameter which operate as proportionality counters. When charged particles are passing through the tubes they ionize the gas atoms within, the resulting electrons drift towards the anode where a measurable signal is collected. The three stations of the OT consist of 4 layers each, whereas a layer is made of two rows of straw tubes each as shown in Figure 16. The inner layers are tilted by $\pm 5^\circ$ around the z -axis as for the TT in order to determine the y -position. The layers have a pitch of 5.25 mm and a resolution of less than 200 μm is achieved perpendicular to the tubes. The OT covers an area of 5 x 6 m^2 and is together with the IT crucial for a precise momentum measurements of tracks.

The hits from the tracking sub-detectors are combined to tracks by using different algorithms, as further explained in [77]. There are different track types according to which sub-detector finds hits, as shown before in Figure 14. For this analysis only so-called *long tracks* are used, which traverse the full tracking system. They consist of at least hits in the VELO and the tracking stations behind the magnet and optionally also in the TT. Long tracks have the highest momentum resolution with $\delta p/p$ of 0.5% for low momentum (20 GeV/c) to 0.8% at 200 GeV/c [77]. The average number of reconstructed tracks depends on the run condition but is of order of 100 tracks per event where 1.5-1.7 PVs are on average

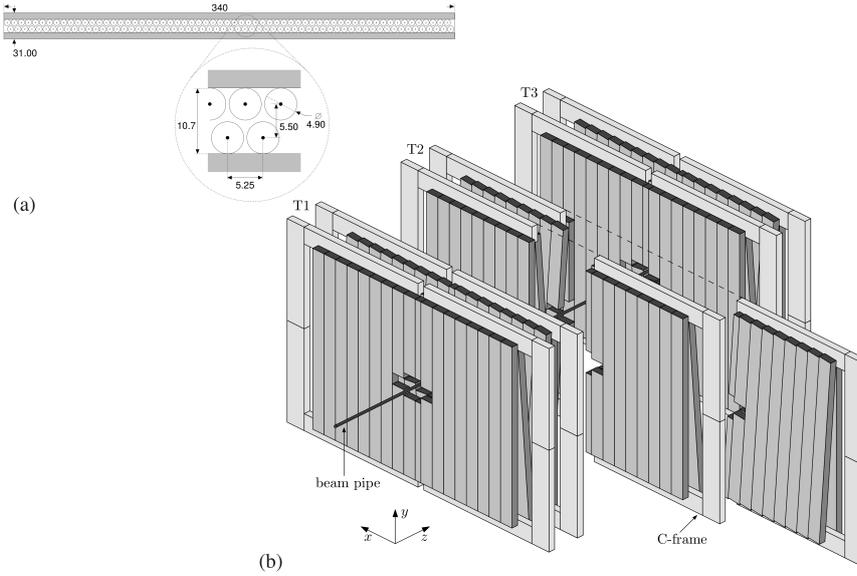


Figure 16: Schematic view of the cross section of a single OT layer (a) compared to the OT straw tube modules arranged in three stations (b). [76].

reconstructed per proton bunch crossing. In general reconstruction artefacts can lead to so-called *ghost tracks* if unrelated hits are combined or to so-called *clones* if two tracks share more than 70% of their hits. In order to select good-quality tracks, information from the tracking system is combined to compute a per track ghost probability (Track `GhostProb`) and a track fit χ^2 per degrees of freedom (Track $\chi^2/\text{n.d.f.}$) which are then usable in the offline analysis.

3.2.2 Particle identification

It is important in flavour physics to discriminate between the different particle species. Especially for semileptonic decays which are not fully reconstructed it is essential to identify the different particle types of the decay products precisely and to suppress decays with similar kinematic or topology as much as possible. The LHCb detector has a dedicated particle identification (PID) system that allows to distinguish between kaons, muons, charged pions, protons, electrons and photons. Information from several sub-detectors are combined into this PID system which will be further discussed in the following.

Two **Ring-Imaging Cherenkov detectors** (RICH1 and RICH2) [78] are used to identify charged hadrons. They are based on Cherenkov radiation which occurs when a charged particle traverses a material with a velocity greater than the speed of light in that medium. Photons are emitted in form of a cone with a specific angle depending on the velocity β of the particle and the refractive index of the material n given by $\cos \theta = 1/\beta n$. From measuring the angle θ the velocity of the particle can be determined and by matching the RICH signal to a track, the momentum estimate can be related to its mass such that the particle species can be identified. The left of Figure 17 shows the Cherenkov angle as a

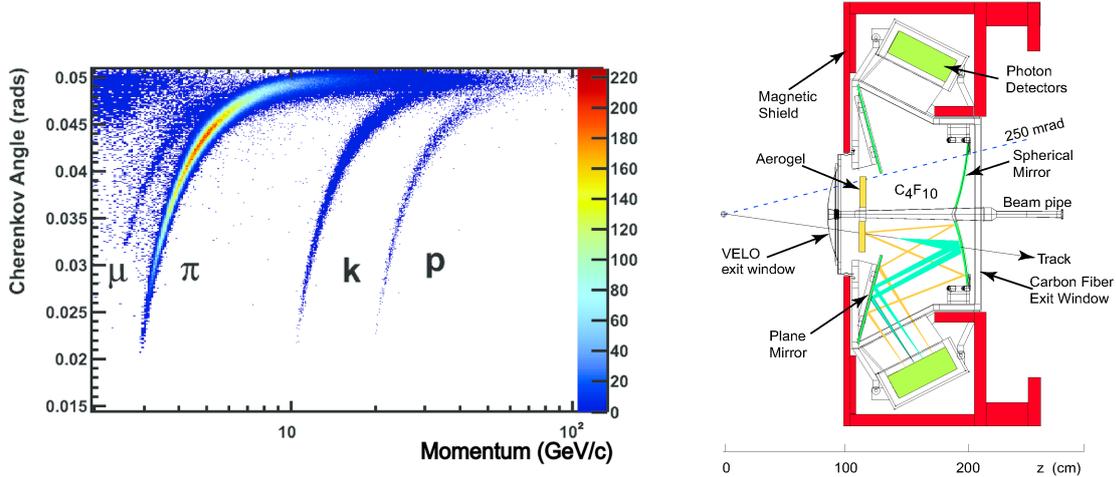


Figure 17: The measured Cherenkov angle θ as a function of the particle momenta [78] on the left and the schematic view of the RICH1 detector [64] on the right.

function of the particle momentum for isolated tracks where the different particles appear as distinct bands.

The two RICH detectors are placed before and after the magnet, both have a similar setup consisting of a radiator material, a mirror system and photon detectors. The radiator material has a different refractive index, which leads to a large momentum range covered. The RICH1 uses aerogel and C_4F_{10} radiators which cover low momentum particles in the range between 2 and 50 GeV/c. The RICH2 detector located after the magnet contains the gas CF_4 which allows to discriminate charged hadrons up to 100 GeV/c in momentum. The mirrors guide the Cherenkov light away from the beampipe to photon detectors placed outside the detector acceptance which detect the positions of the light rings. The right of Figure 17 shows the different components of the RICH1 detector.

In practise there are many overlapping rings in the RICH detectors which have to be matched to the traversing tracks and identified. In order to do so a likelihood for the measured rings is constructed by assigning the observed patterns different mass hypotheses. Differences between the likelihood for the respective particle are then used to discriminate between the particle types, where as a reference the likelihood of the pion hypothesis is used. Figure 18 shows the kaon identification efficiency compared to the kaon-pion misidentification rates as a function of momentum using different selection criteria of the logarithmic likelihood difference (ΔLL) at 8 TeV.

The main task of the **Calorimeter system** [82] is to measure the energy of charged and neutral particles. The LHCb Calorimeter system consists of four sub-systems as shown in Figure 19, which allow to discriminate light hadrons, electrons and photons. The calorimeters are constructed as sampling calorimeters where absorber layers and scintillating material are alternating each other. The absorber material is needed to induce electromagnetic or hadronic showers while the scintillators detect those. Since the number of produced particles in a shower

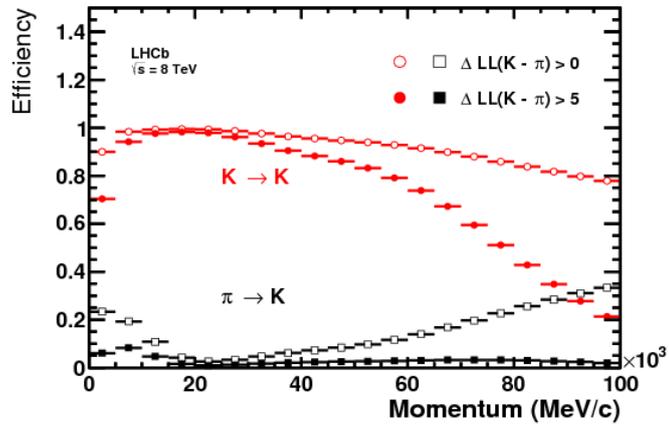


Figure 18: Kaon identification performance as a function of momentum for two different cuts in the log-likelihood of the RICH [79].

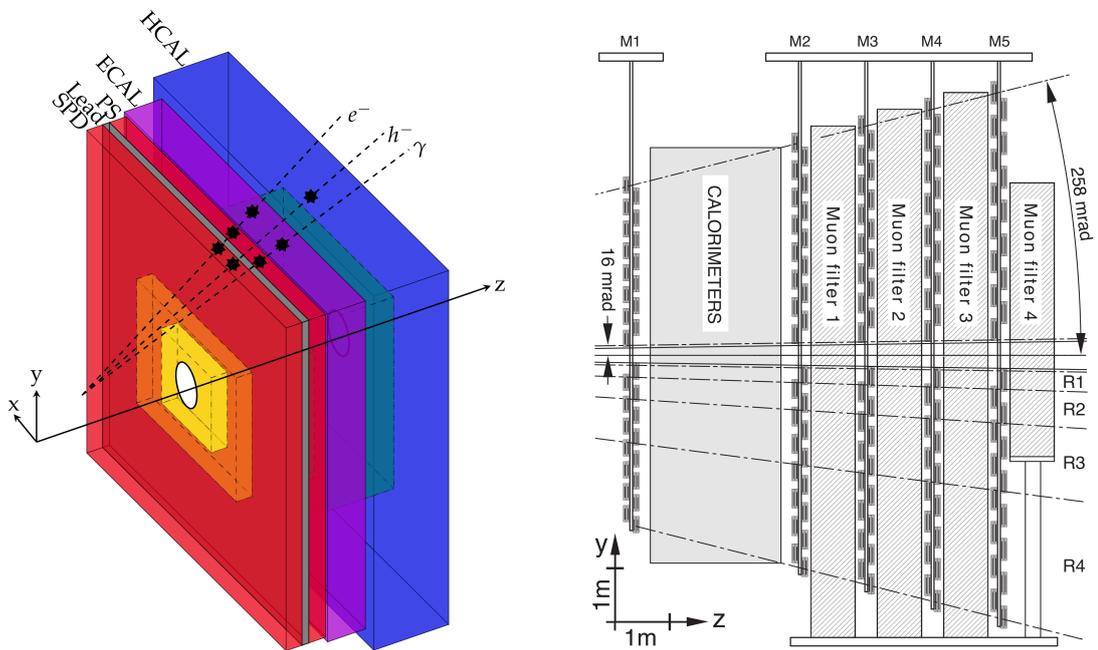


Figure 19: Schematic view of the calorimeter system [80] on the left and the muon system [81] on the right.

is proportional to the energy of the primary particle those sampling calorimeters provide an energy measurement with a resolution of $\Delta E/E \propto 1/\sqrt{E}$.

A *Scintillating Pad Detector* (SPD) and a *Pre-Shower* (PS) detector are separated by a lead plate from the *Electromagnetic Calorimeter* (ECAL). The SPD only produces signal for charged particles and such can distinguish between electromagnetic showers created by electrons and photons. The PS detector can separate electrons from light hadrons from their different shower shape inside the lead layer. The ECAL measures the energy of electrons and photons as they lose their energy via electromagnetic showers which are fully absorbed. It can be also used to reconstruct neutral pions. The *Hadronic Calorimeter* (HCAL) is located behind the ECAL, inside hadronic showers of neutral and charged hadrons are created and detected. The calorimeters play a crucial role for the online selection of events in the trigger as they can be read out at the maximum LHC collision rate of 40 MHz and can be used to select final states with high energetic electrons, photons or hadrons.

The **Muon system** [81, 83] consists of five stations M1-M5 as shown in Figure 19, those are used to identify muons. At LHC energies muons are minimum ionizing particles and as such can traverse the whole detector in contrast to other particle species. Therefore the muon stations are placed behind the calorimeters which is about 15m away from the interaction point, where mostly all other particles are already stopped by interactions with the detector material. The muon stations are made of multi-wire proportional chambers where the passing muon ionises gas atoms in the chamber producing charges which are amplified and collected. In between the muon stations M2-M5 there are 80 cm thick layers of iron stopping high energy hadrons which passed the calorimeter system. The muon chambers identify high energy muons with a large efficiency, together with their fast readout at 40 MHz they are also used as an important ingredient for the trigger system. Especially the muon station M1 which is placed in front of the calorimeter is helpful for extrapolating trajectories back to the PV and it largely improves the transverse momentum resolution in the trigger.

As briefly described for the RICH detectors, every sub-detector used for particle identification calculates per-track likelihoods for different particle hypothesis. Particle identification algorithms then combine these information in order to provide the optimal discrimination between particles. In general the change in logarithmic likelihood between two different mass hypotheses X and Y for a given particle is denoted as

$$\Delta \log \mathcal{L}_{XY} = \log \mathcal{L}(X) - \log \mathcal{L}(Y) \quad (48)$$

and can be used to discriminate between the two hypotheses. Those differences are then linearly combined by particle identification algorithms. Global identification variables are defined for every particle species in order to provide the optimal discrimination between different particles [77]. In order to identify muons a simple algorithm [84] is used which tries to match reconstructed tracks in the tracking

stations to hits in the muon stations. If this was successful a binary flag, the so-called `isMuon` variable, is set which is a powerful criteria to select muons.

3.2.3 Trigger system and data flow

As mentioned earlier the LHC collides proton bunches at a rate of 40 MHz. The LHCb detector can not be fully read out and collect the information from all sub-detectors at this rate so far, this will change for the coming upgrade. Not all of these events are of interest as approximately only in one out of 1000 pp collisions a $b\bar{b}$ quark pair is produced inside the LHCb acceptance. Those events must be identified and separated from other non-interesting events as well as recorded, all these tasks are done by the LHCb trigger system [85–87].

The trigger system of LHCb is made of three consecutive steps, at each step the output rate is reduced while more sub-detector information is used. This allows for a more complex selection criteria at each step. As explained earlier in this section, weakly decaying heavy mesons have similar signatures with a well separated secondary vertex and a corresponding large impact parameter (IP) due to their finite lifetime. Those distinct topologies are used as a discriminant in the trigger.

At the first trigger level, the so-called level zero or **L0**, hardware information based on the calorimeter and muon system is used, as they are the only sub-detectors which can be read out at 40 MHz. This L0 hardware trigger searches for a high transversal energy component (E_T) deposited in the calorimeters or for high transverse momentum (p_T) signatures in the muon stations. Typical calorimeter threshold for the HCAL are $E_T > 30$ GeV, while thresholds for electrons and photons are lower $E_T > 2$ GeV. For muon triggers the threshold are $p_T > 1.2$ GeV/ c for single muons while $\sqrt{p_T^1 p_T^2} > 1.1$ GeV/ c for a pair of muons. Those selections reduce the total rate to about 1MHz, only events passing the L0 trigger step are send further to the second step of the trigger.

The first stage of the software based high level trigger, the so-called **HLT1**, uses information from the tracking system to reconstruct primary vertices (PV) and tracks. This allows to identify high momentum or transverse momentum particles with a large IP which are likely to come from a relatively long-living B or charm hadron. Also events containing two muons forming a good quality vertex with a high invariant mass are selected. The HLT1 reduces the event rate to 80 kHz.

The second stage of the high level trigger, the so-called **HLT2**, is able to perform a full event reconstruction based on the information of all sub-detectors. Two different trigger categories are used: *inclusive* and *exclusive* signatures of multi-body B - and charm-hadron decays. Inclusive sequences select generic topological properties of two-, three- and four-body decays from displaced B hadrons, based on a large IP of tracks with a good-quality secondary vertex as well as loose requirements on individual and combined masses and momenta. Exclusive selections are designed for a specific final states, where the complete reconstruction of decay chains is performed. The HLT2 output rate is 5kHz and all events passing HLT2 are written to disk and stored permanently. In order to make optimal use of the available resources, HLT1 was running asynchronously

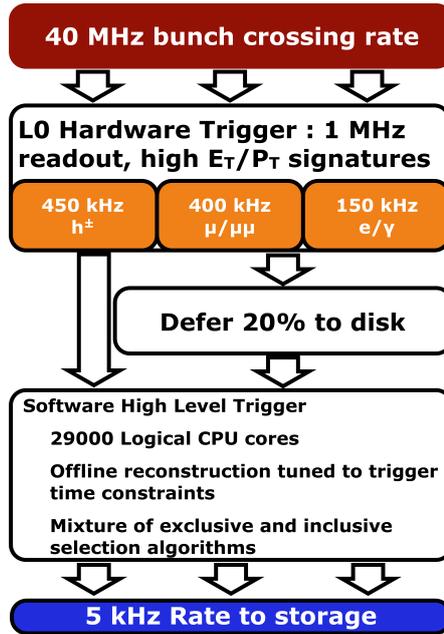


Figure 20: Schematic Layout of the LHCb trigger strategy in Run I [88].

with the pp collisions in 2012. Therefore during this data taking 20% of the L0 output was temporarily buffered to disk to be processed by the HLT1 when the LHC was not providing stable beam as shown in Figure 20. A more detailed discussion of the trigger strategy used this thesis is described later in Section 6.1.

After this online processing step provided by the trigger, there is an **offline processing** step of the data. The following offline step is needed to achieve the best reconstruction quality of the online-selected data for data analysis. In HLT2 the reconstruction algorithms have to be simplified in some cases to meet the necessary timing requirements, but in general those algorithms are close to the offline reconstruction. In order to do so all information from the sub-detectors is saved to be able to do a separate offline reconstruction, which is done centrally by the LHCb computing team performed with the BRUNEL software package [89] based on the GAUDI framework [90]. During that reconstruction measured hits of the tracking stations are used to reconstruct charged tracks and particle identification information from the PID system is added. As a second step of the online processing a standardised selection sequence is performed using the so-called DAVINCI software package [91]. It is also used as the physics analysis software to combine reconstructed tracks to secondary vertices as well as to fit those. Also combined objects such as charm or B hadron candidates are constructed and based on geometric and kinematic variables specific decay topologies of interest are selected. This step of the offline processing is also run centrally as it is computational expensive and the analysts are provided with preselected data sets where specified selection requirements were applied using the so-called *stripping* selections. The output of the stripping is a file containing events passing these selections with a much reduced size. Those can then be processed by individual

analysts applying final selection cuts to obtain the decay of interest. The stripping lines used in this thesis as well as the specific selection cuts applied are further discussed in Section 6.2.

3.2.4 Simulation and data samples

Simulated data samples of the decay processes of interest as well as their detector response are crucial ingredients to be able to perform an analysis of experimental data. Those Monte Carlo (MC) data sets are invaluable to test analysis strategies as well as to study and understand detection efficiencies and assess systematic uncertainties, which are key aspects of an analysis. Due to their importance it is indispensable that the generated MC samples emulate the real data as much as possible. The full MC simulation is built on the GAUDI framework, simulated events are generated using the GAUSS [92] software package. The generation process of MC data sets are split into two parts, in the first proton-proton collisions are simulated with the event generator PYTHIA [93] with a LHCb [94] specific configuration, subsequent decays of hadronic particles are performed via the EVTGEN [95] library in which final-state radiation is generated using PHOTOS [96]. The second step is more time consuming as it involves the interaction of the generated particles with the LHCb detector as well as its response. This is realised using the GEANT4 4 [97] toolkit which is further described in Ref. [98]. The output of the detector response is then digitized by the BOOLE [99] software package and the trigger emulation is performed using the MOORE [100] package in the LHCb software. The simulated samples are generated under the same run conditions and processed in the same way as the real data collected by the LHCb experiment.

In this thesis simulated samples are mainly used to estimate the reconstruction and selection efficiencies as well as to extract the line shape of the distributions used to fit the data. The generated MC sampled of signal and possible background used for this thesis are summarised in Table 5. The analysis performed in this thesis is using part of Run I data recorded in 2012 by LHCb, which corresponds to an integrated luminosity of 2 fb^{-1} of data.

Decay	Type	Size
Samples for signal fit		
$B_s^0 \rightarrow K^+ \mu^- \nu$	Signal	6M
$B_s^0 \rightarrow K^{*+}(\rightarrow K^+ \pi^0) \mu \nu$	Background	4M
$B_s^0 \rightarrow K_2^{*+}(1430)(\rightarrow K^+ \pi^0) \mu \nu$	Background	4M
$B_s^0 \rightarrow K_0^{*+}(1430)(\rightarrow K^+ \pi^0) \mu \nu$	Background	4M
$B^0 \rightarrow J/\psi(\rightarrow \mu\mu) K^{*0}(K^+ \pi^-)$	Background	10M
$B^+ \rightarrow J/\psi(\rightarrow \mu\mu) K^+$	Background	20M
$B^+ \rightarrow J/\psi(\rightarrow \mu\mu) K^{*+}(K^+ \pi^0)$	Background	20M
$B^+ \rightarrow J/\psi(\rightarrow \mu\mu) \phi$	Background	120M
$B^+ \rightarrow c\bar{c}(\rightarrow \mu\mu) K^+ X$	Background	2M(filtered)
$B^0 \rightarrow \pi \mu \nu$	Background	4M
$B^0 \rightarrow \rho^+(\pi^+ \pi^0) \mu^- \nu$	Background	4M
$B^+ \rightarrow \rho \mu^+ \nu$	Background	5M
$\Lambda_b^0 \rightarrow p \mu \nu$	Background	5M (FF from LQCD)
$\Lambda_b^0 \rightarrow p \mu \nu$	Background	10M (FF from LCSR)
inclusive $b \rightarrow c$ decays		
$H_b \rightarrow H_c(\rightarrow K^+ \mu^- X) X'$	Background	144k (filtered)
$H_b \rightarrow K^+ \mu^+ X$	Background	120k (filtered)
$H_b \rightarrow H_c(\rightarrow K^+ X) \mu^- X'$	Background	4.96 M (filtered)
Samples for normalisation fit		
$B_s^0 \rightarrow D_s^+ \mu \nu_\mu X$ cocktail	Normalisation	6M
$B^+ \rightarrow D_s^{*+} D^*$	Background	5M
$B_s^0 \rightarrow D_s^{*+} D_s^{*+}$	Background	5M
$B_s^0 \rightarrow DD$	Background	5M

Table 5: Summary of simulated samples used for this thesis.

4 The Analysis strategy

This chapter describes the strategy followed in this thesis to measure $|V_{ub}|/|V_{cb}|$ with LHCb and explains why the particular decays of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ were chosen for this measurement.

As outlined before $|V_{ub}|$ and $|V_{cb}|$ can be determined from either exclusive or inclusive decays of B -hadrons. Here the exclusive approach is employed due to the large background contamination from many different particles produced at the same time at a hadron machine in contrast to the relatively clean environment at e^+e^- -machines from B -factory experiments BaBar and Belle. Due to their B hadron production from mostly $\Upsilon(4S)$ decays the beam energy and initial state is precisely known and the fully hermetic detector can be used to constrain the kinematic of decays to perform inclusive measurements by summing over all possible hadronic final states. Inclusive determinations of $|V_{ub}|$ in particular are not feasible at LHCb due to the large background from $|V_{cb}|$ decays which can be only separated from $b \rightarrow u$ transitions at the kinematic endpoint of the lepton energy spectrum reconstructed with a high precision by the B -factories.

The advantage of the hadron machine is the higher production cross-section of $b\bar{b}$ -pairs compared to e^+e^- -machines, which leads to a larger data sample and thus allows to study such small $b \rightarrow u$ transition of the order of 10^{-4} . In addition all different kind of B -hadrons can be produced at a hadron collider due to the sizable fragmentation fraction of b -quarks into B_s^0 -mesons and Λ_b -baryons [101]. Therefore other exclusive decays can be explored to measure $|V_{ub}|$ and $|V_{cb}|$ as already performed in $\Lambda_b \rightarrow p\mu\nu$ decays for the first time at a hadron collider in Reference [5] and in this thesis using B_s^0 -decays. These provide important complementary input to the long-standing tension between exclusive and inclusive $|V_{ub}|$ and $|V_{cb}|$ measurements performed by the B -factories measured from B^0 and B^+ decays.

For a precise exclusive determination of $|V_{ub}|$ and $|V_{cb}|$ small form factor uncertainties from either Lattice QCD or LCSR are necessary, thus ground state hadrons in the final state are preferred. Since the 'golden' channel $\bar{B}^0 \rightarrow \pi^+ l^- \bar{\nu}$ for the B -factories suffers from a high pion background at the LHC, $\Lambda_b \rightarrow p\mu\nu$ decays were used for the first $|V_{ub}|/|V_{cb}|$ measurement where the proton makes experimentally a much more distinctive final state which less background contributing [102]. In contrast to that the decay of interest in this thesis $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ suffers from higher background contamination since final state kaons can be produced in decays from many more B hadrons. In addition B_s^0 mesons have a smaller production fraction than Λ_b baryons, such that they are more abundant at the LHC [101]. Theoretically semileptonic B_s^0 decays are more advantageous than B^0 or B^+ mesons due to the larger mass of the valence s-quark which makes LQCD calculations of form factors less computationally expensive and thus more precise [50, 103, 104]. Therefore possible allowing for a more precise determination of $|V_{ub}|$ and $|V_{cb}|$. Taking into account the large data sample collected by LHCb

in 2012 with over 200 billion $b\bar{b}$ -pairs together with fragmentation fraction into B_s^0 -mesons of around 8% and the expected rate of the $b \rightarrow u$ transition of the order of 10^{-4} it is therefore possible to reconstruct a large number of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ candidates and to extract $|V_{ub}|$ from it.

To precisely determine $|V_{ub}|$ alone from the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ branching fraction would require a precise measurement of the $b\bar{b}$ cross-section and the B_s^0 fragmentation fraction at the LHC as well as the precise measurement of the integrated luminosity. The current precision of these measurements is not enough to perform a competitive extraction of $|V_{ub}|$, therefore the decay of interest has to be normalised. The decay $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ is chosen as a normalisation channel as it is very similar to the signal channel, thus many experimental uncertainties cancel in the ratio and in particular the total production rate of B_s^0 mesons. A measurement of the ratio of branching fraction of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ is performed in different regions of q^2 of the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay as explained in Section 2.5 due to the different form factor predictions from LQCD and LCSR. In contrast for the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ the total available q^2 range is used as the form factor predictions were precisely determined for the full q^2 spectrum from Lattice QCD recently. Experimentally the ratio of branching fractions is measured as

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} = \frac{N(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{N(B_s^0 \rightarrow D_s^- (\rightarrow K^+ K^- \pi^-) \mu^+ \nu_\mu)} \times \epsilon_{rel} \quad (49)$$

$$\times \mathcal{B}(D_s^- \rightarrow K^+ K^- \pi^-),$$

where $N(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)$ and $N(B_s^0 \rightarrow D_s^- (\rightarrow K^+ K^- \pi^-) \mu^+ \nu_\mu)$ are the measured yields for the respective signal and normalisation decay. Those are determined from fitting the corrected mass distribution after all selection cuts are applied to the selected $K^+ \mu^-$ and $D_s^- \mu^+$ candidates in the corresponding q^2 bin.

The relative efficiency for selecting the two modes ϵ_{rel} is given as

$$\epsilon_{rel} = \frac{\epsilon(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{\epsilon(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}. \quad (50)$$

It is determined from simulated events after data driven corrections are applied to correct for possible differences between simulation and data. In addition the branching fraction of the decay $D_s^- \rightarrow K^+ K^- \pi^-$ is taken from the PDG [12] to be $\mathcal{B}(D_s^- \rightarrow K^+ K^- \pi^-) = (5.45 \pm 0.17)\%$, which includes measurements from BaBar, Belle and CLEO.

Combining this measurement with form factor predictions from Lattice QCD and LCSR thus allows to directly determine $|V_{ub}|^2/|V_{cb}|^2$ following this formula

$$\frac{|V_{ub}|}{|V_{cb}|} = \left(\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \times R_{FF} \right)^{\frac{1}{2}}, \quad (51)$$

where \mathcal{B} stands for the corresponding branching fraction. R_{FF} is the ratio of form factors which is obtained from integrating the form factor dependent predicted

decay width using Equations 42 and 46,

$$R_{\text{FF}} = \frac{\int_0^{q_{\text{max}}^2} \frac{d\Gamma(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{dq^2} / |V_{cb}|^2 dq^2}{\int_0^{q_{\text{max}}^2} \frac{d\Gamma(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{dq^2} / |V_{ub}|^2 dq^2}, \quad (52)$$

for different ranges of q^2 corresponding to the kinematic of the decay.

The total branching fraction of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ can be measured by extrapolating the measured branching fraction ratio to the full q^2 region using the following formula

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = & \tau_{B_s^0} \times \frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \times |V_{cb}|^2 \\ & \times \int_{0 \text{ GeV}/c^2}^{q_{\text{max}}^2} \frac{d\Gamma(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{dq^2} / |V_{cb}|^2 dq^2. \end{aligned} \quad (53)$$

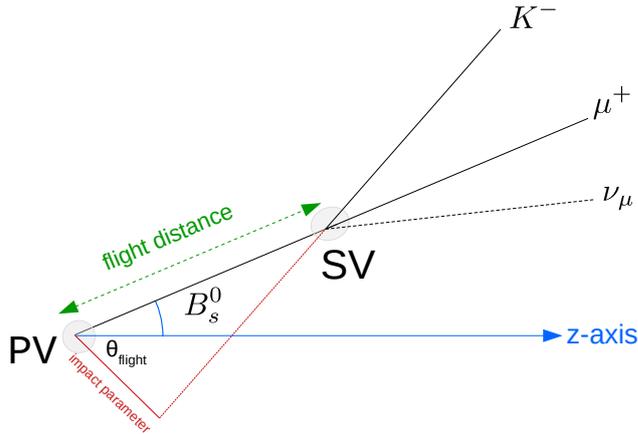


Figure 21: Sketch of the decay topology of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decays. The flight distance of the B_s^0 is indicated together with the impact parameter of the K^+ meson with respect to the primary vertex. The polar angle of the flight vector θ_{flight} with respect to the z -axis is also defined.

5 Methods and kinematics

The decays of interest $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ for the signal and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ for the normalisation channel are both semileptonic B_s^0 decays, which can only be partially reconstructed due to the missing neutrino. Thus the invariant mass of the B_s^0 meson cannot be reconstructed from the four momentum of its decay products. This presents a big challenge at a hadron collider compared to the e^+e^- B-factories Belle and BaBar. The kinematic of the B meson is not determined by the beam energy and can thus not be used to kinematically constrain the neutrino. Different methods and techniques were developed to overcome this problem which are explained in the following section. Section 5.1 discusses the kinematic of semileptonic decays using the so-called corrected mass and 5.2 explains the tools used to reconstruct the missing neutrino. Other methods exploited by this analysis such as the *sPlot* technique and Boosted Decision Trees are introduced in Section 5.3 and Section 5.4.

5.1 Corrected mass

LHCb has a very good precision of the vertex reconstruction as explained in Section 3 which allows to distinguish the production and decay vertex of the B_s^0 meson. The B_s^0 signal decay vertex as the secondary vertex (SV) can be reconstructed from its charged decay products, the kaon and muon. From the pp interaction point as the B_s^0 production vertex or primary vertex (PV) and its decay vertex the B_s^0 flight direction can be determined. Figure 21 shows a sketch of the resulting decay topology. From the symmetry of the decay the transverse momentum of the missing neutrino can be reconstructed as shown in Figure 22, where the event is rotated such that the B_s^0 flight direction is in z -direction. Exploiting the symmetry of the decay the transverse momentum of the neutrino perpendicular to the B_s^0 flight di-

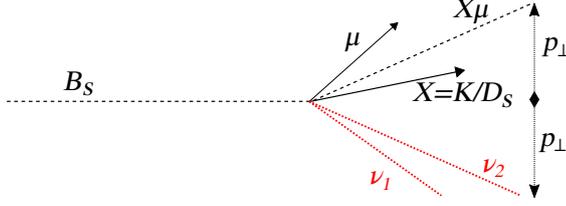


Figure 22: The neutrino p_{\perp} component determined from momentum conservation with respect to the B_s^0 flight direction for either the signal or the normalisation channel. Figure taken from [9].

rection must balance the sum of the transverse momenta of the reconstructed B_s^0 daughters:

$$\vec{p}_{\perp}(X\mu) = -\vec{p}_{\perp}(\nu_{\mu}), \quad p_{\perp} \equiv |\vec{p}_{\perp}(X\mu)| \quad (54)$$

where X is either the kaon for the signal or D_s^+ for the normalisation channel. The neutrino momentum parallel to the B_s^0 flight direction remains unknown and will be further discussed in Section 5.2. With that a lower limit of the mass of the B_s^0 meson can be calculated, the so-called corrected mass. It is defined as

$$m_{corr} = \sqrt{m_{X\mu}^2 + p_{\perp}^2} + p_{\perp}, \quad (55)$$

here $m_{X\mu}^2$ is the invariant mass of the reconstructed B_s^0 daughters. If a neutrino is the only missing particle the corrected mass peaks at the B_s^0 mass, which can be seen from the following derivation. In the rest frame of the B_s^0 , its mass can be written as:

$$\begin{aligned} m_{B_s^0} &= E_{vis} + E_{miss} \\ &= \sqrt{M_{vis}^2 + p_{\perp,vis}^2 + p_{\parallel,vis}^2} + \sqrt{M_{miss}^2 + p_{\perp,miss}^2 + p_{\parallel,miss}^2}, \end{aligned} \quad (56)$$

where the index vis denotes the visible, reconstructed particles and $miss$ stands for missing particles. If the missing particle is massless and using by Equation 54 of the perpendicular momentum balance and that in the B_s^0 rest frame $p_{\parallel,vis} = p_{\parallel,miss} = p_{\parallel}$, the B_s^0 mass becomes

$$m_{B_s^0} = \sqrt{M_{vis}^2 + p_{\perp}^2 + p_{\parallel}^2} + \sqrt{p_{\perp}^2 + p_{\parallel}^2}. \quad (57)$$

Neglecting the longitudinal component in B_s^0 rest frame gives the corrected mass m_{corr} as an approximation for the invariant B_s^0 mass. Therefore the corrected B_s^0 mass lies around its invariant mass of $m_{B_s^0} = 5366.88 \pm 0.17$ MeV [12] if only a neutrino is missing.

The corrected mass is lower for decays where massive particles are missing. Figure 23 shows the corrected mass distribution of the signal together with the normalisation channel reconstructed only through the $K^+\mu^-$ pair. This shows how powerful this variable is to distinguish $B_s^0 \rightarrow K^-\mu^+\nu_{\mu}$ from partially reconstructed background.

Even though the corrected mass is highly discriminating, there is a substantial

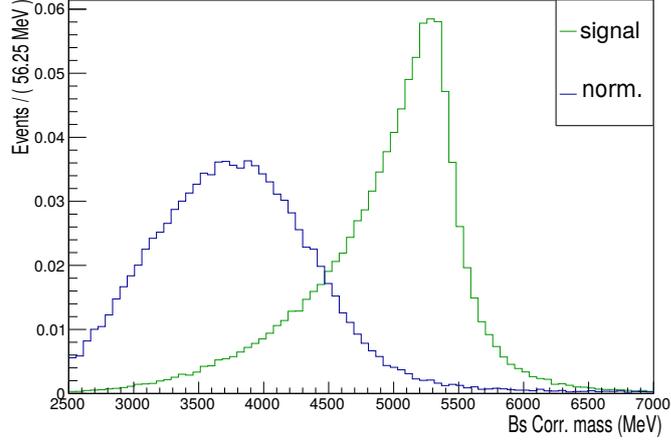


Figure 23: The B_s^0 corrected mass for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ in compared to the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ normalisation channel reconstructed as signal through $K^+ \mu^-$.

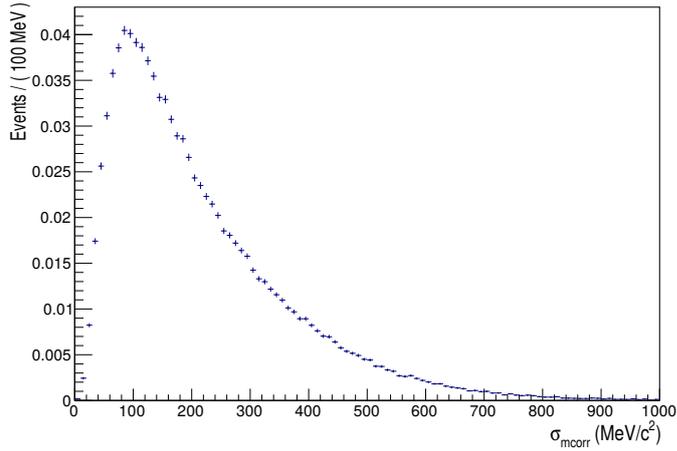


Figure 24: The corrected mass uncertainty for signal decays.

uncertainty associated with its reconstruction which results in a less prominent peaking structure for the signal. This can be seen from the tail towards lower corrected masses and gives a greater leakage of background into the signal region. The overlap with partially reconstructed backgrounds with only one additional missing massive particle is largest, therefore these backgrounds are the most dangerous ones. The uncertainty on the measurement of the B_s^0 flight direction results in a tail of the corrected mass above the B_s^0 mass. The uncertainty on the corrected mass gets contributions from an uncertainty of the visible mass measurement as well as from the uncertainty to determine p_\perp . Since the mass uncertainty is very small with respect to the latter one it is neglected in the following discussion,

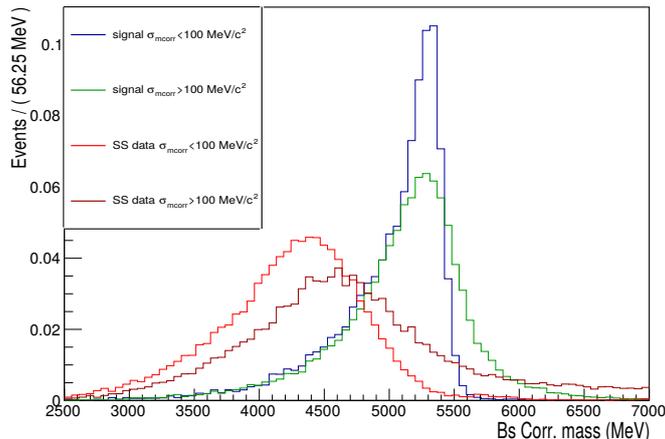


Figure 25: The corrected mass for signal decays and SS data selected with small and large corrected mass uncertainty.

therefore the uncertainty on the corrected mass is given by

$$\sigma_{m_{corr}} = \left(\frac{p_{\perp}}{\sqrt{m_{X\mu}^2 + p_{\perp}^2}} + 1 \right) \sigma_{p_{\perp}}. \quad (58)$$

For the total uncertainty of p_{\perp} several uncertainties contribute, the four momenta of the visible $p\mu$ pair and the imperfect knowledge of the B_s^0 flight direction. The latter one is dominant and results in a large uncertainty on the perpendicular momentum of the neutrino $\sigma_{p_{\perp}}$. It comes from the uncertainty on the positions of the primary and secondary vertex, which must be propagated through to the uncertainty in p_{\perp} . The propagation of uncertainties to $\sigma_{p_{\perp}}$ is not trivial, its full derivation can be found in Reference [102] which shows that uncertainties due to the four momenta can also be safely neglected. The distribution of the corrected mass uncertainty for the signal can be seen in Figure 24.

The resolution on the corrected mass can be improved by rejecting events with large corrected mass uncertainties. Figure 25 shows the signal corrected mass distribution compared to SS data selected with small and large corrected mass uncertainty. The SS sample consists of $K^+\mu^+$ data events including combinatorial, misidentified particles as well as CKM favoured $|V_{cb}|$ decays. The selection improves the separation between signal and SS data, as it removes the combinatorial component and signal events passing the selection have a significant sharper peak. The additional selection power gained by this selection results in a reduced systematic uncertainty in the fit to the corrected mass, which is further studied in Section 6.4.4.

5.2 Neutrino reconstruction

As detailed in Section 4 the $|V_{ub}|/|V_{cb}|$ measurement will be performed in two bins of q^2 , where q^2 is defined as the four momentum squared of the muon and neutrino combination $m_{\mu\nu}^2$. Therefore the neutrino 4-momentum needs to be reconstructed.

As explained in the previous section (5.1), the transverse component of the neutrino with respect to the B_s^0 flight direction p_\perp can be determined from the transverse momenta of the visible $K^-\mu^+$ pair. The neutrino momentum parallel to the B_s^0 flight direction p_\parallel can be determined up to two-fold ambiguity using the B_s^0 mass constraint:

$$(p_\nu + p_{X\mu})^2 = m_{B_s^0}^2 \quad (59)$$

where $X = K, D_s^+$ stands for the signal or normalisation channel and

$$\begin{aligned} p_\nu &= \left(\sqrt{p_\parallel^2 + p_\perp^2}, 0, -p_\perp, p_\parallel \right) \\ p_{X\mu} &= \left(\sqrt{p_\parallel^2(X\mu) + p_\perp^2(X\mu) + m_{X\mu}^2}, 0, p_\perp, p_\parallel(X\mu) \right). \end{aligned} \quad (60)$$

Here $m_{X\mu}$ is the visible mass of the $X\mu$ system and $p_\parallel(X\mu)$ the parallel momentum component with respect to the B_s^0 flight direction of the $X\mu$ pair. Inserting the four-momenta into Equation 59 and solving for the unknown p_\parallel leads to a quadratic solution for the parallel neutrino momentum:

$$p_\parallel = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad (61)$$

in which the parameters a , b and c are defined as

$$\begin{aligned} a &= p_\perp^2 + m_{X\mu}^2, \\ b &= p_\parallel^2(X\mu)(m_{miss}^2 - 2p_\perp^2), \\ c &= 4p_\perp(p_\parallel^2(X\mu) + m_{B_s^0}^2) - |m_{miss}^2|^2, \\ m_{miss}^2 &= m_{B_s^0}^2 - m_{X\mu}^2. \end{aligned} \quad (62)$$

After the preselection cuts are applied as described in detail in Section 6, around 20% of simulated $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ candidates have an unphysical solution for p_\parallel ($b^2 < 4ac$). This is due to detector resolution effects. These unphysical events correspond to a corrected mass region above the physical B_s^0 mass. In the previous V_{ub} analysis performed on $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ decays [5] these unphysical events were rejected. For this analysis those events are kept as they are needed to determine the shape for the combinatorial background for the signal fit which will be further explained in Section 9. Those unphysical events were further studied to test that they have the same q^2 resolution as the physical events and that they are evenly distributed along the selected q^2 bins. Therefore for those events the term $\sqrt{b^2 - 4ac}$ is set to zero, leading to a unique q^2 solution.

Using Equation 61 the q^2 variable can now be reconstructed up to a two fold ambiguity [105, 106]. A choice has to be made which of the two solutions is picked as the right one for this analysis and it is important to resolve this ambiguity without introducing a bias on q^2 . The easiest approach would be to pick randomly one solution, which is unbiased but also has a poor resolution in q^2 . Another approach is to use a linear regression to predict the B_s^0 momentum and then select the solution closest to the output of the regression. The regression method improves the resolution in q^2 as shown in Reference [107] and will be further discussed in the next section.

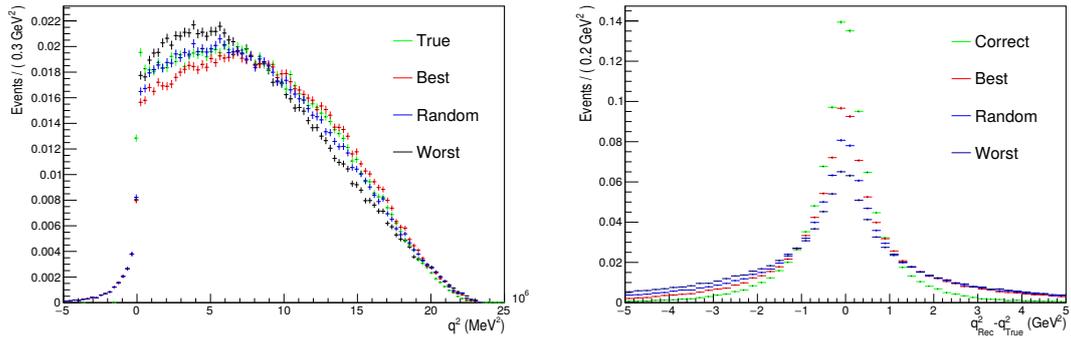


Figure 26: True q^2 compared to different reconstructed q^2 values (left) and q^2 resolution for different methods selecting q^2 (right).

Solution	RMS
Correct	1.55 GeV ²
Best regression	2.43 GeV ²
Random	2.99 GeV ²
Worst regression	2.99 GeV ²

Table 6: RMS of q^2 resolution with different methods selecting the reconstructed q^2 .

5.2.1 Linear regression and q^2

In general the linear regression analysis is a statistical method to predict the value of a response variable based on its relationship with a set of independent variables. More specific for this analysis, Reference [107] infers the B_s^0 momenta from B_s^0 flight information variables such as the flight length $|\vec{F}|$ and the polar angle θ_{flight} as defined in Figure 21. The B_s^0 momentum can be inferred from these quantities as $p = M|\vec{F}|/t$ or $p = \bar{p}_T/\sin\theta_{\text{flight}}$, where t is the decay time and \bar{p}_T is the average of the transverse momentum. Using a least squares linear regression algorithm [108] the regression gives a B_s^0 momentum estimate with a resolution of 60% which leads to the correct solution for the quadratic equation in around 70% of the cases and is thus better than random picking a solution with a success rate of 50% [107]. The predicted B_s^0 momentum is then compared to the two solutions derived from Equation 61 and the solution closest to the regression value, q^2_{Best} , is chosen. Figure 26 (left) compares the different reconstructed q^2 values to the true q^2 value which is derived from exploiting MC truth information, where q^2_{Worst} refers to the solution opposite to the regression value. Using the regression method to select a solution of the reconstructed q^2 improves its resolution, as shown in right of Figure 26. Quantitatively this can also be seen from the RMS of the q^2 distribution given in Table 6. Using a linear regression to chose the reconstructed q^2 solution improves the resolution by 23% compared to the random picking.

5.3 *sPlot* technique

The so-called *sPlot* technique [109] is a statistical method used to unfold the contributions of different sources in a data sample in a specific variable. In this variable the background candidates can be statistically subtracted in order to obtain the distribution of a pure signal sample. This technique is used in this analysis to obtain pure $B^+ \rightarrow J/\psi K^+$ and $D_s^- \rightarrow K^+ K^- \pi^-$ data samples as control channels to be able to compare them with MC distributions and obtain corrections for the simulation, as further explained in Section 7.

The *sPlot* technique will be further explained. Consider a data set with entries N consisting of signal and background samples, having each values for two uncorrelated variables x and y . Here the distributions are known in the control variable y and the aim is to get an estimate of the signal distribution in x without knowing the underlying distribution in this variable. This allows to statistically subtract the background sample from the x -distribution, which can either be done through a sideband subtraction or the *sPlot* technique.

The latter assigns each event of the data set a weight w_i in such a way that the weighted distribution re-samples the background-subtracted distribution. If N_k is the number of events of the signal sample, $f_k(y)$ is the known distribution of the signal in the control variable y and V_{nj} is the corresponding covariance matrix, the weights to obtain the background subtracted sample are given by:

$$w_i^n(y) = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(y)}{\sum_{k=1}^{N_s} N_k f_k(y)}, \quad (63)$$

where the yield N_k and the covariance matrix can be obtained from an extended maximum likelihood fit to the control variable y .

In this analysis the *sPlot* technique is used to distinguish signal and background-like events based on the invariant mass variable. The sidebands of the mass peak in the invariant mass distribution contains pure background, whereas the signal region is a mix of signal and background. A likelihood fit is performed on the data sample to determine the yields of both sources. From the invariant mass signal weights are calculated which subtract the background as a sidebands subtraction and returns the unfolded signal distribution. As an illustration the maximum likelihood fit to the $D_s^- \rightarrow K^+ K^- \pi^-$ invariant mass distribution in data is shown in Figure 27 together with the calculated sweights. This fit is used to obtain a pure D_s sample for the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$.

5.4 Boosted Decision Tree

Different techniques exist to classify observations into classes, such as to separate data into categories like signal or background. Multivariate analysis methods are especially effective for separation since they allow to combine several discriminating variables into one final discriminator. This analysis uses a so-called *Boosted Decision Tree* (BDT) classifiers to distinguish signal from background in the data sample collected by the LHCb experiment. A schematic view of a decision tree is shown in Figure 28, where a consecutive set of questions c_N are asked on the

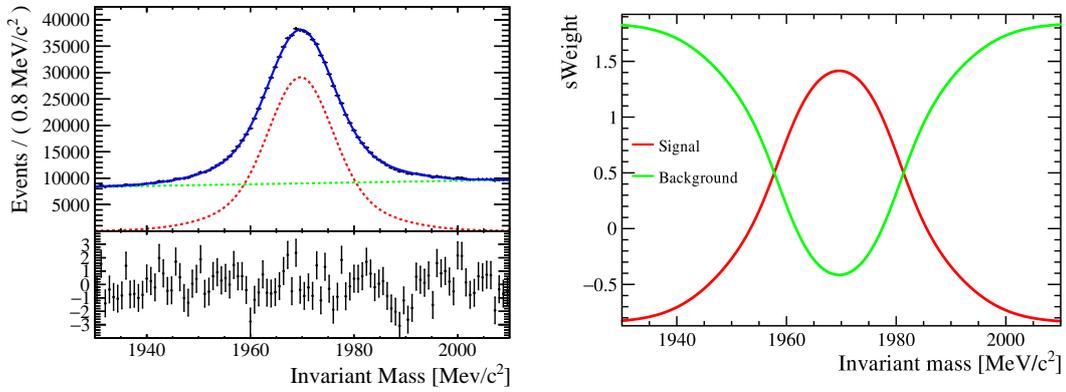


Figure 27: A maximum likelihood Fit to the $K^- K^+ \pi^-$ invariant mass (left) and the *sPlot* weights calculated from the fit results. Figure taken from [9].

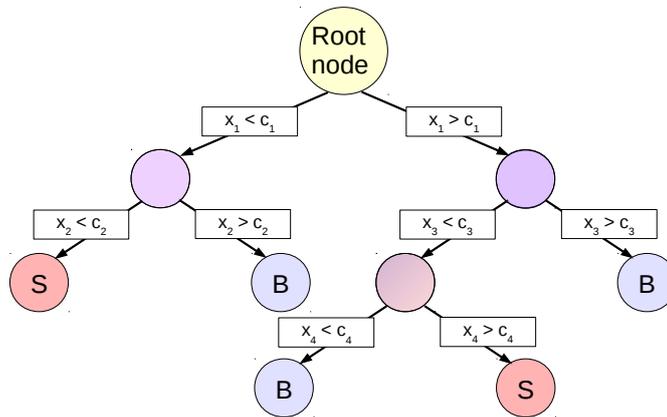


Figure 28: Schematic view of a decision tree. Starting from a root node applying a sequential set of selection cuts, maximising the separation between signal and background at each node. The final nodes are identified as signal and background.

chosen variables x_N representing the data with a binary outcome at each node. At each stage selection cuts are chosen to give the greatest separation between signal and background, a common metric for separation gain is the so-called Gini-Index defined as $p(1 - p)$ where p is the signal purity. The separation gain between the nodes is recursively maximised to optimise the node structure of the classifier. The final response is then classified as a signal or background leaf.

The decision tree has to be trained on a data set which already provides the outcome, such as pure signal and background samples from MC simulation. Even though decision trees are easy to interpret and fast to train, a single tree does not have a strong separation power. Therefore a BDT uses an ensemble of different trees, where the outcome of all trees are combined for its final decision. This is also known as *random forest* with the idea that the sum of many weak learners results in a stronger classifier. A typical tree contains several hundred weak decision trees with a maximum depth depending on the critical number of events per node.

There are different methods to train a random forest, for this analysis a *boosting* technique is used which gives mis-classified events a higher weight after each iteration. This method shifts the focus on the mis-classified samples over the correctly assigned events for the training of the following tree. For this analysis the *AdaBoost* method [110] (adaptive boosting) is used which assigns a boost weight derived from the misclassification rate, err , of the previous tree

$$w = (1 - err)/err. \quad (64)$$

The weights of the whole event sample are renormalised so that the sum of weights remains constant. If $h_i(\vec{x})$ is the result of an individual classifier with \vec{x} the input variables and $h_i(\vec{x}) = +1$ and -1 for signal and background, the final BDT output is then:

$$h_{BDT}(\vec{x}) = \frac{1}{N_{tree}} \sum_i^{N_{tree}} \ln(w_i) h_i(\vec{x}) \quad (65)$$

where N_{tree} is the number of trained decision trees.

BDT classifiers can suffer from over-training, where the BDT makes decisions due to statistical fluctuations rather than real differences in data. This would lead to a bigger separating power of the classifier on the training sample with respect to what is truly achieved. Therefore the data sample is split into training and testing samples and the performance of the training samples should not be better than the test sample. In addition one can approach this issue by using k -fold cross validations [111], where the training data is divided into k subsamples and the classifier is trained k times using $k - 1$ subsamples. Then the classifier is tested on the subsamples which are independent of the training samples.

Differences between the data used in training and classification can lead to a biased BDT outcome. This is the case if there are fundamental differences between MC and data which can be picked up by the algorithm and then wrongly classified. Therefore possible differences need to be corrected for, which is explained in more detail in Section 7.

The performance of a classifier can be illustrated by the *Receiver Operating Characteristic* (ROC) curve. This curve shows the background rejection with respect

to the signal efficiency of the remaining sample. The area under curve (AUC) is largest for the best classifier. BDTs are frequently used in this analysis to separate signal from background events as much as possible as explained further in Section 6.

6 Analysis selection

In the proton-proton collisions many different particles are produced at the interaction point, where most of them are pions or other light particles. Only a small fraction of these events contains the B_s^0 meson required for this analysis. Especially the decay of interest as a $b \rightarrow u$ transition is very suppressed compared to large background contributions from $|V_{cb}|$ -decays. In order to find the decay we are interested in, it has to pass several stages of selection criteria which are outlined in this chapter. First the data passes the general event reconstruction of the online and offline processing steps, as explained in Section 3. The specific trigger and striping line used in this analysis are introduced in Section 6.1 and 6.2 respectively. In general the selections are applied to maximise the signal efficiency and reject background as much as possible by exploiting differences in the topology of signal and background events. Signal events consist of long lived particles originating from a secondary vertex with a significant distance from the primary vertex with a large transverse momentum. The analysis specific offline selection such as additional selection cuts applied for the signal and normalisation channel are further described in Section 6.3. Those consist of applying vetoes to suppress specific backgrounds, tighter particle identification requirements and in addition charged track isolation variables. Two multivariate classifiers are trained and applied to further suppress different remaining background components as explained in Section 6.4.

6.1 Prerequisites

Several selection variables must be defined in order to understand the selection requirements:

- DOCA Distance of closest approach between two particle tracks.
- IP Impact parameter defined as the minimal distance between a track with respect to the primary vertex (PV).
- IP χ^2 Impact parameter χ^2 defined as the difference between the χ^2 obtained when fitting the PV with and without the considered particle. It is a measure of the probability that a particle is coming from the PV.
- FD The flight distance (FD) is the length of displacement between the primary and secondary vertex.
- FD χ^2 The flight distance χ^2 is the difference between the χ^2 obtained when fitting the secondary vertex (SV) with and without the constrain of zero flight distance.
- DIRA Cosine of the direction angle. The direction angle is the angle between the momentum vector of the reconstructed particle with respect to the flight distance vector.

6.2 Trigger selection

Events are first processed online by the L0 hardware trigger. For this analysis candidates need to pass the L0MuonTos trigger line at hardware level, which selects events containing at least one muon with a transverse momentum of more than 1.76 GeV/c for 2012 and a SPD hit multiplicity of below 600 [86].

In the first stage of the software-based trigger, HLT1, signal events have to pass the Hlt1SingleMuonHighPT which requires minimum transverse momenta of 4.8 and a momenta of larger than 8 GeV/c [87].

As a second part of the software trigger, events must pass either the Hlt2SingleMuonDecisionTOS or the Hlt2TopoMu2BodyDecisionTOS trigger line. The first selects candidates with a transverse momenta larger than 1.3 GeV/c, a good quality track ($\chi^2/\text{n.d.f.} < 2$) which is well separated from the PV via the impact parameter ($\text{IP} > 0.5\text{mm}$, $\text{IP}\chi^2 > 200$). This trigger line is pre-scaled with a factor of 0.5 in data, which has to be applied also to simulated events exclusively triggered by this line to match the data. The TopoMu2BodyBBDT is an inclusive trigger designed to select partially reconstructed decays of B hadrons containing a high likelihood muon and one additional good quality track [87]. Two-body objects are created from these tracks if their distance of closest approach (DOCA) is less than 0.2 mm, and if they form a displaced secondary vertex. The following kinematic variables are used to efficiently select signal candidates: $\sum |p_T|$, p_T^{\min} , invariant mass, corrected mass, DOCA, IP significance $\text{IP}\chi^2$ and flight distance significance ($\text{FD}\chi^2$). Here the $\sum |p_T|$ variable sums up the transverse momenta of 2 candidate tracks and p_T^{\min} is the minimum transverse momentum of one of these tracks. Those variables are combined into a multivariate selection rather than a simple cut-based trigger, since higher rejection power of background for the same signal efficiency is achieved [86]. A boosted decision tree (BDT) is used with discretised input variables given in Table 7, also called a Bonsai BDT (BBDT). The BDT can only apply selections at the specific intervals listed in the last column of Table 7, the chosen cut values are given in the second column.

Variable	Cuts	BBDT Intervals
$\sum p_T [\text{GeV}/c]$	> 3	3.5, 4, 4.5, 5, 6, 7, 8, 9, 10, 15, 20
$p_T^{\min} [\text{GeV}/c]$	> 0.5	0.6, 0.7, 0.8, 0.9, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 5, 10
$m [\text{GeV}/c^2]$	< 7	2.5, 4.75
$m_{\text{corr}} [\text{GeV}/c^2]$		2, 3, 4, 5, 6, 7, 8, 9, 10, 15
DOCA [mm]	< 0.2	0.05, 0.1, 0.15
$\text{IP}\chi^2$		20
$\text{FD}\chi^2/100$	> 1	2, 3, 4, 5, 6, 7, 8, 9, 10, 25, 50, 100

Table 7: The variables, the selection cuts applied and the intervals used in the BBDT for the TopoMu2BodyBBDT trigger selecting 2 body decays. Table slightly modified from [86].

Variables	Stripping cuts
Event	long track multiplicity < 250
Kaon	
Track p_T	> 0.5 GeV/ c
Track p	> 10 GeV/ c
Track $IP\chi^2$	> 16
Track $\chi^2/n.d.f.$	< 6
Track GhostProb	< 0.5
$\Delta\log\mathcal{L}_{K\pi}$	> 5
$\Delta\log\mathcal{L}_{Kp}$	> 5
$\Delta\log\mathcal{L}_{K\mu}$	> 5
muon	
Track p_T	> 1.5 GeV/ c
Track p	> 6 GeV/ c
Track $IP\chi^2$	> 12
Track $\chi^2/n.d.f.$	< 4
Track GhostProb	< 0.35
$\Delta\log\mathcal{L}_{\mu\pi}$	> 3
$\Delta\log\mathcal{L}_{\mu p}$	> 0
$\Delta\log\mathcal{L}_{\mu K}$	> 0
$K\mu (B_s^0)$	
m_{corr}	[2.5, 7] GeV/ c^2
SV $\chi^2/n.d.f.$	< 4
DIRA	> 0.994
FD χ^2	> 120
HLT2	Hlt2SingleMuonDecisionTOS or Hlt2TopoMu2BodyDecisionTOS

Table 8: Stripping selection criteria for $B_s^0 \rightarrow K^+\mu^-\nu$ signal candidates.

6.3 Preselection

The first offline selection stage to process the data are the so-called stripping selections. For this analysis dedicated stripping lines were written, one for the signal (**StrippingB2XuMuNuBs2KLine**) and several others for background studies. The **StrippingB2XuMuNuBs2KSSLine** is used to select same sign kaon and muons ($K^+\mu^+$) as well as stripping lines with looser particle identification (PID) selection cuts for the muon, kaon and both (**StrippingB2XuMuNuBs2K_FakeKLine**, **StrippingB2XuMuNuBs2K_FakeMuLine**, **B2XuMuNuBs2K_FakeKMuLine**) which are used for particle misidentification studies.

The selection cuts for the signal are further specified in Table 8. The stripping line uses as input the trigger selection specified in the previous section. Events with at least one reconstructed PV are selected and events with more than 250 long tracks as defined in Section 3.2.1 are rejected. First requirements on the final state particles are made and good quality tracks $\chi^2/n.d.f.$ are selected. To reduce the amount of artificial ghost tracks created by the reconstruction software,

tracks need to have a small ghost track probability, as introduced in the detector section. Kinematic cuts for the muon and kaon to select high momentum and high transversal momentum tracks are applied as well as requirements based on the particle ID, such that they have a high likelihood to be identified as a kaon or a muon as discussed in Section 3.2.2. Based on the topology of the decay the final state tracks have to be well separated from the primary vertex by requiring a large impact parameter χ^2 ($\text{IP}\chi^2$).

In the next step the previously selected final state particles are combined to build the partially reconstructed B_s^0 candidate. The reconstructed secondary vertex from this combination has to be of good fit quality (Vertex $\chi^2 < 4$) and needs to be well separated from the primary vertex (flight distance $\chi^2 > 120$). The combined B_s^0 candidate has to fulfil the requirement $\text{DIRA} > 0.994$. For signal decays this quantity should be close to one. In addition the corrected mass of the B_s^0 candidate has to be within the range of $[2.5, 7]$ GeV/c^2 .

For the normalisation channel a different Stripping line `B2DMuNuX_Ds` is used to efficiently select $B_s^0 \rightarrow D_s^+ \mu^- \nu$ decays. All selection cuts are summarised in Table 10. For the muon similar cuts as for the signal are used to reduce systematic uncertainties, whereas for the kaon less tight kinematic cuts are used due to the additional charged tracks coming from the D_s^+ decay, such as the pion. The latter inherits the same cuts as the kaon apart from the PID cut. The selected kaon and pion particles are then combined to build the D_s^+ candidate using a window around the nominal D_s^+ mass together with topological requirements such as the direction angle (DIRA) and a well separated, good quality decay vertex. From this the B_s^0 candidate is built together with the muon track accepting candidates in a large mass window around the B_s^0 mass. Also the B_s^0 vector is required to be in flight direction before the D_s^+ vector.

In addition a dedicated line has been written for background studies by relaxing the muon PID requirements `B2DMuNuX_Ds_FakeMu`.

6.4 Offline selection

In order to reduce the background contamination as much as possible, further selection cuts have to be placed on the data. As explained earlier one of the dominant backgrounds for this analysis are $|V_{cb}|$ decays such as $B^+ \rightarrow D^0 \mu \nu$, which are due to the CKM-matrix structure and the higher fragmentation fraction about 40 times larger than the signal decay we are interested in. Also background from $b \rightarrow c \bar{c} s$ decays such as $B^+ \rightarrow J/\psi K^+$ needs to be reduced as much as possible. They can be wrongly identified as signal decays if the positively charged muon is not reconstructed. Also higher excited K^* resonances can be misidentified as signal, when they further decay into a kaon and neutral pion. To suppress these kind of backgrounds vetoes are applied, but due to the small reconstruction efficiency of the neutral pion those are not very effective.

Variables	Stripping cuts
Event	long track multiplicity < 250
Muon	
Track p_T	> 1 GeV/c
Track p	> 6 GeV/c
Track GhostProb	< 0.35
Track $\chi^2/\text{n.d.f.}$	< 3.0
Track $\text{IP}\chi^2$	> 12
$\Delta\log\mathcal{L}_{\mu\pi}$	> 0
Kaon	
Track p_T	> 250 MeV/c
Track p	> 2000 MeV/c
Track GhostProb	< 0.35
Track $\chi^2/\text{n.d.f.}$	< 3.0
Track $\text{IP}\chi^2$	> 4
$\Delta\log\mathcal{L}_{K\pi}$	> -2.0
Pion	
Track p_T	> 250 MeV/c
Track p	> 2000 MeV/c
Track GhostProb	< 0.35
Track $\chi^2/\text{n.d.f.}$	< 3.0
Track $\text{IP}\chi^2$	> 4
$\Delta\log\mathcal{L}_{K\pi}$	< 20.0
D_s^+ candidate	
$ m_{\text{Cand.}} - m_{D_s^+} $	< 80 MeV/c ²
DOCA χ^2	< 20
Vertex $\chi^2/\text{n.d.f.}$	< 6.0
FD χ^2	> 25
DIRA	> 0.99
$D_s^+ \mu^-$ (partial B_s^0)	
m_{Cand}	[2.2, 8] GeV/c ²
Vertex $\chi^2/\text{n.d.f.}$	< 9.0
DIRA	> 0.999
Vertex(D_s^+) _Z - Vertex(B_s^0) _Z	> -0.05

Table 9: Stripping selections applied to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ candidates using the B2DMuNuX_Ds line.

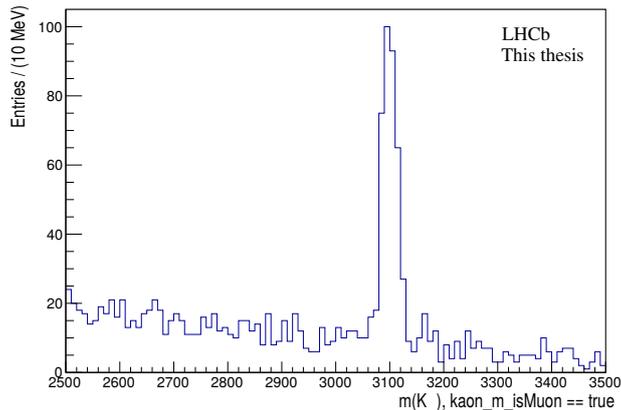


Figure 29: Invariant $K^+\mu^-$ mass for kaons identified as muons in data. Events consistent with the J/ψ mass are rejected.

6.4.1 Background vetoes for $B_s^0 \rightarrow K^-\mu^+\nu_\mu$

Several vetoes are applied to reduce different background components for the signal channel, they are summarized at the end of the section in Table 11.

The opposite sign kaon and muon pair used to reconstruct the signal can originate from semileptonic charm meson decays. Those backgrounds are reduced by requiring the invariant mass of the $K^+\mu^-$ pair to be greater than the mass of the D meson (>1.9 GeV/ c).

At the stripping level PID requirements are applied to select a $K^+\mu^-$ pair, but residual muons can remain which are misidentified as kaons. Therefore $J/\psi \rightarrow \mu^+\mu^-$ decays can be wrongly identified as $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ signal decays. These candidates are rejected by requiring the $K^+\mu^-$ invariant mass to be outside the nominal J/ψ mass, if the kaon is reconstructed under the muon mass hypothesis and is identified as a muon. Figure 29 shows the invariant $K^+\mu^-$ mass of these events in data, they are vetoed if they fall into the J/ψ mass window of $3072 < m(\mu K \rightarrow \mu) < 3130$ MeV/ c^2 .

Background from higher excited kaon decays ($K^{*+} \rightarrow K^+\pi^0$) can be rejected by searching for neutral pions in a cone around the reconstructed kaon track. The π^0 candidates are reconstructed using two photons with an invariant mass around the nominal π^0 mass. The reconstructed π^0 mass is shown in the left of Figure 30 for different simulated K^* resonances. Higher excited kaon candidates are rejected if a neutral pion is found and the invariant mass of the $K^+\pi^0$ pair is consistent with coming from either the $K^{*+}(892)$ or the $K_{0,2}^{*+}(1430)$ with the mass ranges ($|m(K^+\pi^0) - m(K^*(892))| < 65$) and ($|m(K^+\pi^0) - m(K^*(1430))| < 90$) respectively. The right of Figure 30 shows the invariant $K^+\pi^0$ mass for the simulated K^* resonances, those are reconstructed as the signal through the $K^+\mu^-$ pair where a neutral pion is added with the requirements given above. This veto only rejects about 20% of the background from higher excited kaon resonances due to the low reconstruction efficiency of neutral pions [112]. Its remainder has

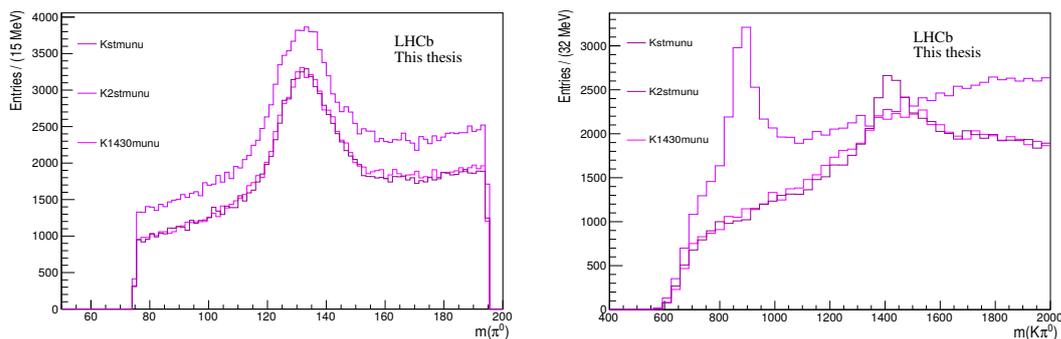


Figure 30: (Left) The invariant mass of the neutral pion reconstructed from two photons $m(\pi^0)$ on the left and the reconstructed invariant mass of a kaon together with a neutral pion $m(K^-\pi^0)$ for different simulated K^* resonances: $B_s^0 \rightarrow K^{*-}(892)(\rightarrow K^-\pi^0)\mu^+\nu_\mu$, $B_s^0 \rightarrow K_0^{*-}(1430)(\rightarrow K^-\pi^0)\mu^+\nu_\mu$ and $B_s^0 \rightarrow K_2^{*-}(1430)(\rightarrow K^-\pi^0)\mu^+\nu_\mu$.

to be taken into account in the signal fit to the corrected mass. The veto rejects around 4% of signal events.

Combinatorial background comes from randomly combined K^+ and μ^- tracks. It can arise from $b\bar{b}$ production where the kaon and muon originate from the decay of different quarks which hadronise into different B mesons. This background has to be suppressed further. This can be done by exploiting the different topology of $b\bar{b}$ production with respect to the decay of interest as displayed schematically in Figure 31. The two b -quarks fragment into B -mesons which are back-to-back in the rest frame. For a boosted $b\bar{b}$ pair in the longitudinal direction, as produced at the LHC, this results in B mesons having opposite momenta in the transverse plane. Therefore their corresponding decay products, the $K^+\mu^-$, also have opposite transverse momenta and they end up in opposite quadrants in the xy plane. This results in a low relative transverse momenta of the reconstructed fake B_s^0 meson candidate as shown in Figure 32 for data together with true B_s^0 candidates. These events are rejected by requiring $p_x(K^+) \times p_x(\mu^-) < 0$ AND $p_y(K^+) \times p_y(\mu^-) < 0$.

6.4.2 Additional selection cuts for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$

In addition to the selection cuts applied at the stripping level, tighter PID cuts are applied to the kaon and muon of the normalisation channel. Also the kaons need to have a momenta larger than 10 GeV/c and the reconstructed D_s^+ candidate should be within the nominal D_s^+ mass. All additional selection cuts are summarized in Table 10.

For the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ with $D_s^+ \rightarrow K^+ K^- \pi^+$ the background mainly originates from semileptonic B_s^0 decays to higher excited D_s^+ resonances, such as $B_s^0 \rightarrow D_s^{*-} \mu^+ \nu_\mu$, $B_s^0 \rightarrow D_{s0}^{*-} \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_{s1}^{*-} \mu^+ \nu_\mu$. Other background consists of partially reconstructed decays such as $B_q^0 \rightarrow D_s^- D_q^{(*)} X$ where the D_q further decays semileptonically.

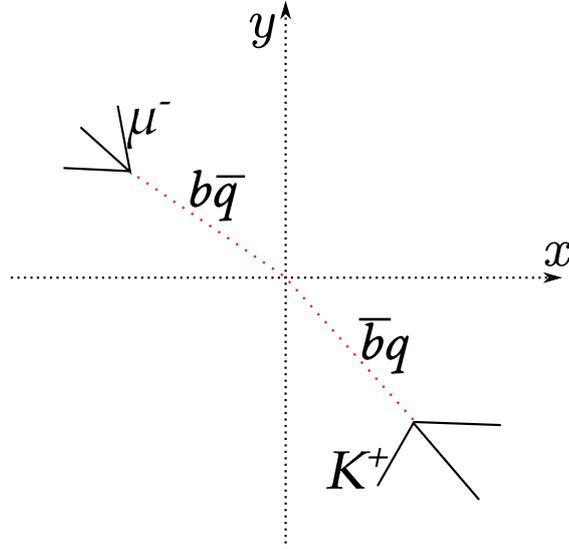


Figure 31: Topology of the combinatoric background with kaon and muon originating from the decay of different B mesons looking down the beam line. Figure taken from [9].

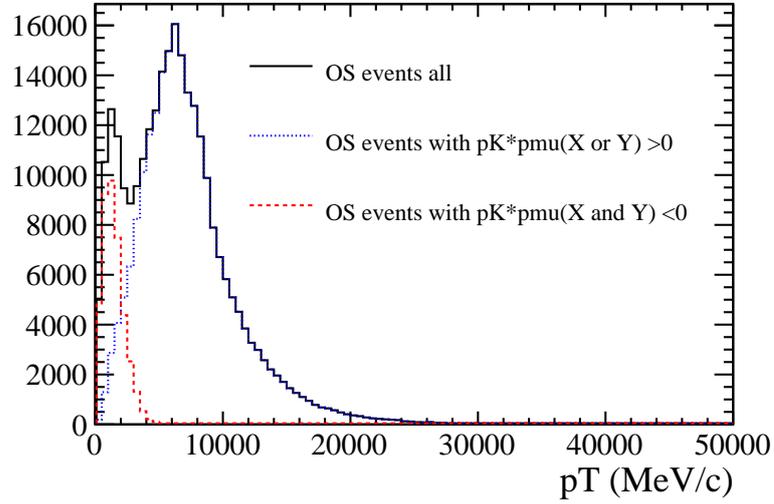


Figure 32: p_T distribution of reconstructed B_s^0 candidates in data (black solid), split up into combinatorial background candidates with $p_x(K^+) \times p_x(\mu^-) < 0$ AND $p_y(K^+) \times p_y(\mu^-) < 0$ (dashed red) and potential signal candidates with $p_x(K^+) \times p_x(\mu^-) < 0$ OR $p_y(K^+) \times p_y(\mu^-) < 0$ (blue). Figure taken from [9].

Additional Selections	
K^-	$\Delta \log \mathcal{L}_{K\pi} > 5.0$
K^-	$\Delta \log \mathcal{L}_{Kp} > 5.0$
K^-	$\Delta \log \mathcal{L}_{K\mu} > 5.0$
K^-	$p > 10000 \text{ MeV}$
μ^+	$\Delta \log \mathcal{L}_{\mu\pi} > 3.0$
μ^+	$\Delta \log \mathcal{L}_{\mu p} > 0$
μ^+	$\Delta \log \mathcal{L}_{\mu K} > 0$
D_s^+	$ m_{\text{Cand.}} - m_{D_s^+} < 40 \text{ MeV}/c^2$
Veto	
$D^{*-} \rightarrow D^0 \pi^+$ Veto	$ m_{KK\pi} - m_{KK} > 148 \text{ MeV}$
$B_s^0 \rightarrow D_s^- \pi^+$ Veto	$ m_{Ds(\mu \rightarrow \pi)} - m_{B_s^0} > 70 \text{ MeV}$
Isolation	$\min(\text{IsoMinBDT}_K, \text{IsoMinBDT}_\mu) > -0.8$

Table 10: Additional selections and vetoes applied to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ candidates.

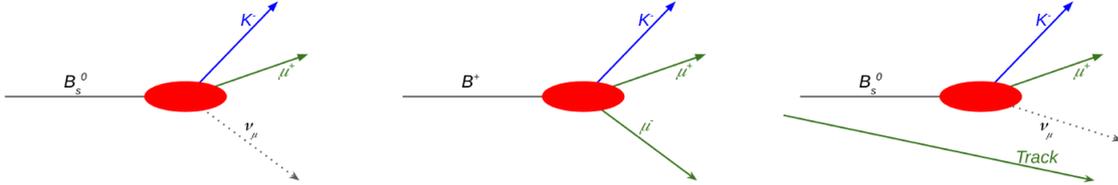


Figure 33: The topology of Signal decays (left) compared to $B^+ \rightarrow J/\psi K^+$ background (middle) where the second muon is a non-isolated track and isolated tracks coming from a different vertex (right). Figure taken from [9].

Background from $B \rightarrow (D^* \rightarrow (D^0 \rightarrow K^+ K^-) \pi) \mu \nu X$ is rejected using the mass difference between the D^* and D^0 which is only slightly higher than the pion mass ($|m(KK\pi) - m(KK)| > 148 \text{ MeV}$). This selection efficiently rejects all $B \rightarrow D^* \mu \nu$ decays.

Possible background from $B_s^0 \rightarrow D_s^- \pi^+$ can be reduced by reconstructing the muon under the mass hypothesis of a pion and rejecting $D_s^- \pi^+$ candidates with an invariant mass around the nominal B_s^0 mass ($|m(Ds(\mu \rightarrow \pi)) - m(B_s^0)| < 70 \text{ MeV}$). The applied isolation cut mentioned in Table 10 is explained in the following section.

6.4.3 Charged track isolation

Charged track isolation variables are used to further discriminate signal from background with additional charged tracks. They are used by many LHCb analysis [113–115]. For true B_s^0 signal decays no further charged tracks are expected nearby the kaon and muon tracks. Instead for partially reconstructed background such as $B^+ \rightarrow J/\psi K^+$ there are additional charged tracks originating from the same decay vertex, as displayed in Figure 33.

The algorithm searches through all underlying tracks of the event for additional charged tracks (other than the signal tracks) which build a good vertex with any

signal track. In addition it uses cone isolation variables by drawing a cone around the candidate tracks in $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, where ϕ is the azimuthal angle and η the pseudorapidity, and checking for tracks and energy deposits in the calorimeter within this cone. The algorithm returns variables determining the activity around the track either from neutral or charged particles. True $B_s^0 \rightarrow K^+\mu^-\nu$ candidates are well isolated and have very little detector activity within the cone. Several variables are built based on cone isolation of the candidate track:

$p_T(\text{cone}) = \sum_i \vec{p}_T$	Transverse momenta of a cone defined as the vector sum of p_T of all tracks within the cone.
$A_{p_T} = \frac{ \vec{p}_T(\text{track}) - \vec{p}_T(\text{cone}) }{ \vec{p}_T(\text{track}) + \vec{p}_T(\text{cone}) }$	Momentum asymmetry between the cone and the candidate track.
$TI = \frac{p_T(\text{track})}{p_T(\text{track} + \text{Cone})}$	Transverse isolation of the cone with respect to the candidate track.

The isolation algorithm used in this analysis is a BDT adapted from a previous analysis [116]. It uses as input kinematic and topological information of the B candidate and the reconstructed tracks and returns for each track an output variable specifying the likelihood to be consistent with originating from the same vertex. The BDT is trained with simulated samples of $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ as signal and $B^+ \rightarrow D^{*-}\pi^+\mu^+\nu_\mu$ as background. The following variables are used in the training to separate signal and background, here 'test track' refers to all reconstructed tracks of the event other than the signal tracks which are tested :

track minIPchi2	Minimum of impact parameter χ^2 of the test track with respect to any PV.
track η, ϕ, p_T	pseudorapidity, azimuthal angle and transverse momentum of the test track.
track_MatchChi2	Track quality determined from the track fit when downstream and upstream track segments are matched.
track_pvdis_mu	Distance between (signal track, test track) vertex with respect to the PV.
tracksvdis_mu	Distance between (signal track, test track) vertex of with respect to the SV.
track_DOCA_mu	Distance of closest approach between the signal track and the test track.
track_angle_mu	Angle between signal track and the test track.
track_fc_mu	$FC = \frac{(p_{sig} + p_{track}) \times \alpha^{sig+track, PV}}{(p_{sig} + p_{track}) \times \alpha^{sig+track, PV} + p_{t_{sig}} + p_{t_{track}}}$

where $\alpha^{sig+track, PV}$ is the angle between the sum of momenta ($p_{sig} + p_{track}$) and the straight line from the primary vertex to the vertex reconstructed using the signal track and the test track.

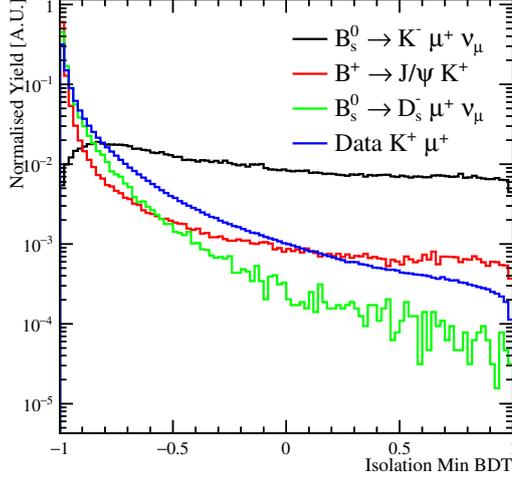


Figure 34: The per-event distribution of the minimum output of the isolation BDT for signal (black) compared to background MC from $B^+ \rightarrow J/\psi K^+$ (red), $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ (green) and SS data (blue). Figure taken from [9].

Distributions of these discriminating variables are shown in Figure 35 for signal and background. The output of the Gradient Boosted Decision Tree (BDTG) classifier is shown in Figure 36 along with its ROC curve. In the next step these track-based quantities have to be converted into candidate based quantities. When processing an event each reconstructed track is assigned a weight based on the BDT output to be either a signal or an underlying track. Different combinations of the BDT values are possible, for this analysis two combined variables are created for each candidate track: `IsoMinBDT` and `IsoSumBDT`. The first one is the minimum BDT output value of the test track which relates to a high probability to originate from the same vertex as the candidate and is assigned to both the kaon and muon. The minimum BDT output for both kaon and muon is shown in Figure 34 for simulated signal and background events as well as for data. There is a clear separation visible with respect to the background events which peak at low values, therefore a loose selection cut is placed on this variable of $\min(\text{kaon_IsoMinBDT}, \text{muon_IsoMinBDT}) > -0.9$ for $B_s^0 \rightarrow K^+ \mu^- \nu$ and $\min(\text{kaon_IsoMinBDT}, \text{muon_IsoMinBDT}) > -0.8$ for the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. The `IsoSumBDT` variable is the sum of the BDT output over all underlying tracks when compared to the candidate track. The isolation algorithm also saves the kinematics of the least isolated track ($\text{track}_{\text{least iso}}$) with respect to both candidate tracks. These charged track isolation variables are used as input for two additional BDTs trained to discriminate signal and background even further and are explained in more detail in Section 6.5.

6.4.4 Corrected mass uncertainty

As mentioned already in Section 5.1 rejection of events with a large uncertainty on the corrected mass improves the separation between signal and background. In this analysis candidates with a corrected mass uncertainty larger than $100 \text{ MeV}/c^2$ are rejected which results in an efficiency of $\approx 30\%$ for both signal and partially

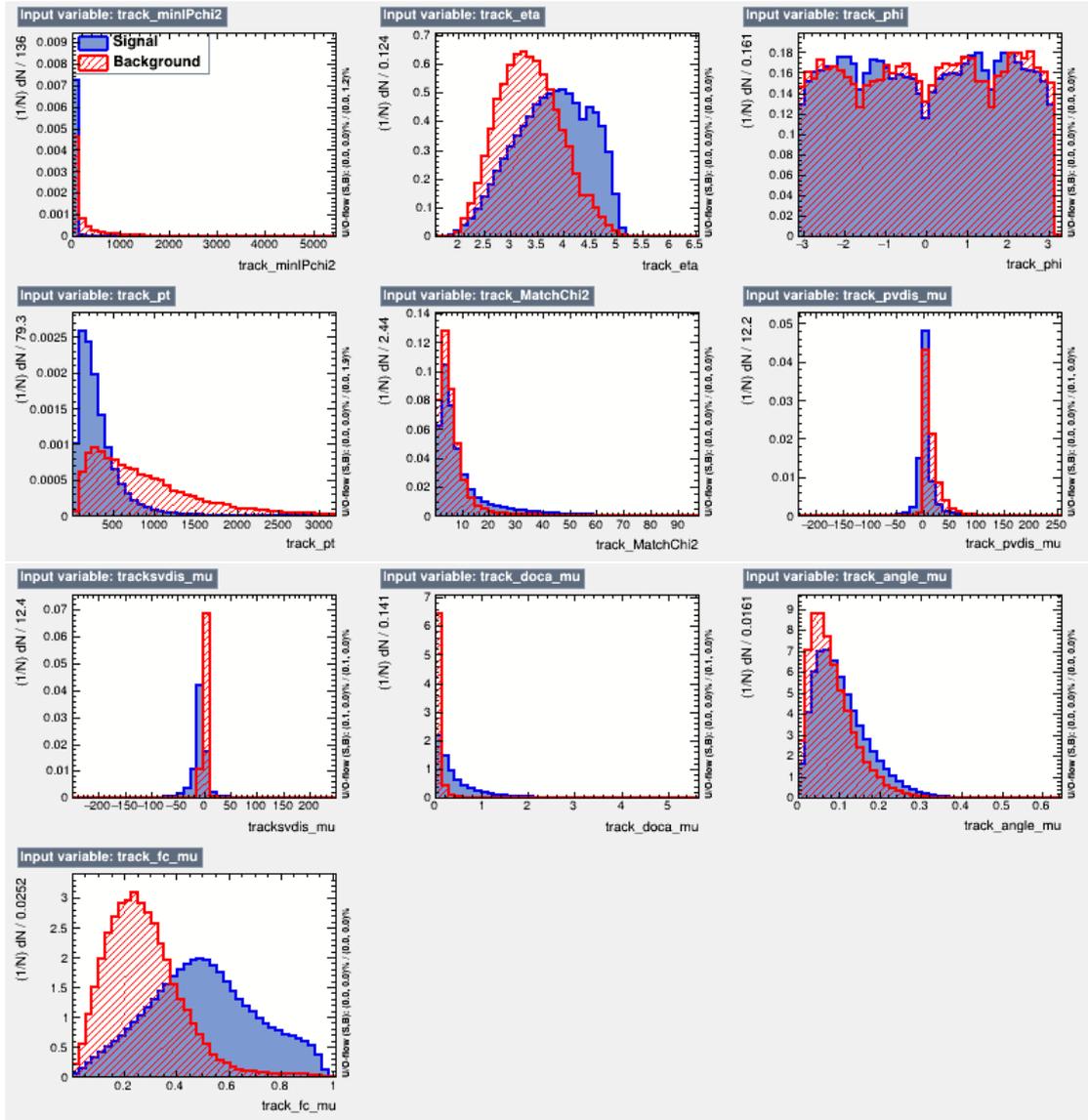


Figure 35: Input variables for the BDT training of the charged track isolation. Figure taken from [9].

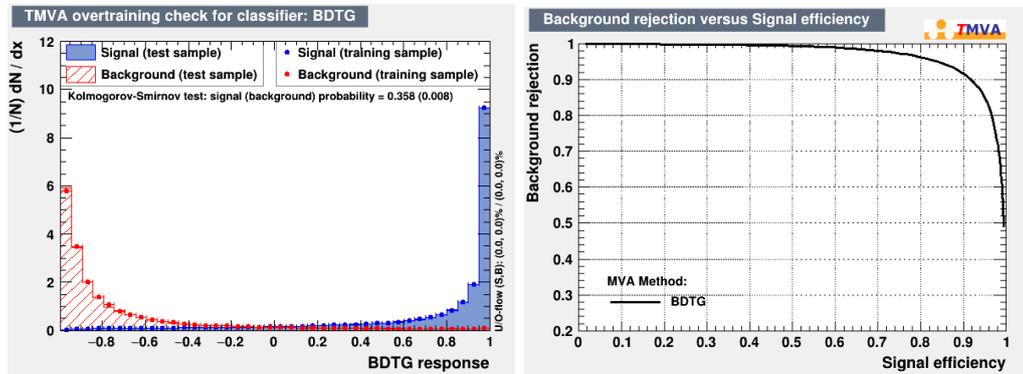


Figure 36: (Left) BDTG output distribution for signal (blue) and background (red). (Right) ROC curve corresponding to the BDTG output. Figure taken from [9].

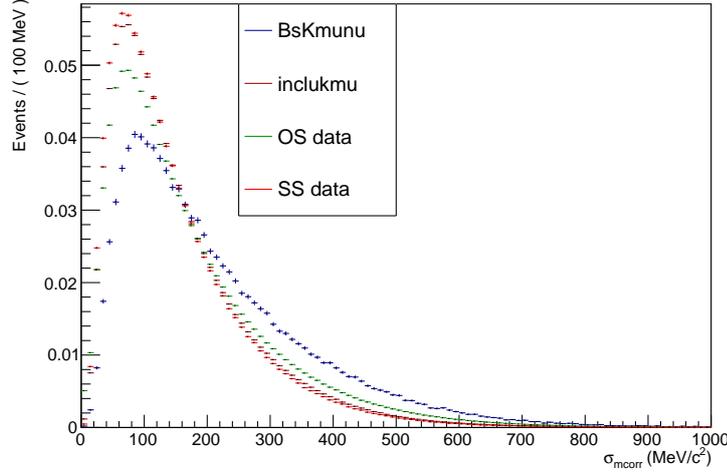


Figure 37: The corrected mass uncertainty for signal decays compared to background decays.

reconstructed background decays while backgrounds from random combinations are significantly reduced. The cut value of $100 \text{ MeV}/c^2$ corresponds to the peak of the signal distribution of the corrected mass uncertainty, as shown in Figure 37. Although this selection does not increase signal purity the separation between signal and background decays is significantly improved in the corrected mass, this results in a reduced systematic uncertainty when performing a fit to the corrected mass.

The variable is verified using a control channel as explained further in Section 7. The distributions of the corrected mass uncertainty for $K^-\mu^+$ combination from the decay $B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+$ are plotted in Figure 38 for Monte Carlo samples and data, together with simulated signal events. For the normalisation mode $B_s^0 \rightarrow D_s^-\mu^+\nu_\mu$ no selection on the corrected mass uncertainty of the $K^-\mu^+$ pair or the $D_s^-\mu^+$ pair is made.

6.4.5 Summary of the offline selection

The offline requirements applied to $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ signal decays which were described in the previous subsections are summarized in Table 11.

6.5 Selection BDT

For this analysis two different BDTs are trained to separate signal from background, both of them use $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ signal MC samples for training. The first BDT is used to discriminate signal from partially reconstructed background with additional charged tracks, the so-called *charged BDT*. The second BDT is trained to remove background found in SS data, such as combinatorial and other feed-down decays of higher excited particles, it is referred to as *SS BDT*. All previously defined selection cuts, including the loose cut on the output of the isolation BDT, are applied to

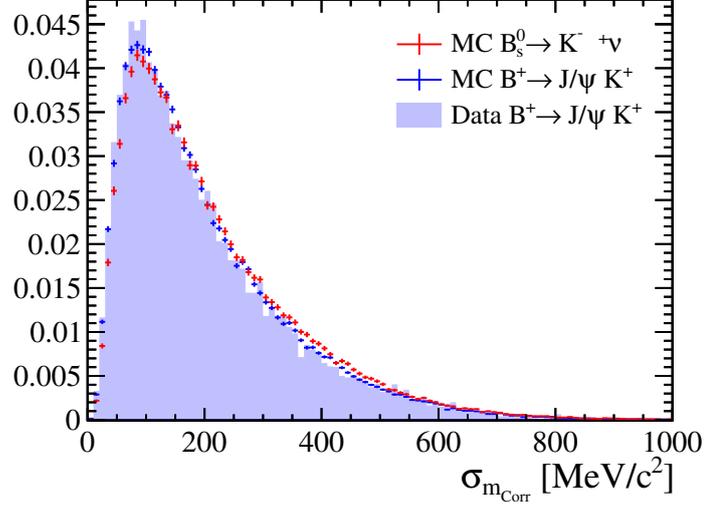


Figure 38: The corrected mass uncertainty for signal decays and $B^+ \rightarrow J/\psi K^+$ decays reconstructed as $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. Figure taken from [9].

Selection	Variables
Candidates per event	= 1
D meson rejection	$m_{K\mu} > 1900 \text{ MeV}/c$
π^0 and K^* veto	$ m(\pi^0) - 135 < 30 \text{ MeV}$ & \mathcal{E} $(m_{K^+\pi^0} - m_{K^*(892)} > 65 \text{ MeV}$ OR $ m_{K^+\pi^0} - m_{K^*(1430)} > 90 \text{ MeV})$
J/ψ misID veto	K^+ IsMuon & $ m_{(K \rightarrow \mu)\mu} - m_{J/\psi} > 30 \text{ MeV}$
combinatorial cut	$p_x(K^+) \times p_x(\mu^-) > 0$ OR $p_y(K^+) \times p_y(\mu^-) > 0$
track isolation	$\min(\text{kaon_IsoMinBDT}, \text{muon_IsoMinBDT}) > -0.9$
corrected mass uncert.	$\sigma_{m_{corr}} < 100 \text{ MeV}/c^2$

Table 11: Offline selection criteria for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$

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

Table 12: Simulated background samples used for training of the charged BDT.

the MC and data samples. Then the charged BDT is trained on MC samples from partially reconstructed backgrounds, a selection based on its output is chosen and afterwards the SS BDT is trained to provide additional discriminating power on other backgrounds.

The charged BDT is trained on a mix of background MC samples, reconstructed as $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$, their contribution is given in Table 12. All of these samples used for training, are kinematically corrected using the reweighting procedure described in Section 7. It uses as input the variables displayed in Table 13 to separate signal from background, the separation power of the different variables with respect to the background is also given as computed by the BDT. It is zero for identical distributions and one if distributions do not overlap. For the training 850 trees with a maximum depth of 3 and a minimum node size of 2.5% of all events are used with the AdaBoost method [110]. To remove possible overtraining the data is divided by magnet polarity for training and testing samples.

The same-sign (SS) BDT is trained with $K^- \mu^-$ candidates in data as the background sample. The input variables used in the training are shown in Table 14 together with their separation power. To minimise correlations in the training between the two BDTs a selection is placed on the output of the charged BDT before the training of the SS BDT. Its training follows the same procedure as for the charged BDT described above.

The BDT response for both BDTs are shown in Figures 39. In the left plot the charge BDT response for simulated signal events is compared to background from simulated $B_s^0 \rightarrow K^{*-} \mu^+ \nu_\mu$, $B^+ \rightarrow J/\psi K^+$, $|V_{cb}|$ decays, combinatorial background as well as SS data. In the right plot the SS BDT response of the signal is compared to simulated $|V_{cb}|$ decays as well as SS data. The corresponding ROC curves to illustrate the performance of the classifiers by plotting the signal efficiency against

Variable	Separation
$\min(\text{kaon_IsoMinBDT}, \text{muon_IsoMinBDT})$	1.90×10^{-1}
$m(K^- + \text{track}_{\text{least iso}})$	1.00×10^{-1}
$\max(\text{kaon_IsoSumBDT}, \text{muon_IsoSumBDT})$	9.39×10^{-2}
Minimum of kaon & muon cone Isolation	3.54×10^{-2}
Transverse isolation (TI) between K^- and cone	3.00×10^{-2}
Kaon p_T	2.84×10^{-2}
Transverse isolation (TI) between K^- and charged cone	2.58×10^{-2}
$p_T(B_s^0) - 1.5 \times p_T(\mu^+)$	2.17×10^{-2}
$B_s^0 p_T$	1.94×10^{-2}
$\Delta\eta$ between K^- and charged cone	1.89×10^{-2}
Momentum asymmetry (A_{p_T}) between μ^+ and charged cone	1.85×10^{-2}
$\text{kaon_IsoMinBDT} - \text{muon_IsoMinBDT}$	1.63×10^{-2}
B_s^0 Decay vertex fit χ^2	1.11×10^{-2}
$m(\mu^+ + \text{track}_{\text{least iso}})$	7.67×10^{-3}
B_s^0 helicity angle	2.03×10^{-3}

Table 13: The input variables used for the charged BDT are listed with their separating power. Several variables use information obtained from a cone drawn around candidate tracks with $\Delta R = 0.5$ and were defined in Section 6.4.3.

Variable	Separation
Kaon p_T	2.76×10^{-02}
$\text{DIRA}_{B_s^0}$	1.98×10^{-02}
Momentum asymmetry (A_{p_T}) between K^- and neutral cone	1.36×10^{-02}
Transverse isolation (TI) between K^- and neutral cone	1.30×10^{-02}
$m(K^- + \pi^0)$	1.02×10^{-02}
$p_T(B_s^0) - 1.5 \times p_T(\mu^+)$	9.31×10^{-03}
B_s^0 Flight distance significance	7.56×10^{-03}
$B_s^0 p_T$	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 14: The input variables for the SS BDT training. Several variables use information obtained from a cone draw around candidate tracks with $\Delta R = 0.5$.

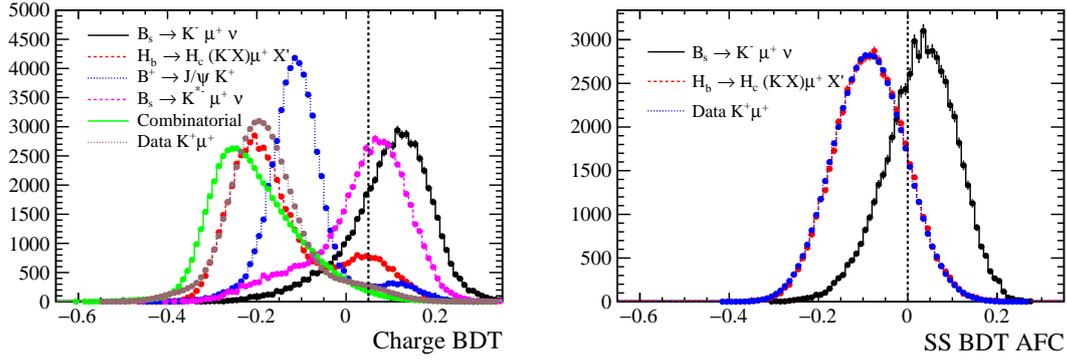


Figure 39: Response for the charged BDT (left) and SS BDT (right). In the left the signal response in black is compared to background from simulated $|V_{cb}|$ decays (indicated by $H_b \rightarrow H_c (K X) \mu^+ X'$ in red), $B_s^0 \rightarrow K^{*-} \mu^+ \nu_\mu$ in purple, $B^+ \rightarrow J/\psi K^+$ in blue, combinatorial background in green as well as SS data in grey. In the right the SS BDT response for signal in black is compared to simulated $|V_{cb}|$ decays in red as well as SS data in blue. Figure taken from [9].

the background rejection are given in Figure 40. A selection of `Charged_BDT` $>$ 0.05 is placed on the charged BDT and `SS_BDT` $>$ 0 on the same sign BDT.

The BDTs are validated by comparing the BDT response in *sPlot* unfolded data with simulated $B^+ \rightarrow J/\psi K^+$ decays using a fully reconstructed $K^+ \mu^- \mu^+$ decay or only the $K^+ \mu^-$ pair. The validation is shown in Figure 41, where a slight discrepancy is visible in the BDT response between simulation and data. Therefore a correction factor and a systematic uncertainty needs to be applied to the calculated BDT efficiency using the $B^+ \rightarrow J/\psi K^+$ decay. The details are given in Section 10. More validation plots of the BDT input variables are given in Appendix B.

6.6 Selection on data

Figure 42 shows the B_s^0 corrected mass distribution for opposite-sign $K^+ \mu^-$ and same-sign $K^+ \mu^+$ candidates in data after applying all consecutive selection steps. The first plot shows the data where only the trigger and stripping selection as detailed in Sections 6.1 and 6.2. Then the different vetoes are applied as well as the combinatorial background cut of the offline selection detailed in Section 6.4.2. The charge isolation requirement introduced in Section 6.4.3 reduces the data sets even more and by applying the two selection BDTs defined in Section 6.5 a structure appears in the $K^+ \mu^-$ data set corresponding to the B_s^0 mass. When applying the corrected mass error cut in the final plot the signal decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ is visible by eye as a shoulder at the B_s^0 mass. Figure 42 compares the opposite-sign $K^+ \mu^-$ data candidates to the simulated $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decay after the same consecutive selections are applied. Here the top four plots are drawn normalized otherwise the signal distribution would not be visible, in the bottom two the distributions are not normalized such that the evolving signal structure in opposite-sign data is overlaid with the shape from the signal simulation.

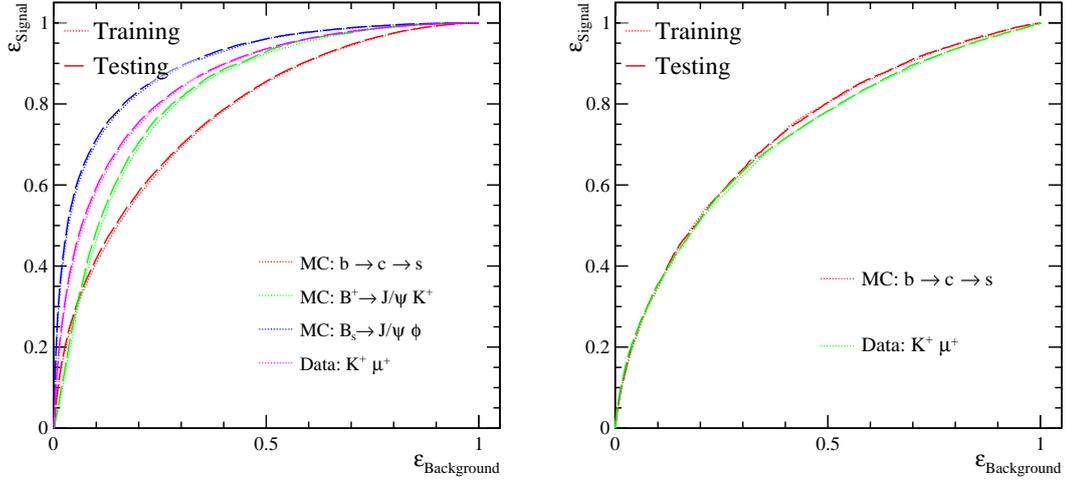


Figure 40: ROC curves for the charged BDT (left) and SS BDT (right). The curves for training and testing are both shown. The SS BDT is applied after the charge BDT. Figure taken from [9].

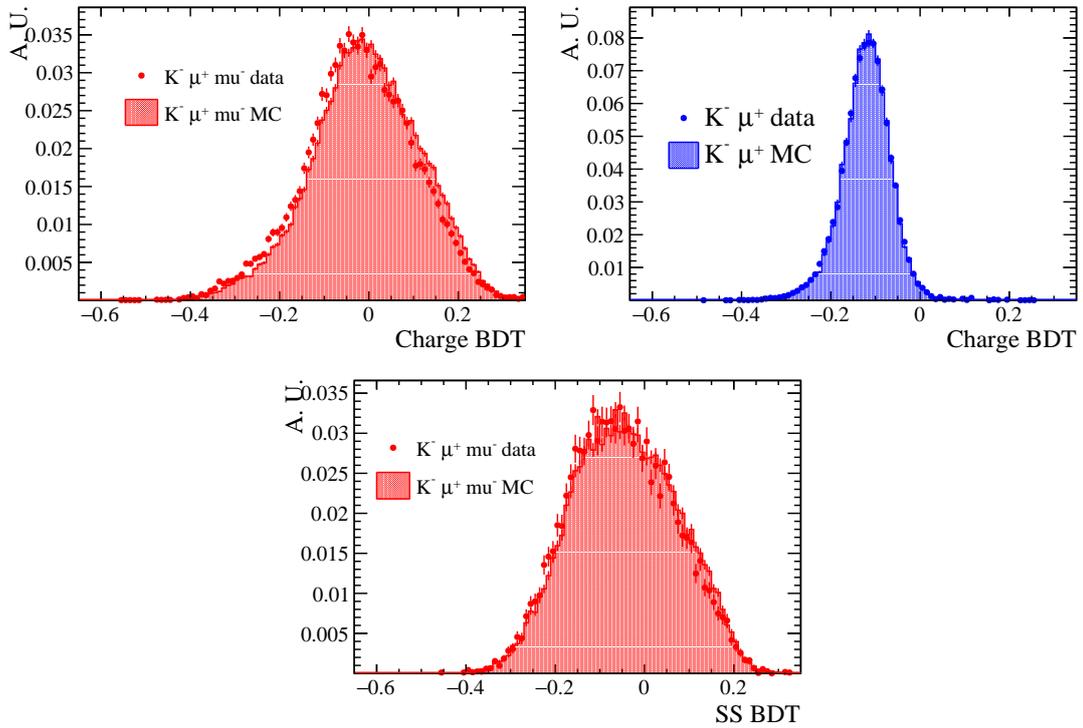


Figure 41: The charged BDT response on top and SS BDT on the bottom plotted for $B^+ \rightarrow J/\psi K^+$ decays reconstructed as $B^+ \rightarrow J/\psi K^+$ (red) and the $K^+ \mu^-$ pair in (blue) for Monte Carlo (line) and background subtracted data (points). Figure taken from [9].

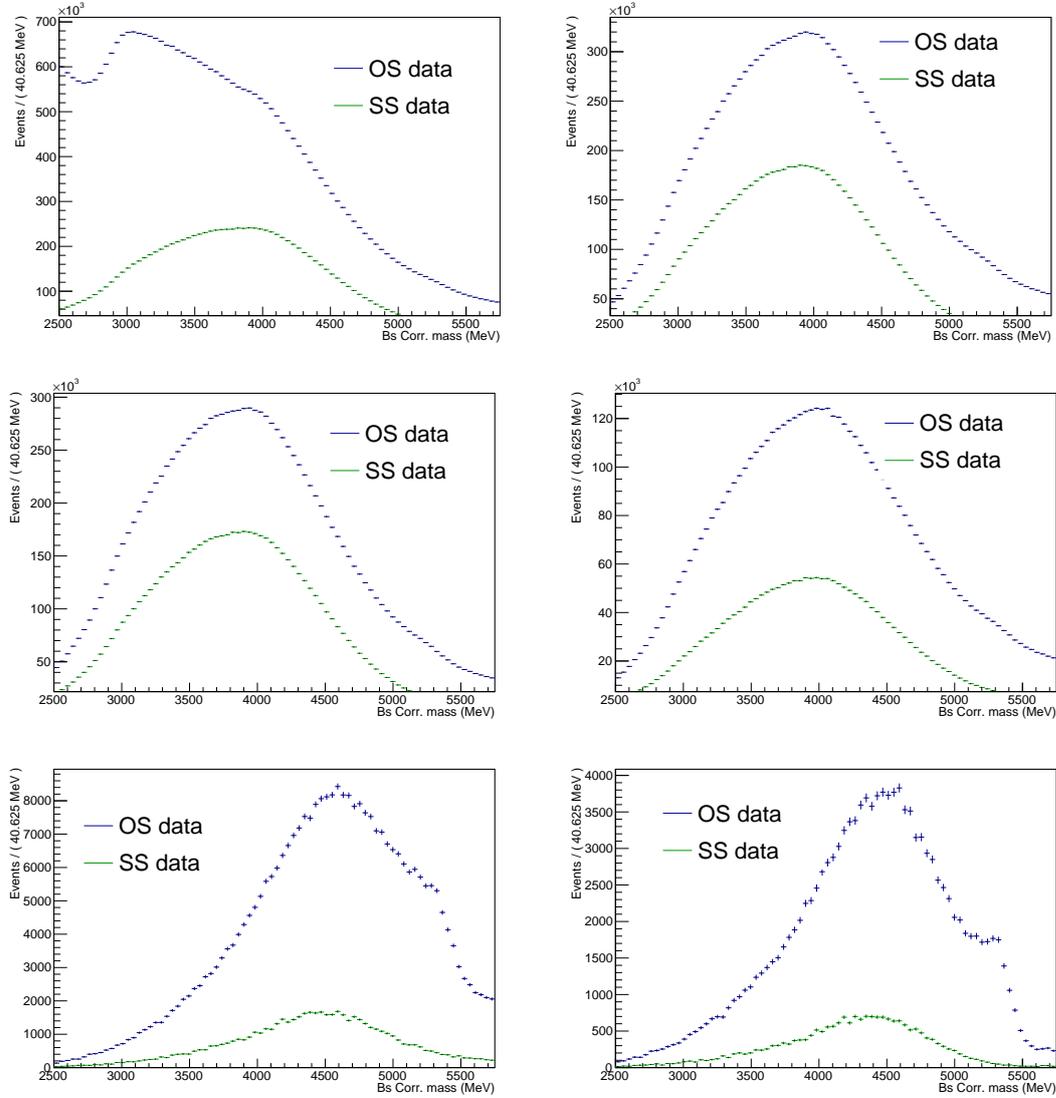


Figure 42: Corrected mass distribution for kaon and muon candidates from same-sign (SS) and opposite-sign (OS) data passing consecutive selection cuts. The top left plot shows the distributions after trigger and stripping cuts are applied, top right for adding the vetoes and middle left plot with the combinatorial cut applied. The middle right plot has the isolation cut and the bottom left the selection BDTs applied. For the bottom right plot the corrected mass error cut is added on top as the full selection, here a clear peaking structure at the B_s^0 mass can be seen.

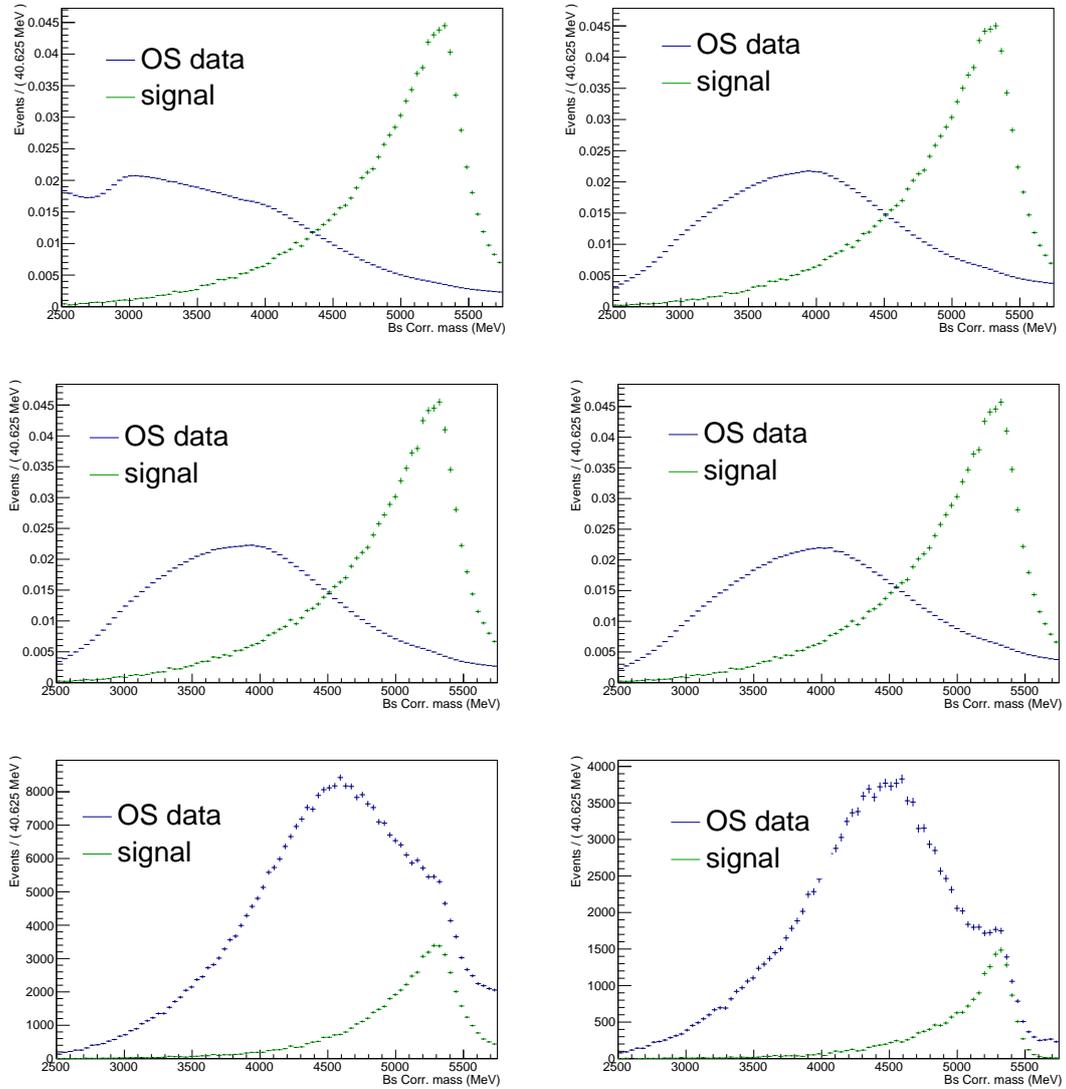


Figure 43: Corrected mass distribution for $K^+\mu^-$ candidates from opposite-sign (OS) data compared to simulated signal events passing consecutive selection cuts. The same sequence of cuts are applied as in the previous Figure from top left to bottom right: trigger and stripping, adding vetoes, removing combinatorial background, isolation requirements, applying BDT cuts and full selection cuts including the low corrected mass uncertainty selection.

7 Control channel

In general the agreement between simulated samples and data is not perfect and several variables show disagreement. As many Monte Carlo (MC) samples are used in this analysis either to evaluate efficiencies or as fitting variables it is important that these differences between data and MC are corrected. Therefore the accuracy with which the MC samples describe the data has to be verified in control channels.

A simple solution for the MC correction would be to reweight the MC sample to agree with data in the variable of interest showing discrepancies. This approach however fails when more than one variable needs to be reweighted. Low statistic bins and the unknown optimal binning choice to resolve structures within a bin often cause problems for multidimensional reweighting. Therefore for this analysis a BDT is trained to separate signal candidates in data and MC. If the simulation perfectly models the data, the BDT returns an output variable with no separation power. Overtraining and statistical fluctuations would return a false separation power. For disagreement between data and MC samples the BDT returns an output variable with separation power. This approach assumes that this output variable combines all differences between simulation and data in different variables into a single discriminating variable. A reweighting is performed to match the BDT output variable in simulation to data. This one dimensional correction on the BDT output is able to correct all variables used in the training simultaneously [117].

The control channel used for this correction should have similar topology and kinematics as the signal channel $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. Here the decay $B^+ \rightarrow J/\psi K^+$ is selected as the control channel due to its high yield. Furthermore it is a high purity data sample with only minor background contributing. Depending on the study under investigation, $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ is either fully reconstructed from $\mu^- \mu^+ K^+$ or only through the $\mu^- K^+$ pair similar to the signal. In order to validate the kinematic variables of the signal decay the $B^+ \rightarrow J/\psi K^+$ decay is reconstructed as $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ due to the similarity to the signal with a non reconstructed neutrino with respect to the non reconstructed second muon in $B^+ \rightarrow J/\psi K^+$. The B^+ candidate can be fully reconstructed together with the additional muon found by the charge isolation algorithm as the least isolated track as defined in Section 6.4.3. The reconstructed B^+ mass peak is then used to extract a pure data sample, subtracting background by exploiting the *sPlot* method as explained earlier in Section 5. The mass distribution is shown in left of Figure 44.

For the correction of the simulated normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ a simulated cocktail sample $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$ is used to compare to a well reconstructed D_s^+ peak in data together with a muon. Possible backgrounds are reduced by selecting regions of the isolation BDT output (`muon_p_IsoMinBDT` > -0.8 AND `kaon_m_IsoMinBDT` > -0.8) as well as tight PID cuts for the kaon and muon. The *sPlot* technique is used on the D_s^+ peak to get a clean D_s^+ sample as shown in the right of Figure 44.

The BDT training for both samples uses 200 trees with a maximum depth of 3

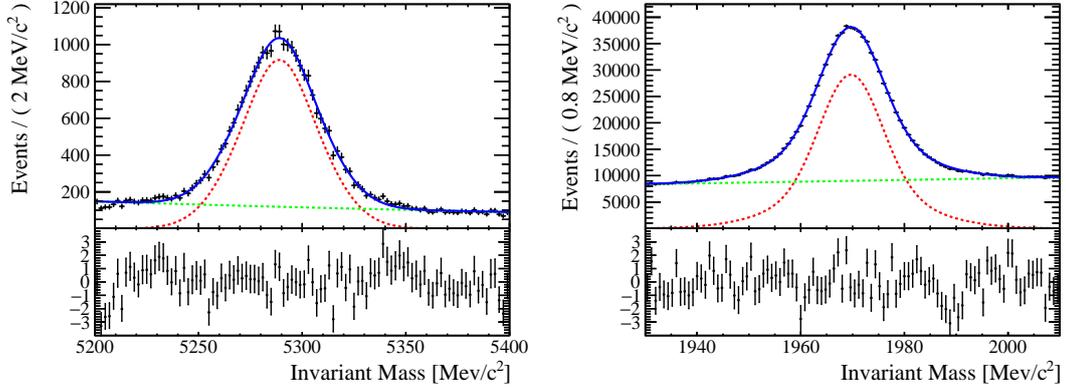


Figure 44: Invariant Mass of B^+ reconstructed from $K\mu\mu$ (left) and D_s^+ reconstructed from $K^+K^-\pi$ (right) used to generate the sWeights for the signal and normalisation channel. Figure taken from [9].

$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$
Track multiplicity	Track multiplicity
$B_s^0 \eta$	$B_s^0 \eta$
$B_s^0 p_T$	$B_s^0 p_T$
$K^- p_T$	$D_s^- p_T$
$\mu^+ p_T$	$\mu^+ p_T$

Table 15: Input variables used to train the BDT to correct for MC and data differences.

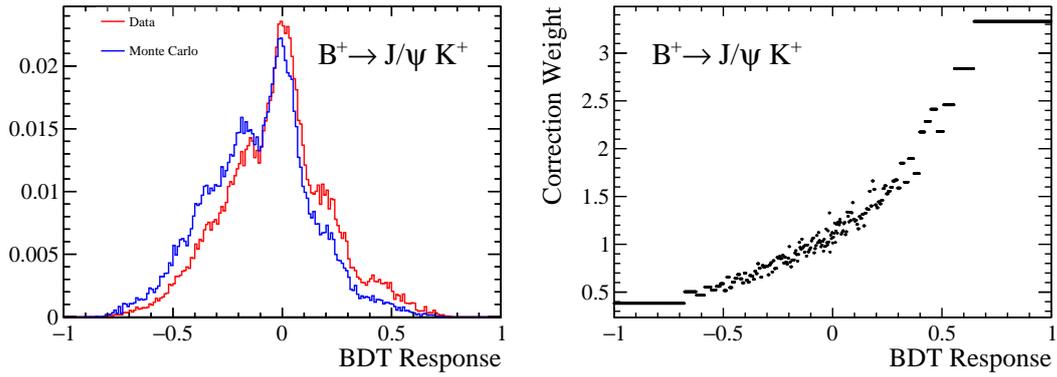


Figure 45: The BDT response for $B^+ \rightarrow J/\psi K^+$ used to separate Monte Carlo and data (left) and the weights used to correct the simulation (right). The decay is reconstructed through only the $\mu^- K^+$ pair. Figure taken from [9].

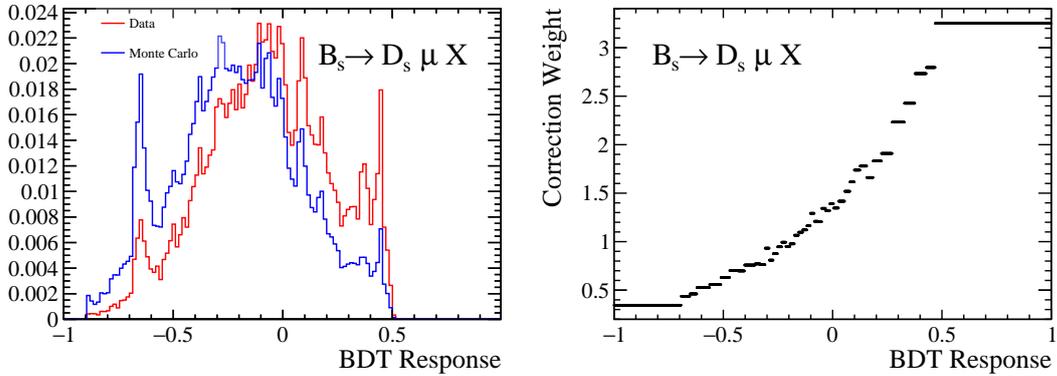


Figure 46: The BDT response and the weights for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$ used to correct the simulation. Figure taken from [9].

and a minimum leaf size of 6% for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ and 4% for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ to separate simulation and data with the AdaBoost method. Overtraining is removed by using again the 2 k -folding method for training and testing as explained in Section 5.4. The BDT is trained with the input variables given in Table 15 which are the ones showing differences between simulation and data and should be corrected in simulation by the BDT.

The left of Figure 45 shows the BDT response for $B^+ \rightarrow J/\psi K^+$ candidates in data and MC samples. In the right of Figure 45 the corresponding correction weights are shown which are applied to the simulated sample. Figure 46 shows the BDT response in MC and data as well as the corresponding correction weights for the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$. Figure 47 shows the effect of the BDT reweighting for some variables of $B^+ \rightarrow J/\psi K^+$ decays in simulation together with background subtracted data. Figure 48 shows kinematic variables for the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$ before and after the BDT reweighting corrections are applied. After the corrections the MC and data samples agree very

well. Additional validation plots can be found in Appendix C. The control channel for the signal and normalisation channel are also used to evaluate systematic uncertainties due to the mis-modelling of selection cuts between MC and data as further explained in Section 10.2.

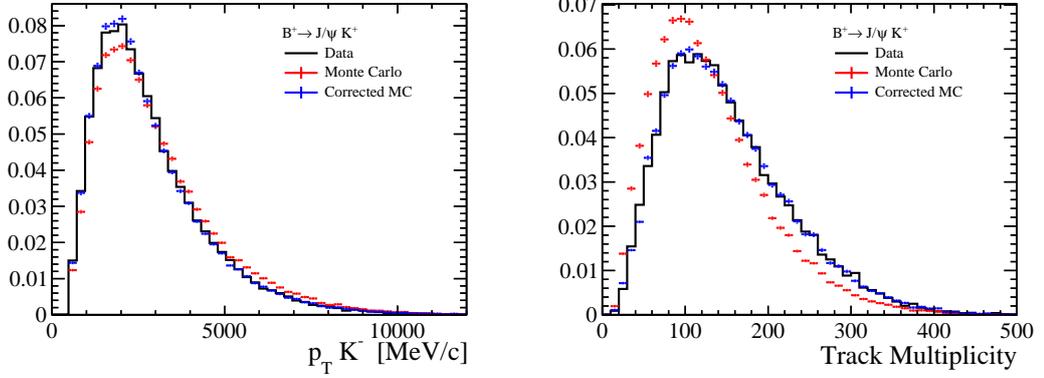


Figure 47: Kinematic distributions for $B^+ \rightarrow J/\psi K^+$ reconstructed using the $K^+ \mu^-$ pair for simulation before and after correction, plotted against background subtracted data. Figure taken from [9].

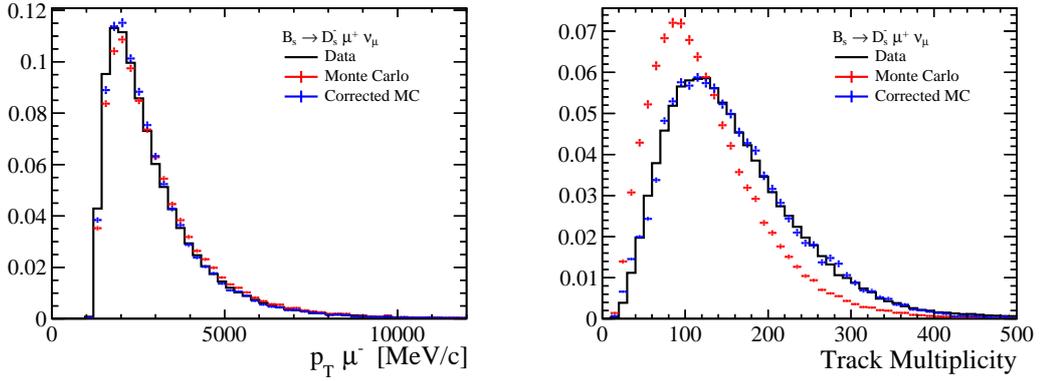


Figure 48: Kinematic distributions for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$ before and after correction using the BDT reweighting compared to data. Figure taken from [9].

8 Fit to the normalisation channel

After the signal and normalisation channel are selection and the kinematic corrections are applied using the control channel, as described in the previous section, the corresponding yields of the signal and normalisation channel are needed to measure their relative branching fraction. This chapter presents the maximum likelihood fit performed to extract the yield of the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ normalisation channel. A binned template fit is performed on the corrected mass of the $D_s^- \mu^+$ candidates in data, passing all selection cuts outlined before (Section 6). The Beeston-Barlow method is used for fitting templates taking into account the finite Monte Carlo statistics as discussed in Section 9.1. The specific implementation for this analysis the so-called `HistFactory` tool is outlined in Section 9.2. The MC templates used in the fit to the normalisation mode to model the corrected mass are explained in Section 9.3. A background subtraction is done to remove non- D_s^+ events in data as shown in Section 9.4. The final results of the fit to the reconstructed $D_s^- \mu^+$ corrected mass distribution are given in Section 9.5 together with the discussion of a possible fit bias.

8.1 Beeston-Barlow method

The composition of the data sample is extracted by fitting the corrected mass. However the individual distributions cannot be described analytically. Therefore one has to rely on simulations to provide templates for the shape of the signal and background. A binned template fit is performed to extract the signal yield using histograms made of finite MC samples. Exploiting the so-called the Beesten-Barlow method [118] uncertainties due to the finite statistics of the MC samples used to obtain the templates is taken into account. This method is briefly explained here. Dividing up the data into n bins gives a set of n numbers d_i with $i \in [1, n]$, where d_i is the number of events in data that fall into bin i . If j is a component contained in the data (a fit component), P_j its strength and a_{ji} the number of simulated events from component j in bin i , then the predicted number of events per bin i is given by:

$$f_i = N_D \sum_{j=1}^m \frac{P_j a_{ji}}{N_j}, \quad (66)$$

where N_D is the total number of events in the data sample, N_j the number of simulated events of component j and m the total number of components (MC templates) contained in the data

$$N_D = \sum_{i=1}^n d_i, \quad N_j = \sum_{i=1}^n a_{ji}. \quad (67)$$

The P_j are the actual proportions of the different components which should sum to unity. To estimate those one can perform a χ^2 minimisation using

$$\chi^2 = \sum_i \frac{(d_i - f_i)^2}{d_i}, \quad (68)$$

but this assumes that the d_i follow a Gaussian distribution, which is a good approximation for large numbers. Instead for binned fits there is only a small number of data points per bin, which makes this approximation inappropriate and leads to the use of a Poissonian distributed probability to observe d_i :

$$\mathcal{L} = \prod_{i=1}^n e^{-f_i} \frac{f_i^{d_i}}{d_i!}. \quad (69)$$

The proportions are then found by maximising the logarithm of the likelihood

$$\ln(\mathcal{L}) = \sum_{i=1}^n (d_i \ln(f_i) - f_i) \quad (70)$$

where constant factorials are dropped. This likelihood function is also known as the *binned likelihood*. It accounts correctly for small numbers of data events per bin but it does not include uncertainties due to the finite number of events in simulated samples.

Due to a large computation time simulated samples have often a small number of events which leads to non-negligible statistical fluctuations of a_{ji} . Those have to be taken into account which can be incorporated into the likelihood by modifying it as follows. For each source j per bin i there is an unknown expected number of events A_{ji} , the predicted number of data events per bin (Equation 66) is then modified to be

$$f_i = N_D \sum_{j=1}^m \frac{P_j A_{ji}}{N_j}. \quad (71)$$

The observed number of simulated events a_{ji} is actually generated from A_{ji} by a Poisson distribution. Therefore the probabilities of observing d_i and a_{ji} have to be combined and the total likelihood function to be maximised is

$$\ln(\mathcal{L}) = \sum_{i=1}^n d_i \ln(f_i) - f_i + \sum_{i=1}^n \sum_{j=1}^m a_{ji} \ln(A_{ji}) - A_{ji}. \quad (72)$$

The maximisation of this likelihood equation is referred to as the *Beeston-Barlow Method* throughout this thesis.

8.2 HistFactory and Beeston-Barlow light method

Introducing the nuisance parameters A_{ji} to include statistical uncertainties of the simulation into the template fit, increases the number of floating parameters and gives an additional nuisance parameter for each bin for each Monte Carlo template. The values $A_{ji}(p_j)$, where $p_j = N_D P_j / N_j$, can only be determined by numerically solving a large number of independent equations, one for each bin. This has to be done at each iteration step of the value p_j and leads to a large computational effort with a very slowly converging fit. Also this can lead to discontinuous jumps in the nuisance parameters [119], which is problematic for the numerical minimization routine `Minuit` [120].

The `HistFactory` tool [121] is used in this analysis to perform template fits to the corrected mass, which is also part of the `Roofit/RooStats` framework based on `ROOT` [122]. This framework is designed for binned fits using histogram templates and it builds a likelihood function which can be minimized with the usual combination of minimizers built into `Minuit` (`Migrad`, `Hesse`, `Minos`). Due to the computational complexity the `HistFactory` package uses a modified Beeston-Barlow method, the so-called *Beeston-Barlow light* technique. This is implemented in such a way that instead of a separate nuisance parameter A_{ji} for each bin for each Monte Carlo template, only a single nuisance parameter γ_i per bin is included which accounts for the total MC estimate and the total statistic uncertainty in that bin: $A_{ji} = \gamma_i a_{ji}$. Such that each bin of the combined model is assigned a statistical uncertainty according to the sum of the relative uncertainties from each template. Each bin of the combined model is allowed to fluctuate up or down by a fractional amount γ_i .

As this method does not keep track of separate fluctuations of the individual templates but only the summed template this light method may underestimate uncertainties from small statistic templates as it mainly accounts for uncertainties in the largest fit components. Thus the uncertainty from the Beeston-Barlow light method should be cross-checked using the bootstrap technique and creating bootstrapped fit templates where one samples from the original individual templates using random selection with replacement to create alternative templates.

8.3 Normalisation fit model

A maximum likelihood, binned template fit is performed on the corrected $D_s^- \mu^+$ mass distribution to extract the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ normalisation yield. This is done using the Beeston-Barlow light method implemented in the `HistFactory` tool as described in the previous chapter. The fit to the corrected $D_s^- \mu^+$ mass is able to separate the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ component from background contributions. Here the D_s^- is reconstructed through the $K^- K^+ \pi^-$ channel and in order to remove $K^- K^+ \pi^-$ combinatorial contributions not coming from the D_s^- decay, a background subtraction is performed using the $K^- K^+ \pi^-$ invariant mass as described in more detail in the next section.

Background to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ mainly comes from semileptonic B_s^0 decays containing higher excited D_s^- resonances. Also backgrounds from doubly charmed final states such as $B_q \rightarrow D_s^{(*)-} D_q^{(*)}$ are contributing, where the D_q is decaying semileptonically and the D_s^+ further decays into $K^- K^+ \pi^-$ and $q = (u, d, s)$. Also partially reconstructed tauonic decays like $B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$ are considered.

A potential background contamination from misidentified muons is studied using data passing the stripping selection `B2DMuNuX_Dp_FakeMuon`. The contribution is found to be of the order of three per mille of the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ yield, its yield is constrained in the fit. Misidentified D_s^- background does not contribute to the fit since it is subtracted by the fit to the invariant $K^- K^+ \pi^-$ mass as described in Section 8.4.

Random combinations of real muons with real D_s^+ mesons can also lead to a combinatorial background contribution. Those are investigated using the same

Component	Source
$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	Monte Carlo
$B_s^0 \rightarrow D_s^{*-} \mu^+ \nu_\mu$, with $D_s^{*-} \rightarrow D_s^- \gamma$	Monte Carlo
$B_s^0 \rightarrow D_{s0}^{*-} \mu^+ \nu_\mu$, $B_s^0 \rightarrow D_{s1}^{*-} \mu^+ \nu_\mu$	Monte Carlo
$B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$, with $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	Monte Carlo
$B_q \rightarrow D_s^{(*)-} D_q^{(*)}$, with $D_q \rightarrow \mu^+ \nu_\mu X$, $q = (u, d, s)$	Monte Carlo
Misidentified Muons	Fake Muon Data

Table 16: Fit components of the normalisation fit.

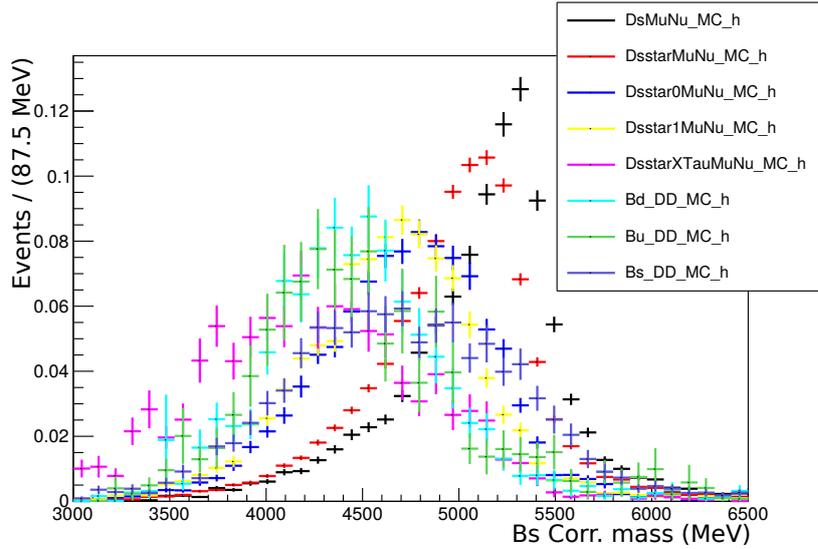


Figure 49: Individual fit templates from MC in the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ fit normalized to unit area. Figure modified from [9].

sign $D_s^- \mu^-$ sample which is assumed to be proxy for a purely combinatorial component since only very few decays contain a same sign D_s^- and muon. Since no D_s^- peak is seen in the $K^- K^+ \pi^-$ invariant mass distribution of this sample, this contribution is neglected in the fit to the corrected mass.

All fit components together with their template sources used in the fit to the corrected $D_s^- \mu^+$ mass are summarised in Table 16 and shown in Figure 49 for comparison.

Backgrounds with similar shape in the corrected mass distribution are combined into a common template since they can not be disentangled by the fit to the corrected mass distribution. A comparison of similar template shapes are given in Figure 50. On the left side backgrounds from $B_s^0 \rightarrow D_{s0}^{*-} \mu^+ \nu_\mu$, $B_s^0 \rightarrow D_{s1}^{*-} \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} X$, where one D_s is decaying into $K^- K^+ \pi^-$ and the other one semileptonically, are shown. Those are combined into a single template in the fit.

Another combined component consists of background from $B^0 \rightarrow D_s^{(*)-} D^{(*)+}$, $B^- \rightarrow D_s^{(*)-} D^{(*)0}$ decays, where $D^{(*)-,0}$ are decaying semileptonically, together

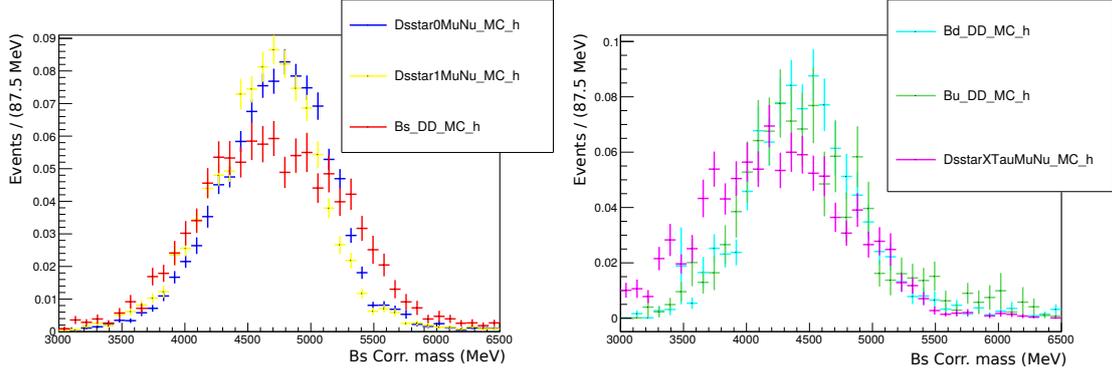


Figure 50: Similar background shapes are combined into a single template in the fit. Those are $B_s^0 \rightarrow D_{s0}^{*-} \mu^+ \nu_\mu$, $B_s^0 \rightarrow D_{s1}^{*-} \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ on the left side as well as $B^0 \rightarrow D_s^{(*)-} D^{(*)+}$, $B^- \rightarrow D_s^{(*)-} D^{(*)0}$ and $B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$ decays on the right side. Figure modified from [9].

with $B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$ decays, here the X stands for additional neutral particles such as pions and photons coming from higher excited D_s^- states. Their templates are shown on the right of Figure 50.

The MC templates are corrected for kinematic differences between data and simulation using the BDT reweighter as explained in Section 7, as well as for PID and tracking corrections as explained in Section 10.4 and 10.5. All component are left free in the fit apart from the misidentified muon background which is fixed to the yield determined from the fake muon data of around 570 events.

8.4 Background subtraction

In order to remove the $K^+ K^- \pi^+$ combinatorial contribution from real D_s^+ in data a background subtraction is needed. Due to correlations between the $K^+ K^- \pi^+$ invariant mass and the $D_s^- \mu^+$ corrected mass the *sPlot* method for subtracting backgrounds cannot be used. Instead the data is divided into n smaller subsets where each set corresponds to a specific bin in the $D_s^- \mu^+$ corrected mass spectrum. Then a binned maximum likelihood fit is performed to the $K^+ K^- \pi^-$ invariant mass distribution for each of those datasets. The D_s^- shape is modelled by a double-Gaussian component and the combinatorial background by an exponential function. Therefore the yield in each bin of the corrected $D_s^- \mu^+$ mass template as plotted in green in Figure 50 (right) is the result of this fit to the $K^- K^+ \mu^-$ invariant mass to determine the D_s^+ yield and to subtract the combinatorial background. The templates used in the fit contain 40 bins in corrected mass ranging from 3000 MeV to 6500 MeV. The corresponding 40 fits to the $K^- K^+ \mu^-$ invariant mass are shown in Figure 51.

8.5 Fit results

A maximum likelihood fit to the $D_s^+ \mu^-$ corrected mass distribution for all events passing the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ selection, as described in Section 6, is performed. The

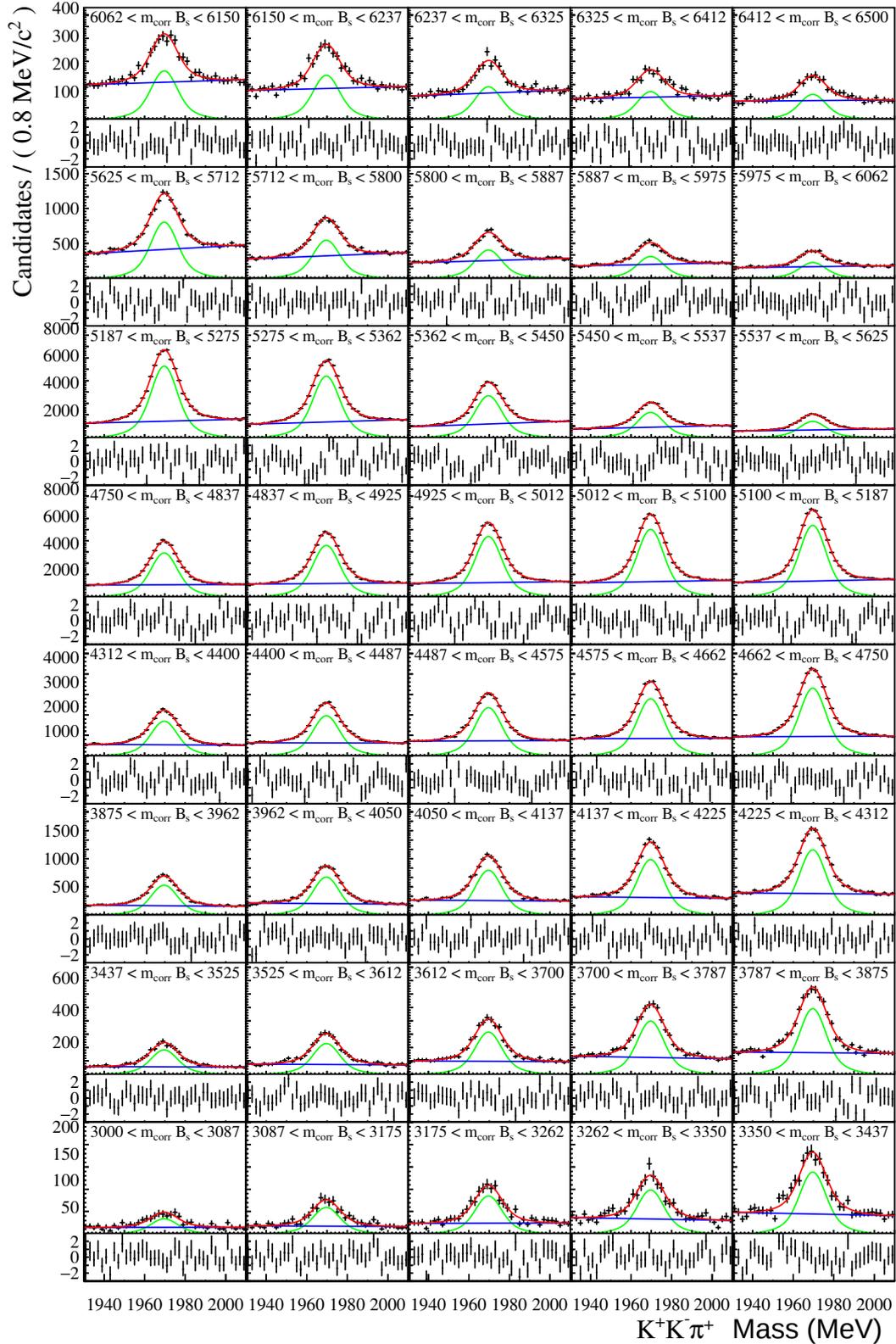


Figure 51: Fits performed as part of a combinatorial background subtraction on the $K^+K^-\pi^+$ invariant mass with pulls underneath. Figure taken from [9].

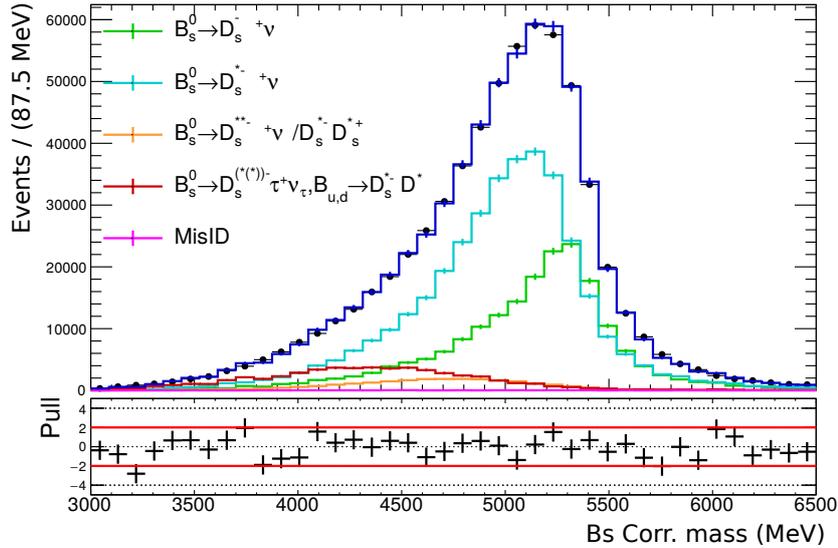


Figure 52: Fit to the corrected $D_s^- \mu^+$ mass. The error bars display the template uncertainty due to the finite Monte Carlo statistics. Figure modified from [9].

Sample	Yield / 10^3
$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	191.0 ± 6.5
$B_s^0 \rightarrow D_s^{*-} \mu^+ \nu_\mu$	385.5 ± 9.8
$B^{0,-} \rightarrow D_s^{(*)-} D^{(*)0,-}, B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$	54.6 ± 2.8
$B_s^0 \rightarrow D_{s0,1}^{*-} \mu^+ \nu_\mu, B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	23.8 ± 6.7

Table 17: Fit results for all components of the fit to the $D_s^- \mu^+$ corrected mass distribution.

result is shown in Figure 52 together with its pulls including the data and template statistical uncertainties defined as $(n_{\text{data}}^i - n_{\text{model}}^i) / \sqrt{(\sigma_{\text{data}}^i)^2 + (\sigma_{\text{model}}^i)^2}$. The signal and background yields obtained from the fit are given in Table 17. Their correlation matrix is given in Table 18, where one can see a large anti-correlation between the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^{*-} \mu^+ \nu_\mu$ component.

The result of the $D_s^- \mu^+$ corrected mass fit is validated by performing 1000 toy fits to pseudo-data. The data template in each pseudo-data fit is replaced with a *toy template* generated from the fit templates with the same number of events per template as given in Table 17. Therefore the yield of each fit component is known precisely. The distribution of the pulls for all 1000 fits to the pseudo-data is plotted in Figure 53. 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 $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ events in the pseudo-data. The pulls should be centred at zero and follow a Gaussian distribution with a width of one. As shown in Figure 53 the pull distribution of the toy fits is centred around zero but has a width of smaller than one. This is due to the inclusion of the statistical uncertainties of the templates from the Beeston-Barlow method which

	Param.	1	2	3	4	5
$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	1	1	-0.858	0.461	-0.263	-0.002
$B_s^0 \rightarrow D_s^{*-} \mu^+ \nu_\mu$	2		1	-0.769	0.471	-0.004
$B^{0,-} \rightarrow D_s^{(*)-} D^{(*)0,-}, B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$	3			1	-0.819	-0.001
$B^{0,-} \rightarrow D_s^{(*)-} D^{(*)0,-}, B_s^0 \rightarrow D_s^- \tau^+ \nu_\mu X$	4				1	0.001
MisPID	5					1

Table 18: Fit correlation coefficients between the different templates of the $D_s^+ \mu^-$ corrected mass fit which are left free in the fit.

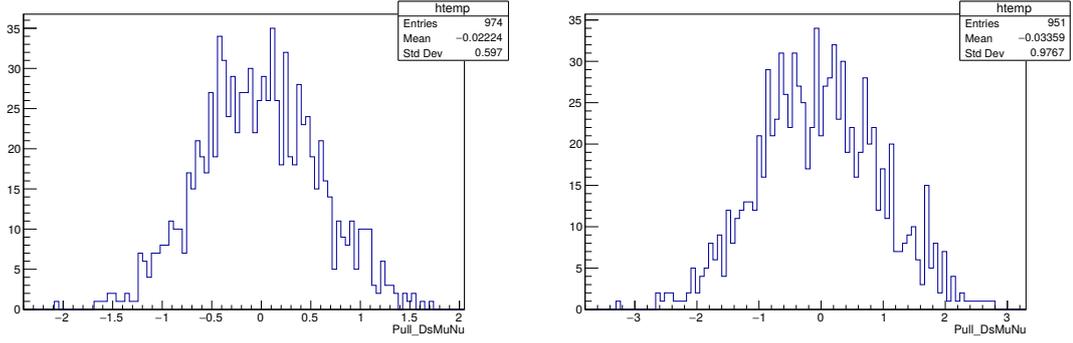


Figure 53: Distributions of pulls from fits to 1000 pseudo datasets for the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ yield, in the left plot including statistical template uncertainty whereas the right plot they are excluded. Figure taken from [9].

overestimates the statistical uncertainties which is a known feature. When turning off the statistical uncertainty in the fit, a pull width value of one is recovered as shown in right of Figure 53.

9 Fit to the signal channel

This chapter presents the maximum likelihood fit performed to extract yield of the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal channel. A binned template fit is performed to the corrected mass of the $K^- \mu^+$ candidates in data, passing all selection cuts outlined before in Section 6. As input to this fit MC samples are used to describe the signal as well the remaining background components, which are corrected by the corrections determined from the control channel as explained in Section 7. An overview of the used fit model as well as the different components to perform this fit is given in Section 9.1, which again makes use of the `HistFactory` tool explained in detail in the previous section. Some of the simulated background templates are combined with each other due to their similar shape in the corrected mass which is described in Section 9.2. The modelling of the combinatorial background for this analysis is very challenging and a new method was developed from data which is outlined in Section 9.3. Misidentified backgrounds are estimated using a specific technique as detailed in Section 9.4. The fit results to the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ channel are shown and discussed at the end of the section.

9.1 Signal fit model

In order to determine the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ yield, a simultaneous fit is performed to two bins of q^2 , with the bin boundary placed at $q^2 = 7\text{GeV}^2/c^4$. The two yields will be used to extract the branching fraction of the signal in the two bins. As for the normalisation fit discussed in the previous section, the signal fit to the corrected mass of the $K^- \mu^+$ is implemented using the `HistFactory` tool included in the `Roofit` package as a binned template fit. The same implementation as for the normalisation channel is used which is based on the Beeston-Barlow light method [118] to incorporate MC template statistical uncertainties. A corrected mass range of $2500 \text{ MeV}/c^2 < m_{\text{corr}} < 5750 \text{ MeV}/c^2$ is used in the fit to $K^- \mu^+$ candidates. Candidates in data are selected by passing all selection cuts for the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal channel described in Section 6. Even after these tight selection cuts a large number of background sources remain.

Dangerous background contributions come from B_s^0 decays to excited K^* resonances, where the K^* further decays into a kaon and a neutral pion. Partially reconstructed decays involving charmonium resonances such as $B_q \rightarrow c\bar{c}K^+ X$, where the charmonium further decays into $(\mu^+ \mu^-)$ are critical. Those are dominated by $B^+ \rightarrow J/\psi K^+$ decays which have a similar template shape as the signal. A large background component comes from $|V_{cb}|$ transitions such as $B_q \rightarrow D_q \mu \nu X$ decays, where the D -meson further decays via a kaon. Those decays have a much larger branching fraction compared to the signal, but their candidates peak at lower corrected mass as they are only partially reconstructed. Combinatorial background from random combinations of a kaon and muon and candidates containing misidentified kaon and muon particles are also present in the data after all selections cuts but their contribution is much smaller.

All fit components used for the signal fit together with their template sources are summarised in Table 19. The templates originating from Monte Carlo simulation

Component	Source
$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ (Signal)	Monte Carlo
$B_s^0 \rightarrow K^{*-} \mu^+ \nu_\mu$	Monte Carlo
$B_s^0 \rightarrow K_0^{*-}(1430) \mu^+ \nu_\mu$	Monte Carlo
$B_s^0 \rightarrow K_2^{*-}(1430) \mu^+ \nu_\mu$	Monte Carlo
$B^- \rightarrow J/\psi K^-$	Monte Carlo
$B^- \rightarrow J/\psi K^{*-}$	Monte Carlo
$B^0 \rightarrow J/\psi K^{*0}$	Monte Carlo
Combinatorics	$K^- \mu^+$ event mixing (data-driven)
Misidentified particles	Fake muon and kaon sample (data-driven)
$b \rightarrow c \rightarrow s$	inclusive Monte Carlo cocktail

Table 19: Fit component of the Signal Fit to extract the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ yield.

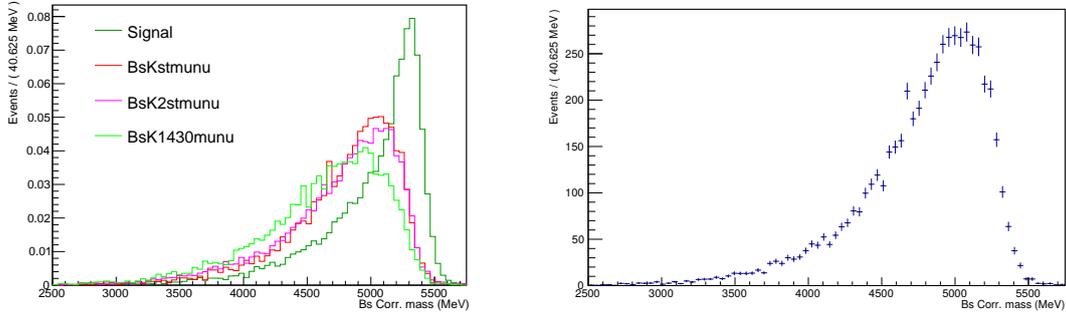


Figure 54: B_s^0 corrected mass template shapes of the three excited kaon resonances considered in the signal fit compared to the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decay (left). In right their combined template is shown which is used as an input template to the signal fit.

are corrected for differences in kinematic and mismodelling between simulation and data using the BDT reweigher introduced in Section 7.

9.2 Combined MC templates

Three excited kaon resonances K^{*-} , $K_2^{*-}(1430)$ and $K_0^{*-}(1430)$ are considered as backgrounds for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decays, where each of the excited kaons decays to $K^- \pi^0$. Their shape in corrected mass is plotted in left of Figure 54 together with the signal template. Due to their similar shape those templates are combined into a single template in the fit as the fit is not able to distinguish between the different excited K^* components. The combination of those templates is done taking into account their different spin structure, such that the K^* is a vector and contributes with 3 different helicity states to the decay, $K_0^{*-}(1430)$ is a scalar and $K_2^{*-}(1430)$ is a tensor. Therefore isospin relations of the PDG [12] from other decays such as $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ with respect to $B^+ \rightarrow K^+ \mu^+ \mu^-$ has be used to relate their relative expected branching fraction as there are no theory predictions available. Including the different branching fractions of the K^* resonances to decay into the final state $K^- \pi^0$ from the PDG [12], this leads to the following fractions between

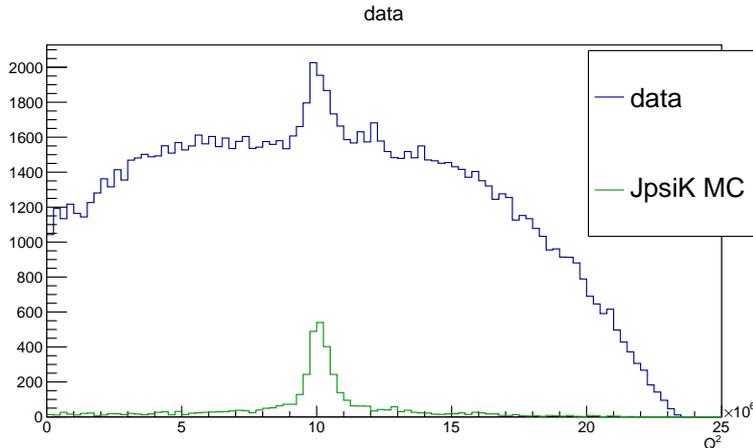


Figure 55: q^2 -distribution of data passing all selection cuts (blue) compared to simulated $B^+ \rightarrow J/\psi K^+$ events (green).

the components in the template:

$$0.64 \times K^{*-} + 0.24 \times K_0^{*-}(1430) + 0.12 \times K_2^{*-}(1430) \quad (73)$$

The final combined template which is used as input template to the signal fit is shown in right of Figure 54. It will be called the *Kstmunu* component from now on.

Another dangerous background source comes from partially reconstructed $B_q \rightarrow c\bar{c}(\rightarrow \mu^+\mu^-)K^+X$ decays, where $c\bar{c}$ stands for different charmonium resonances and the X denotes other light hadrons. Those include decays like $B^+ \rightarrow J/\psi K^+$, $B^+ \rightarrow \psi(2S)K^+$, $B_s^0 \rightarrow J/\psi \phi(\rightarrow K^+K^-)$ but also higher excited kaon resonances such as $B^+ \rightarrow J/\psi K^{*+}$ and $B^0 \rightarrow J/\psi K^{*0}$. Those type of backgrounds are very suppressed by the isolation and BDT cuts described in detail in Section 6, but due to their high branching fraction with respect to the signal decay a large amount still passes the selection. This can be seen from the q^2 -distribution of data passing all selection cuts shown in Figure 55, where a peak at around $10 \text{ GeV}^2/c^4$ is visible which corresponds to the nominal J/ψ mass squared. For illustration purpose it is compared to the simulated q^2 -distribution of $B^+ \rightarrow J/\psi K^+$, which is the dominant contribution of this kind of background as only one particle is missing. The other decays mentioned above which also contribute to the so-called *B2CC* decays are further suppressed by the isolation algorithm and BDTs due to their additional charged particles. Also the branching fraction of higher charmonium states decaying into $\mu^+\mu^-$ is smaller. Therefore only three templates are considered to describe the B2CC component in the fit: $B^+ \rightarrow J/\psi K^+$, $B^+ \rightarrow J/\psi K^{*+}$ and $B^0 \rightarrow J/\psi K^{*0}$. Their shape is compared to the signal in Figure 56. Due to their similar shape they are combined into a single template in the fit, which is shown in right of Figure 56. Their combination is based on their relative branching fraction.

The large background from $|V_{cb}|$ transitions of $B_q \rightarrow (D_q \rightarrow K)\mu\nu X$ decays is described by an inclusive simulated sample selecting $H_b \rightarrow (H_c \rightarrow K^+X)\mu^- X'$

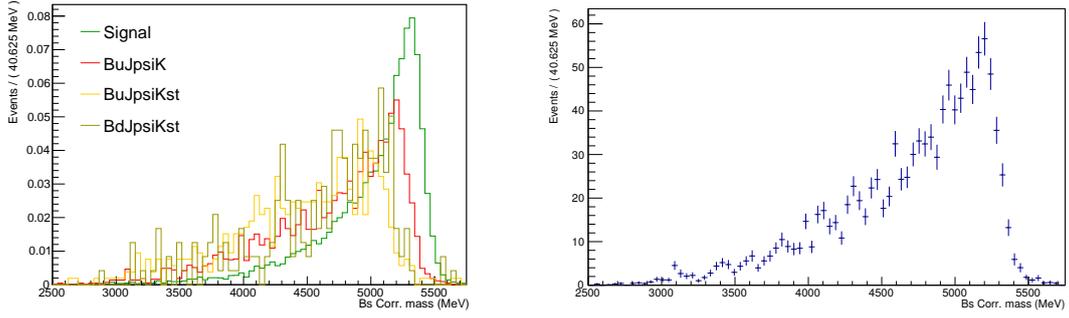


Figure 56: B_s^0 corrected mass template shapes of the three considered B2CC templates in the signal fit compared to the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decay (left). In right their combined template is shown which is used as an input template to the signal fit.

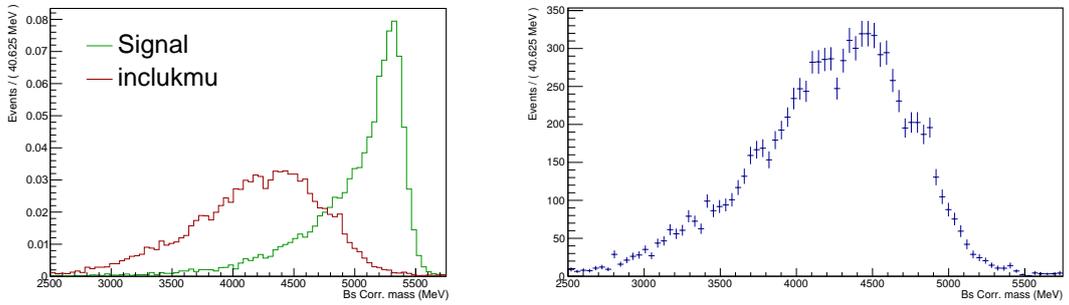


Figure 57: B_s^0 corrected mass template shapes of $|V_{cb}|$ transitions such as $B_q \rightarrow (D_q \rightarrow K) \mu \nu X$ denoted here as inlukmu compared to the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal decay (left). In right their combined template is shown which is used as an input template to the signal fit.

events, where H_b and H_c stands for bottom and charm hadrons. The composition of this sample is given in Table 20, for each listed decay the charm hadron has to further decay into a kaon with a opposite charge to the muon. A large MC sample was generated of around 83M events to describe this component, after all selection cuts outlined in Section 6 only around 10k events are left. The corrected mass shape of this component is shown in Figure 57 compared to the signal template. This type of background will be called *inlukmu* in the following chapter.

9.3 Combinatorial background

The selection cuts discussed in Section 6 remove already a large component of the combinatorial background, especially the quadrant cut removes the low p_T combinatorial and the corrected mass error cut is very powerful to further suppress this kind of background at high corrected mass. The remaining combinatorial background will be discussed in this section.

The usual approach to use $\mu^+ K^+$ SS data as a proxy for this background component and removing the misidentified background from it does not work here, as additional physics background contribute, which are not present in the

Decay	fraction [%]
$B^0 \rightarrow D^0 \mu^- \nu_\mu X$	10.02 ± 0.04
$B^0 \rightarrow D^+ \mu^- \nu_\mu X$	9.92 ± 0.04
$B^+ \rightarrow D^0 \mu^- \nu_\mu X$	19.76 ± 0.05
$B^+ \rightarrow D^+ \mu^- \nu_\mu X$	3.50 ± 0.03
$B_s^0 \rightarrow D^0 \mu^- \nu_\mu X$	1.17 ± 0.01
$B_s^0 \rightarrow D^+ \mu^- \nu_\mu X$	2.69 ± 0.02
$B_s^0 \rightarrow D_s^+ \mu^- \nu_\mu X$	23.67 ± 0.06
b baryon semileptonic	4.72 ± 0.03

Table 20: Decomposition of the inclusive MC sample selecting $H_b \rightarrow H_c(\rightarrow K^+ X)\mu^- X'$ events used to model $|V_{cb}|$ background.

$\mu^+ K^-$ OS data. Those include Cabibbo-favoured decays such as $B^+ \rightarrow \bar{D}^0 \mu \nu$, $B^+ \rightarrow D^- \mu^+ \pi^+ \mu_\nu X$, $B^0 \rightarrow D^- \mu^+ \mu_\nu$, $B^0 \rightarrow D^0 \mu \nu X$, $B^0 \rightarrow D^* \mu^+ \nu_\mu X$, where the charm meson then further decays via 2-, 3- and 4-body decays including neutral pions. As a first attempt it was tried to subtract all these additional components not present in the OS sample from the SS data, but this procedure was discarded later since it is very error prone. Also the subtraction would need templates for every component which is very computing resource intensive and the statistics of the SS sample passing all selection cuts is not enough as it includes a pre-scale factor of 0.1 in data. Therefore another method is needed to model the combinatorial background.

Another common approach is to use events outside the reconstructed mass peak in the sidebands above or below the peak. Due to the only partially reconstructed signal decay including the neutrino there is no defined mass peak. By using the corrected mass there are regions in phase space above the nominal mass of the B_s^0 meson $m_{K^- \mu^+} > m_{B_s^0}$ where a pure combinatorial sample can be obtained. This however can only be used in that region above the B_s^0 mass and thus not in the whole phase space considered in the fit to the corrected mass.

For this analysis a new method was developed to model the combinatorial background, the so-called *event mixing* method. Here a candidate kaon track from one event in data is combined with a candidate muon from another event and the B_s^0 meson is reconstructed from this combination. The method is validated by comparing the kinematics of these mixed data events with pure combinatoric candidates in data above the B_s^0 mass, as outlined before.

The event mixing uses OS 2011 data events passing the stripping selections outlined in Section 6. Those are split based on B_s^0 , K^- , μ^+ candidate information and the underling event. A new combinatoric B_s^0 candidate is created by combining a muon and kaon from different events together, as illustrated in Figure 58 and by recalculating its kinematic quantities. In order to determine the secondary vertex (SV) of the combinatoric B_s^0 , the flight distance (FD) of the combined B_s^0 is randomly sampled from candidates above the B_s^0 mass ($m_{K^- \mu^+} > m_{B_s^0}$) and then used together with the primary vertex to calculate the secondary vertex.

Event #	n-1	n	n+1	n+2	n+3	n+4	n+5	n+6	n+8	n+9	n+10	n+12	n+15
Combi. 1		B_s^0	μ^+	K^-	Ev.								
Combi. 2		B_s^0		μ^+		K^-		Ev.					
Combi. 3		B_s^0			μ^+			K^-		Ev.			
Combi. 4		B_s^0				μ^+			K^-			Ev.	
Combi. 5		B_s^0					μ^+				K^-		Ev.

Figure 58: Each event is used to generate five combinatoric combinations. For the first combination consecutive events are mixed, for the second combination alternate events are used. Figure taken from [9].

Having the SV allows to calculate the corrected mass, whereas the corrected mass uncertainty is determined using a toy study: the position of the secondary vertex, sampled from the FD, is randomly varied within its quoted uncertainty where a new value of the corrected mass calculated each time. The uncertainty on the corrected mass is then the standard deviation of the corrected mass evaluated with those toys. The combined combinatoric B_s^0 candidates are then passed through a loose selection mimicking the preselection cuts given in Section 6.

The event mixing method does not reproduce the kinematics of the B_s^0 meson in the region $m_{K-\mu^+} > m_{B_s^0}$, therefore a two dimensional reweighting is used to correct the momentum and transverse momentum of the B_s^0 candidate. This is shown in Figure 59 with and without the kinematic reweighting applied. The B_s^0 corrected mass distribution as the fit variable and the invariant $K^-\mu^+$ mass are invariant under such a kinematic correction, as shown in bottom of Figure 59. The B_s^0 corrected mass for the whole mass range considered in the signal fit is shown in Figure 60, this template is used as an input to the signal fit. More validation plots of the event mixing method can be found in Appendix D. This component will be called *combi* in the following chapter.

9.4 Misidentified background

Tight PID selections cuts are applied to select the kaon and muon from the $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ signal as described in Section 6. Other particle types such as protons or pions can pass the selection and can be misidentified as kaons or muons. Those kind of backgrounds have to be estimated to determine their yield as well as their corrected mass shape. This can be done using a data-driven method, which is similar to the one used by the $\Lambda_b^0 \rightarrow p\mu^-\nu$ analysis [123]. In general the contribution of the misidentified background will be smaller with respect to the $\Lambda_b^0 \rightarrow p\mu^-\nu$ one due to a better kaon identification efficiency compared with the proton identification used in [123].

Two components contribute to the misidentified background for $B_s^0 \rightarrow K^-\mu^+\nu_\mu$: particles misidentified as kaons and as muons, where the first one is the dominant contribution. Events where both particles are misidentified are very rare and will be neglected. To obtain misidentified background templates, data samples without kaon and muon PID requirements are used, those are selected by the 'fake' K and μ stripping lines: `StrippingB2XuMuNuBs2K_FakeKLine`, `StrippingB2XuMuNuBs2K_FakeMuLine`. From those samples enriched regions of π , μ , p for the misidentified K^+ component and π , K , p for the misidentified μ^-

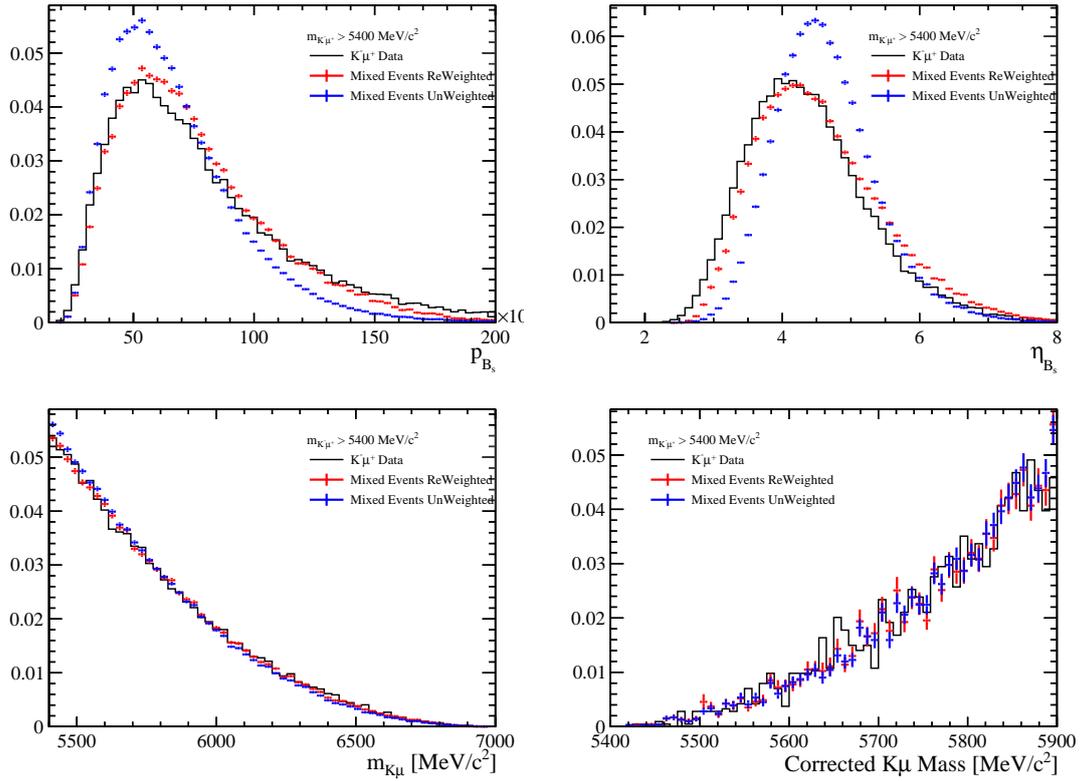


Figure 59: Distributions of mixed events before (blue) and after (red) kinematic reweighting compared to the data in the region above the B_s^0 mass $m_{K-\mu^+} > m_{B_s^0}$ (black) for the B_s^0 momenta (top left), pseudorapidity (top right), invariant mass (bottom left) and corrected mass (bottom right). Figure taken from [9].

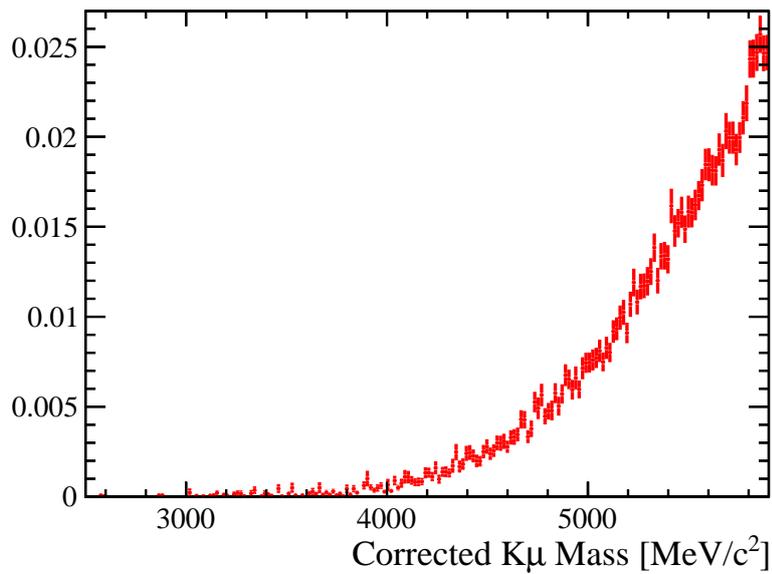


Figure 60: The corrected mass of the $K^-\mu^+$ pair plotted for candidates modelling the combinatoric background. Figure taken from [9].

Region	Cuts applied to FakeKLine
Signal	$\Delta\log\mathcal{L}_{K\pi} > 5 \ \& \ \Delta\log\mathcal{L}_{Kp} > 5 \ \& \ \Delta\log\mathcal{L}_{K\mu} > 5$
π	$\Delta\log\mathcal{L}_{K\pi} < 0 \ \& \ \Delta\log\mathcal{L}_{Kp} < 0 \ \& \ \Delta\log\mathcal{L}_{K\mu} < 0$
p	$\Delta\log\mathcal{L}_{K\pi} < 5 \ \& \ \Delta\log\mathcal{L}_{Kp} > 0 \ \& \ \Delta\log\mathcal{L}_{K\mu} < 0$
μ	$\Delta\log\mathcal{L}_{K\pi} < 0 \ \& \ \Delta\log\mathcal{L}_{Kp} < 0 \ \& \ \Delta\log\mathcal{L}_{K\mu} > 0$
Region	Cut applied to FakeMuLine
Signal	$\Delta\log\mathcal{L}_{\mu\pi} > 3 \ \& \ \Delta\log\mathcal{L}_{\mu p} > 0 \ \& \ \Delta\log\mathcal{L}_{\mu K} > 0$
π	$\Delta\log\mathcal{L}_{\mu\pi} < 3 \ \& \ \Delta\log\mathcal{L}_{\mu p} > 0 \ \& \ \Delta\log\mathcal{L}_{\mu K} > 0$
p	$\Delta\log\mathcal{L}_{\mu\pi} < 3 \ \& \ \Delta\log\mathcal{L}_{\mu p} > 0 \ \& \ \Delta\log\mathcal{L}_{\mu K} < 0$
K	$\Delta\log\mathcal{L}_{\mu\pi} < 3 \ \& \ \Delta\log\mathcal{L}_{\mu p} < 0 \ \& \ \Delta\log\mathcal{L}_{\mu K} > 0$

Table 21: PID requirements applied to obtain enriched regions for different particles for the FakeKLine (top) and μ^- FakeMuLine (bottom).

component can be obtained. Table 21 shows the PID cuts used to define such regions. The yield of misidentified events estimated in the signal region for a given particle type i (e.g π) is:

$$\mathcal{Y}_i = \frac{N(\text{region})_i \times R(i \rightarrow K)}{\epsilon(i, \text{region}_i)}, \quad (74)$$

where $N(\text{region})_i$ is the number of candidates passing the PID cuts in Table 21 for the different particle type enriched regions, $R(i \rightarrow K)$ is the misidentification rate per particle type in the signal region, and $\epsilon(i, \text{region}_i)$ the efficiency to select particle-type i in the corresponding enriched region.

A complication is that the enriched regions themselves are not pure samples per particle type i , they will rather have a contamination of other hadron species. This cross-feed between the different region and various particle types has to be accounted for in the previous equation and leads to

$$N_{\text{region}_i} = \sum_j N_{\text{region}_j}^{\text{True}} \times \epsilon(j, \text{region}_j). \quad (75)$$

Since $\epsilon(j, \text{region}_j)$ can be taken from PIDCalib as given in Table 22 for the different enriched regions, the number of true events per enriched region of particle type $N_{\text{region}_j}^{\text{True}}$ from Equation 75 can be determined using the matrix inversion method. Using Equation 74 and including the misidentification rate ($R(i \rightarrow K)$) from PID-calib, this gives the yield for different misidentified particles in the signal region. The total yield for misidentified kaons and muons is then obtained from summing up the different components, which is around 2600 for kaon and 100 formuon candidates. Their distribution in the B_s^0 corrected mass variable is shown in Figure 61. Both components are added together to get the final yield of the misidentified component in the signal region as well as its template shape as input for the signal fit. The yield of this component is Gaussian constraint in the fit.

particle type eff. [%]	enriched region			
	K	π	μ	p
K	57.4	1.0	0.3	1.4
π	3.2	70.6	4.5	2.0
μ	0.6	10.1	76.4	0.51
p	2.7	9.7	0.9	30.1
K	97.9	18.1	3.3	85.6
π	89.6	89.1	3.9	82.2
μ	0.2	1.8	80.5	0.2
p	70.9	17.4	5.5	93.8

Table 22: PIDCalib efficiency for each particle-type in every enriched region calculated for the kaon (top) and muon (bottom).

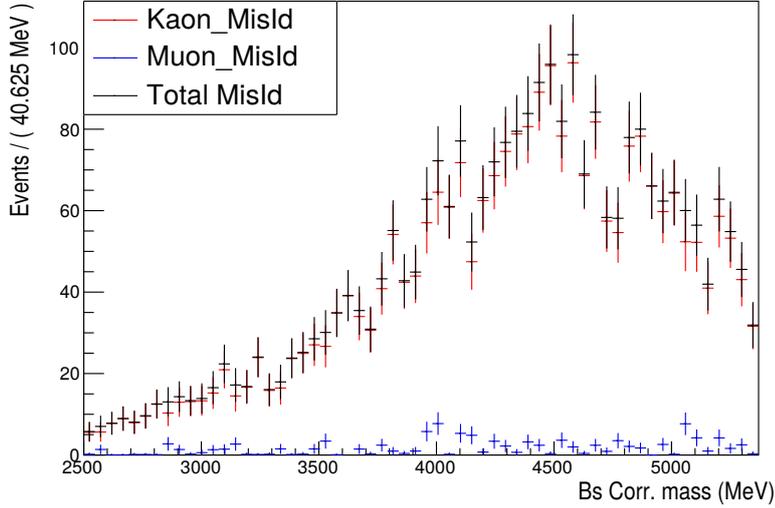


Figure 61: B_s^0 corrected mass distribution for misidentified kaons (red) and muons (blue) and the their total contribution in black.

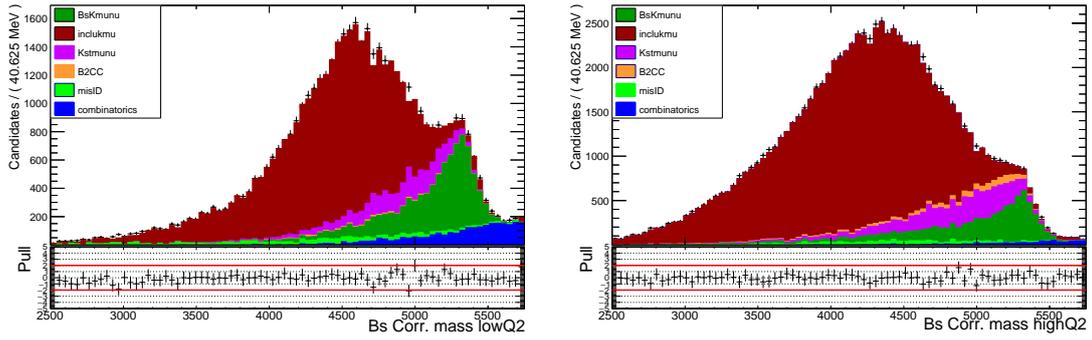


Figure 62: Signal fit to B_s^0 corrected mass in low (left) and high (right) q^2 region. The individual components are stacked on top of each other.

9.5 Fit results

A simultaneous maximum likelihood binned template fit is performed in two bins of q^2 , which are split at $q^2 = 7\text{GeV}^2/c^4$ using the Beeston-Barlow light method implemented in `HistFactory`. The fit contains the six different templates explained before in this section: $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ Signal, the combined Kstmumu and B2CC components, the combinatorial background modelled from event mixing, the so-called includkmu background from $|V_{cb}|$ transitions and the misidentified background obtained from the data driven method. The latter is Gaussian constraint to the yield obtained from this method. One additional constraint is needed for the fit to be able to disentangle the B2CC and Kstmumu components which have a similar shape in corrected mass. The relative fraction of B2CC events between the low and high q^2 bin is constraint from MC simulation to be 12% in the low and 88% in the high q^2 bin. This is a valid assumption since the branching fractions of these decays are well known as well as their kinematic distribution as a function of q^2 . Also due to the J/ψ component in the decay, they mostly peak at the nominal J/ψ mass with only a small tail towards the low q^2 bin as discussed previously in Section 11.3 and shown in Figure 55.

The result of the simultaneous fit to the low and high q^2 bins is shown in Figure 62 together with its pulls including data and simulation statistical uncertainties. Here the different MC templates are drawn stacked on top of each other. The signal component drawn in dark green is clearly visible in both bins, its relative contribution is higher in the low q^2 bin due to a smaller overall background contamination. Also the high q^2 bin shows a clear signal component but here more background from B2CC, Kstmumu and especially from $|V_{cb}|$ transitions are contributing. The numerical fit results are given in Table 23 together with their correlation matrix given in Table 24. There is a large anti-correlation between the B2CC component and the Kstmumu yield in the high q^2 bin.

The fit in Figure 62 is drawn allowing the templates to vary according to their statistical uncertainties per bin according to the Beeston-Barlow method. This deforms templates with large statistical uncertainties and pushes them up or down

Component	low q^2	high q^2
Signal	7277.90 ± 367.98	7415.90 ± 371.39
$B_s^0 \rightarrow K^{*-} \mu^+ \nu_\mu$	3407.16 ± 505.33	6809.36 ± 1399.12
B2CC	224.79 ± 198.00	1674.26 ± 1474.70
Combinatorics	2688.44 ± 135.11	941.78 ± 105.67
Misidentified particles	1431.79 ± 36.89	1534.61 ± 39.13
inclukmu	30401.8 ± 596.61	68798.90 ± 759.52

Table 23: Signal fit results for the simultaneous fit to the B_s^0 corrected mass for both q^2 bins.

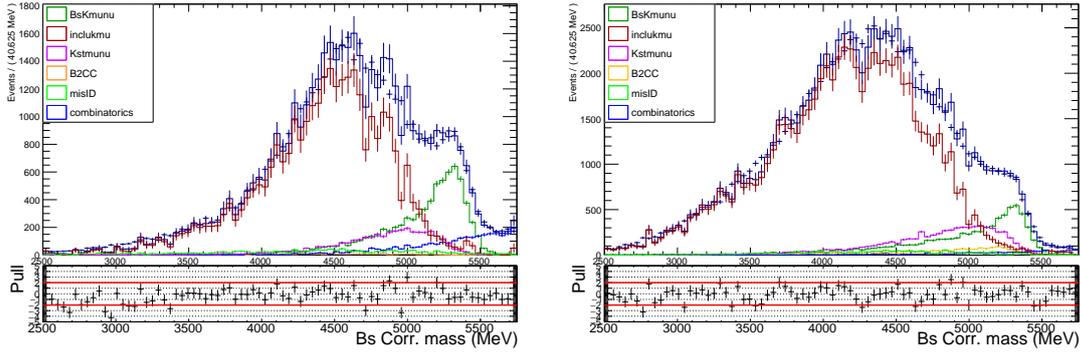


Figure 63: Signal Fit to B_s^0 corrected mass in low (left) and high (right) q^2 region drawn without the Beeston-Barlow method.

to match the data per bin perfectly. Therefore it is difficult to judge if the the fit model describes the data as the template are deformed. As a comparison the fit is also drawn switching off this method, just scaling the input histograms by the fit result given in Table 23, this additional fit is shown in Figure 63. Here the data can also be well described by the different templates and in addition the template statistical uncertainties are indicated as error bars on the templates. The inclkmu template has the largest statistical uncertainties.

The result of the simultaneous signal fit to the B_s^0 corrected mass is validated as for the normalisation fit by performing 1000 fits to pseudo-data. For that the data template in each pseudo-data fit is replaced with a toy template generated by randomly selecting points from the total fit model with the same number of events as in the fitted data. The parameters used to generate each toy are the ones corresponding to the nominal fit results. The Beeston-Barlow procedure is switched off when performing the toys as it overestimates the uncertainties as the limited MC statistics of the templates are propagated to the fit uncertainties. From these 1000 fits to pseudo-data the pull distribution can be determined. The distribution of pulls should follow a normal distribution which is centred at zero with a width of one. An offset would indicate a bias present in the fit and a width differing from one would indicate that the uncertainty of the fitted yield is incorrectly estimated and either over- or underestimated. The distribution of the pulls for all 1000 fits to the pseudo-data of the signal fit is shown in Figure 64.

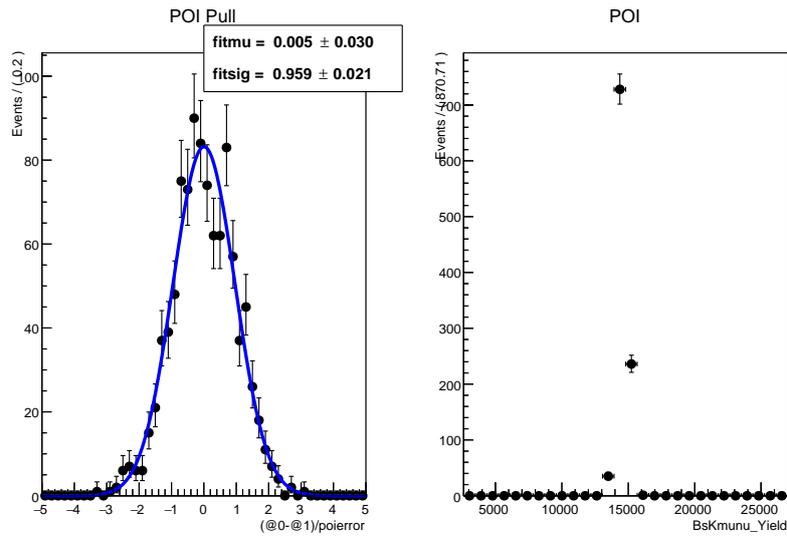


Figure 64: Pull distribution for the signal yield from fits to 1000 pseudo-datasets together with the distribution of the signal yield as the parameter of interest (POI) for these fits.

The pull distribution of the toy fits gives a mean of $\mu = 0.005 \pm 0.030$ and has a width of $\sigma = 0.959 \pm 0.021$ which is compatible with one. Therefore the signal yield and the associated uncertainty are unbiased estimators.

Parameter	1	2	3	4	5	6	7	8	9	10	11
1 B2CC	1.000	-0.346	-0.914	-0.177	0.141	-0.058	0.036	-0.204	0.394	0.009	0.007
2 BsKmuu		1.000	0.086	-0.310	-0.286	-0.192	0.133	0.209	-0.365	-0.008	-0.009
3 Kstmunu highQ2			1.000	0.162	-0.086	0.053	-0.209	0.187	-0.136	-0.013	-0.007
4 Kstmunu lowQ2				1.000	-0.025	0.099	-0.006	-0.540	-0.438	-0.002	0.006
5 combi highQ2					1.000	-0.008	-0.041	-0.029	0.286	-0.006	0.001
6 combi lowQ2						1.000	-0.002	-0.040	-0.233	-0.001	-0.024
7 inclukmu highQ2							1.000	-0.007	-0.127	-0.038	0.000
8 inclukmu lowQ2								1.000	0.057	-0.002	-0.070
9 lowQ2 fraction									1.000	0.008	-0.003
10 misPID highQ2										1.000	0.000
11 misPID lowQ2											1.000

Table 24: Correlation matrix of the yield parameters of the signal fit

10 Relative efficiencies and corrections

In the previous sections we have determined the yields for the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ normalisation channel by fits their corresponding corrected mass distributions. The next step is to convert these yields into a measurement of the relative branching ratio of both decays. Therefore the relative efficiencies have to be determined. The different decay topologies as well as the tight selection cuts applied to the signal lead to different experimental efficiencies for both decays. Moreover for the signal decay three different q^2 regions are selected: $q^2 < 7\text{GeV}^2/c^4$, $q^2 > 7\text{GeV}^2/c^4$ and the full q^2 range. Between the first two regions inwards and outward migration has to be taken into account which is discussed Section 10.6. The relative efficiency is determined from simulated signal and normalisation samples. Differences between data and simulation are corrected from control samples using data-driven methods. The total relative efficiency thus factorises into different components:

$$\begin{aligned} \epsilon_{rel, q^2 \leq 7} &= \frac{\epsilon_{gen}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\epsilon_{gen}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \times \frac{\epsilon_{sel}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\epsilon_{sel}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \times \text{PID corr.} \\ &\quad \times \text{trigger corr.} \times \text{tracking corr.} \times \text{migration corr.}_{q^2 \leq 7\text{GeV}^2/c^4} \end{aligned} \quad (76)$$

These relative efficiencies and applied correction will be described in more detail in this section. The end of the section covers the estimated systematic uncertainties.

10.1 Generator efficiency

Only for MC events where all signal decay products are within the acceptance of the detector, interactions with the detector are simulated. According generator level cuts are directly applied after the generation of the decay and require the polar angle θ_{flight} of all stable charged decay products to be within

$$0.01 < \theta_{\text{flight}} < 0.4. \quad (77)$$

For $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ events generator level cuts are applied to both the kaon and the muon. In the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ events are generated by requiring the muon and the daughters of the D_s^+ to be in acceptance. All other simulated samples used as a proxy for the backgrounds, as described in Section 3.2.4, require all charged tracks to be within the detector acceptance. The generator level cut efficiency has to be determined for the different q^2 regions. To determine this efficiency for the signal and normalisation channel, 250000 events are generated for each decay prior to the generator level cuts. Their generator efficiencies as a function of q^2 are shown in Figure 65. The generator efficiencies are given in Table 25 together with their relative ratio, as explained above for the normalisation channel only the full q^2 range is used whereas for the signal three different q^2 regions are selected.

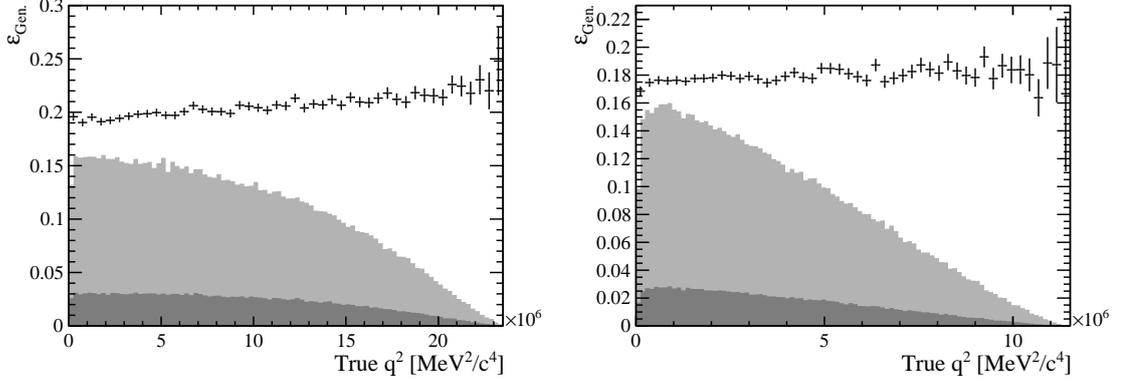


Figure 65: The generator efficiencies plotted as a function of the true q^2 for the signal $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ on the left and the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ on the right. The q^2 distributions are plotted in grey before and after the selections are applied. Figure taken from [9].

Generator efficiency [%]	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	Ratio
Full q^2	20.84 ± 0.083	17.87 ± 0.08	1.166 ± 0.006843
$q^2 < 7 \text{ GeV}^2/\text{c}^4$	19.67 ± 0.14		1.101 ± 0.008975
$q^2 > 7 \text{ GeV}^2/\text{c}^4$	21.13 ± 0.09		1.182 ± 0.007311

Table 25: Generator efficiencies for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decays in different q^2 regions evaluated on Monte Carlo together with their uncertainties coming from the simulation statistics.

Selection eff. [%]	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	Ratio
Full q^2	0.4162 ± 0.0138	0.7817 ± 0.02404	0.5265 ± 0.0270
$q^2 < 7 \text{ GeV}^2/c^4$	0.7790 ± 0.0209		0.9876 ± 0.0405
$q^2 > 7 \text{ GeV}^2/c^4$	0.3242 ± 0.0121		0.4094 ± 0.0233

Table 26: Selection efficiency for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decays in different q^2 regions evaluated on Monte Carlo together with their uncertainties coming from the simulation statistics.

10.2 Selection efficiency

The selection is summarized in Section 6. Its efficiencies are evaluated on simulated samples from the signal and normalisation channel. The resulting efficiencies are given in Table 26 together with their relative ratio.

To ensure that differences between data and simulation are corrected properly taken into account, the decay $B^+ \rightarrow J/\psi K^+$ is used as a control channel as explained in Section 7. The decays of interest, $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$, respectively, are partially reconstructed and have broad distributions making it difficult or even impossible to isolate a pure signal sample in data. Contrary $B^+ \rightarrow J/\psi K^+$ can be selected purely in data using the *sPlot* technique and then its distribution can be compared to simulation. For cuts on kinematics variables such as the corrected mass error cut, the control channel is reconstructed partially by only using one muon of the J/ψ in order to emulate the partially reconstructed signal decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. To evaluate efficiencies for isolation variables, the control channel is fully reconstructed using both muons and the kaon to not have any additional tracks associated to the secondary vertex. The selection efficiency is calculated by performing a fit to the invariant $\mu^+ \mu^- K$ mass before and after the selection cuts are applied, both for s-weighted data and simulation. From this ratio a correction factor is obtained. The simulated signal $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ are then scaled with this correction factor, whereas its uncertainty is applied as a systematic uncertainty. For the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ only the charge isolation cut is applied as explained in Section 6, its correction factor is determined using the same method. The correction factors and the corresponding uncertainties for the different selection cuts are given in Table 27. Due to the kinematic distribution of the control channel $B^+ \rightarrow J/\psi K^+$, its q^2 peaks at the squared of the J/ψ mass, such that only very few events are reconstructed in the low q^2 bin. Therefore only one correction factor for the whole q^2 range can be determined.

Figure 66 shows the distribution of the corrected mass error for simulated signal events together with the simulated control channel and s-weighted data. The black line indicates where the selection cut is placed. The response of the charged track isolation BDT is shown in Figure 67 for the $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ signal and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ normalisation channel together with the control channel. The charged BDT and same-sign (SS) BDT response for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ are shown in Figure 68 alongside with the control channel.

	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$
σ_{mcorr}	1.036 ± 0.005	
Isolation BDT	0.993 ± 0.001	0.988 ± 0.002
Charged BDT	0.989 ± 0.007	
Same Sign BDT	0.959 ± 0.011	

Table 27: Correction factors applied to simulated $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decays determined exploiting the control channel $B^+ \rightarrow J/\psi K^+$.

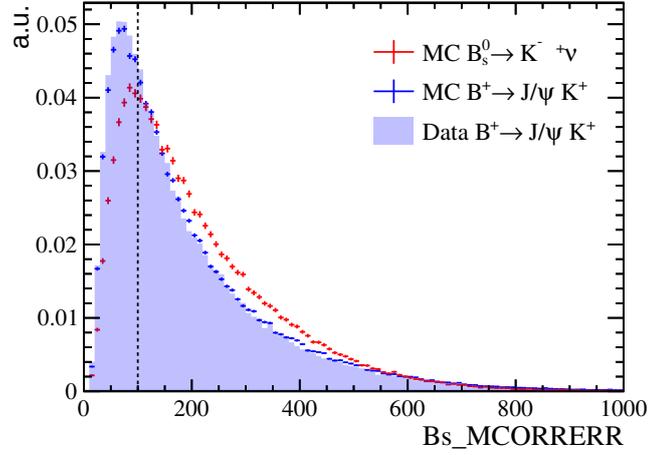


Figure 66: The corrected mass error for $B^+ \rightarrow J/\psi K^+$ in data and simulation together with $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ candidates in simulation. The black line shows where the selection cut is placed. Figure modified from [9].

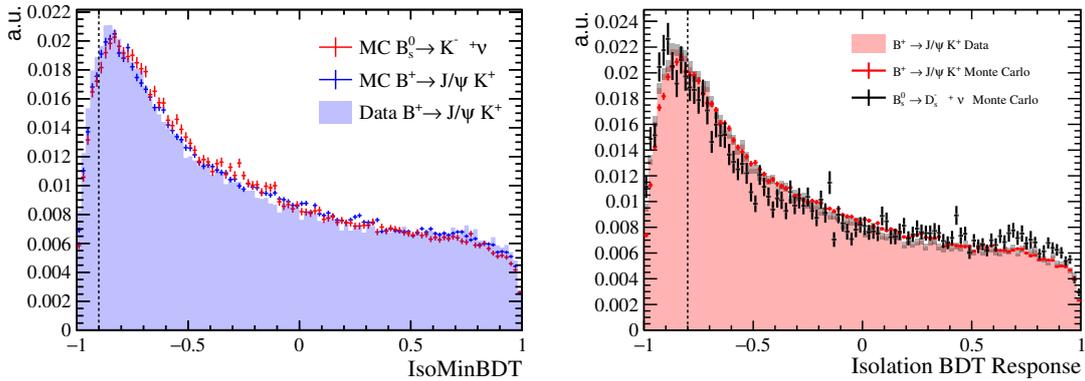


Figure 67: The response of the charged track isolation BDT for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ (left) and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ (right) is plotted together with the $B^+ \rightarrow J/\psi K^+$ calibration samples. The black line shows where the selection cut is placed. Figure modified from [9].

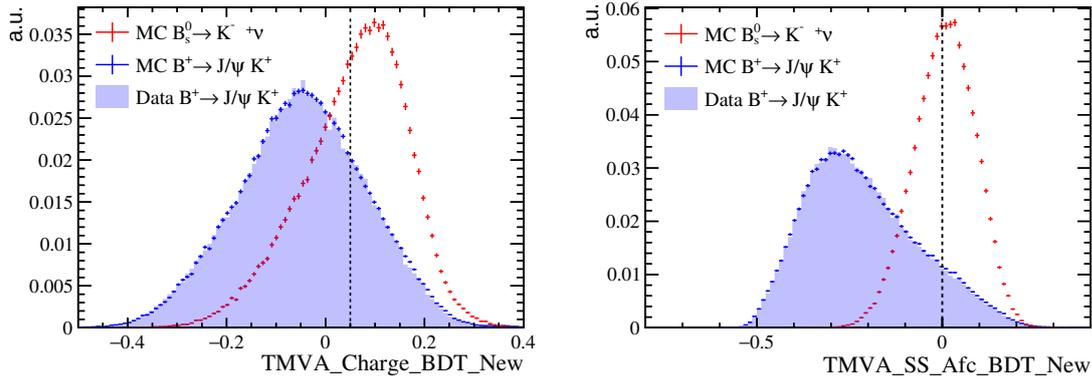


Figure 68: The response of the charged BDT (left) and SS BDT (right) for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ candidates are plotted together with candidates of the $B^+ \rightarrow J/\psi K^+$ control sample. Figure modified from [9].

10.3 Trigger efficiency

In general the trigger efficiency is defined in terms of the selected events

$$\epsilon_{\text{trig}} \equiv \epsilon_{\text{trig|sel}} = \frac{N_{\text{trig|sel}}}{N_{\text{sel}}}, \quad (78)$$

where $N_{\text{trig|sel}}$ are the number of selected events passing the trigger and N_{sel} are the number of events that would be selected without the trigger. Here the problem is that the latter number is unknown in data as only events passing the trigger are stored and can be studied. Instead the so-called *TISTOS* method [125] is used to determine the trigger efficiency in data which will be briefly explained in the following, more details and the validation of the method itself can be found in Reference [125].

Events can be triggered using tracks from the signal candidate, Triggered On Signal (TOS), triggered by tracks not belonging to the signal, Trigger Independent of the Signal (TIS), or triggered on both (TOB). Therefore the trigger efficiency can be rewritten as

$$\epsilon_{\text{trig|sel}} = \frac{N_{\text{trig|sel}}}{N_{\text{TIS|sel}}} \times \frac{N_{\text{TIS|sel}}}{N_{\text{sel}}} = \frac{N_{\text{trig|sel}}}{N_{\text{TIS|sel}}} \times \epsilon_{\text{TIS}}, \quad (79)$$

where $N_{\text{TIS|sel}}$ are the number of TIS events after the selection is applied. The TIS efficiency ϵ_{TIS} cannot be measured from data, but it can be approximated by the TISTOS efficiency

$$\epsilon_{\text{TIS}} \approx \epsilon_{\text{TISTOS}} \equiv \frac{N_{\text{TIS\&TOS}}}{N_{\text{TOS}}} \quad (80)$$

where N_{TOS} are the number of events that are selected by the TOS trigger and $N_{\text{TIS\&TOS}}$ the number of events that are both TIS and TOS. The approximation assumes that the TIS efficiency ϵ_{TIS} is independent of the chosen sup-sample of data, which was proven to be valid in Reference [125].

In this analysis the TISTOS method is used on a subset of its decay product to determine the trigger efficiency. The $K^+ \mu^-$ combinations of the signal

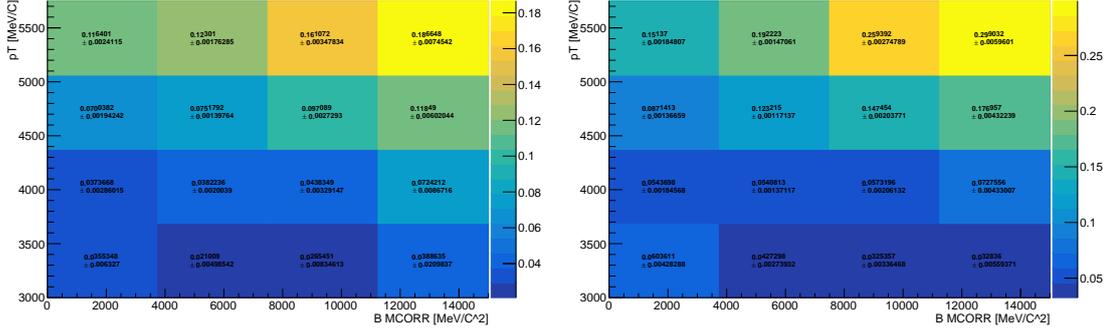


Figure 69: TISTOS efficiency for partially reconstructed $B^+ \rightarrow J/\psi K^+$ decay in bins of B^+ p_T and m_{corr} in background subtracted data (left) and simulation (right). Figure taken from [9].

$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ have to fire the same trigger lines (Hlt2SingleMuonDecisionTOS or Hlt2TopoMu2BodyDecisionTOS), as explained in Section 6. This greatly reduces systematic effects due to the trigger in the measurement of the branching fraction ratio of both decays. The remaining differences in the trigger efficiency between both modes are related to different kinematics or a difference in the corrected mass of the $K^+ \mu^-$ pair.

In order to account for a possible mismodeling of the trigger efficiency in simulation due to kinematic differences between data and MC, the TISTOS efficiency is compared in simulation and data which follows the approach used in the Λ_b decays analysis [123]. The comparison is made by evaluating the TISTOS efficiency in the control channel $B^+ \rightarrow J/\psi K^+$, which is reconstructed as signal such that one muon is ignored. The TISTOS efficiency is evaluated in bins of B^+ p_T and corrected mass for both data and simulation, as shown in Figure 69. Figure 70 shows the ratio of data versus simulation efficiency $TISTOS_{data}/TISTOS_{MC}$ based on the plots from Figure 69.

This difference between simulation and data in the control channel for each 2D bin of B^+ p_T and corrected mass is then used to correct the simulated signal $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ depending on the kinematic distributions. Their simulated 2D kinematic distribution is shown in Figure 71. The correction map in Figure 70 is then used to correct the simulated distributions shown in Figure 71. The calculated correction factors (data/MC) for signal and normalisation are shown on the left side of Figure 72.

To estimate the statistical uncertainty due to this reweighting procedure, the correction map is varied within the bin uncertainties using 1000 pseudo experiments. The resulting distribution of the ratio of signal over normalisation channel is shown on the right side of Figure 72. The mean of the distribution is the trigger correction and the RMS gives the statistical uncertainty. The statistical uncertainty is driven by the statistics of the available data control sample $B^+ \rightarrow J/\psi K^+$. This procedure is repeated for the two q^2 regions and is summarised in Table 28.

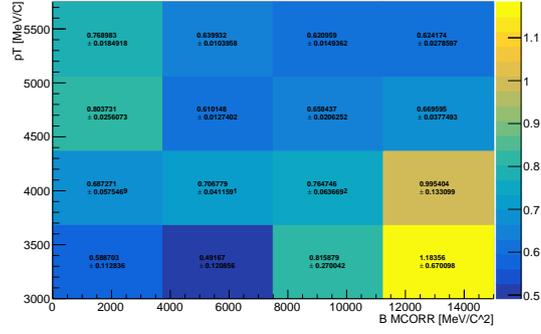


Figure 70: The ratio of data versus MC of the TISTOS efficiency $TISTOS_{data}/TISTOS_{MC}$ for partially reconstructed $B^+ \rightarrow J/\psi K^+$ decays in bins of B^+ p_T and m_{corr} . Figure taken from [9].

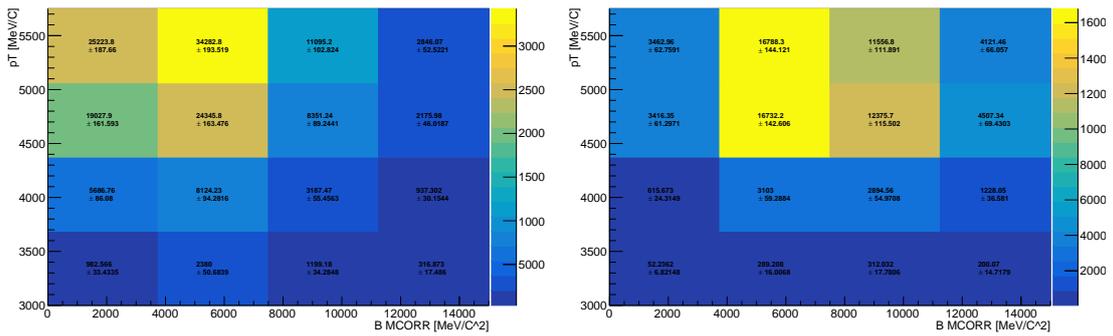


Figure 71: Distributions of B_s^0 p_T and m_{corr} signal (left) and normalisation (right) channel. Figure taken from [9].

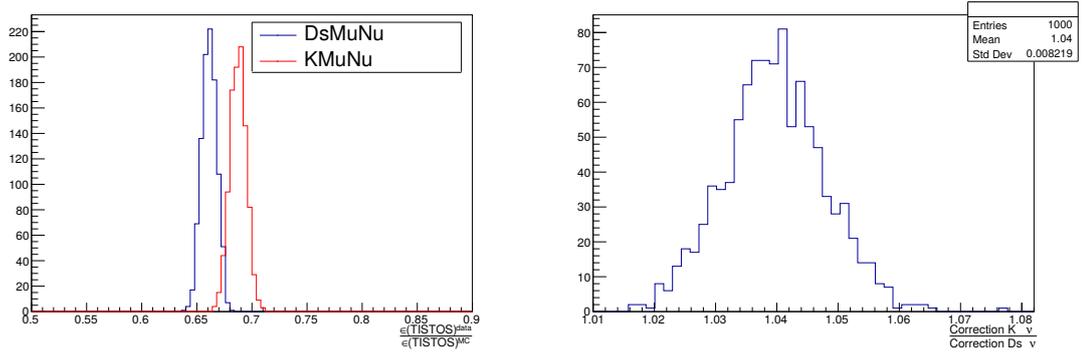


Figure 72: Distribution of correction factors (data/MC) for signal (blue) and normalisation channel (red) on the left side. Distribution of the ratio of correction factors for signal with respect to the normalisation channel for 1000 toy experiments (right). Figure taken from [9].

	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	Ratio
No Sel.	0.687 ± 0.007	0.661 ± 0.007	1.04 ± 0.008
$q^2_{K^- \mu^+} < 7 \text{ GeV}^2/c^4$	0.692 ± 0.007		1.046 ± 0.009
$q^2_{K^- \mu^+} > 7 \text{ GeV}^2/c^4$	0.684 ± 0.009		1.033 ± 0.009

Table 28: Trigger correction factors averaged over all events applied to Monte Carlo samples in bins of q^2 . The ratio in the last column refers to the the ratio of signal versus normalisation channel.

Using $B^+ \rightarrow J/\psi K^+$ decays, other cross checks for TISTOS and TOS efficiencies for individual trigger lines are performed. The variation of the TOS fraction for the trigger decisions in bins of η and p_T are studied, the trends are compatible between data and MC as shown in Appendix E.

10.4 Particle identification

As explained in Section 3.2.2, particles are identified combining the information from different sub-detectors such as the calorimeters, the muon stations and the two RICH detectors. These systems provide an excellent charged particle identification (PID). Modelling their combined response in simulation is not trivial since it depends on many different quantities such as the kinematics of the particles, beam conditions as well as event-level quantities. Therefore data-driven methods were developed to determine PID efficiencies and misidentification rates for different final state particles such as protons, muons, charged pions and kaons, based on calibration data samples. The calibration samples contain candidates which have been selected without the use of any PID information, such that different PID requirements can be tested using the tag and probe method [124]. The calibration samples used are listed in Table 29. Here the PID response is parametrised as a function of particle momentum, pseudo rapidity or track multiplicity of the event. Therefore the determined PID performance of these calibration samples are usable for any tracks. Tools to evaluate those are

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

Table 29: The decays are used by the PIDCalib software package [124] to calibrate particle identification efficiencies. The low momentum (soft) tag π^+ originates from the D^{*+} decay allowing the flavour of the D^0 to be unambiguously identified.

particle type	PID selection cuts
μ^+	$\Delta \log \mathcal{L}_{\mu\pi} > 3$ and $\Delta \log \mathcal{L}_{\mu p} > 0$ and $\Delta \log \mathcal{L}_{\mu K} > 0$
K^-	$\Delta \log \mathcal{L}_{K\pi} > 5$ and $\Delta \log \mathcal{L}_{Kp} > 5$ and $\Delta \log \mathcal{L}_{K\mu} > 5$
K^+	$\Delta \log \mathcal{L}_{K\pi} > -2$
π^-	$\Delta \log \mathcal{L}_{K\pi} < 20$

Table 30: The PID likelihood selections applied to all particles. Selections are aligned between $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ minimising systematics when taking the ratio of efficiencies. While the cuts for μ^+ and K^- are used for the signal and normalisation channel, those for K^+ and π^- are only used for the normalisation mode.

collected within the PIDCalib software package [124] which is provided by the collaboration. This package provides look-up tables for all PID efficiencies and can be used by analysts to extract efficiencies for an analysis specific PID selection.

All PID selections cuts which are applied in this thesis to simulation and data were introduced in Section 6, they are summarized in Table 30. Tight PID selections are only applied to the opposite sign kaon and muon pair to minimise systematic effects, while for the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ also loose selections to the additional opposite sign $\pi^- K^+$ pair are applied. Therefore the PID selection efficiency will be similar for both signal and normalisation decay such that systematic effects mostly cancel when the correction factor the ratio of both is calculated.

Since the PID efficiency depends on track kinematic and underlying event variables, differences in these quantities between the calibration and signal samples could lead to systematic differences in PID efficiencies. Therefore lookup tables are binned in track momentum, pseudo rapidity and track multiplicity providing the PID efficiency in that specific data region. Two dimensional projections of these tables are shown in Figures 73 and 74. In order to minimise systematic effects from intra bin variations in the efficiency and from the *sWeight* background subtraction for the calibration samples, a MC/Data driven correction is used. The MC/Data driven correction returns the ratio of PID efficiencies obtained in data and Monte Carlo samples. This ratio is then used to correct the PID efficiency in the simulation.

To determine the PID efficiency correction each track from each decay in the MC sample is weighted by the correction factor obtained from the ratio of the

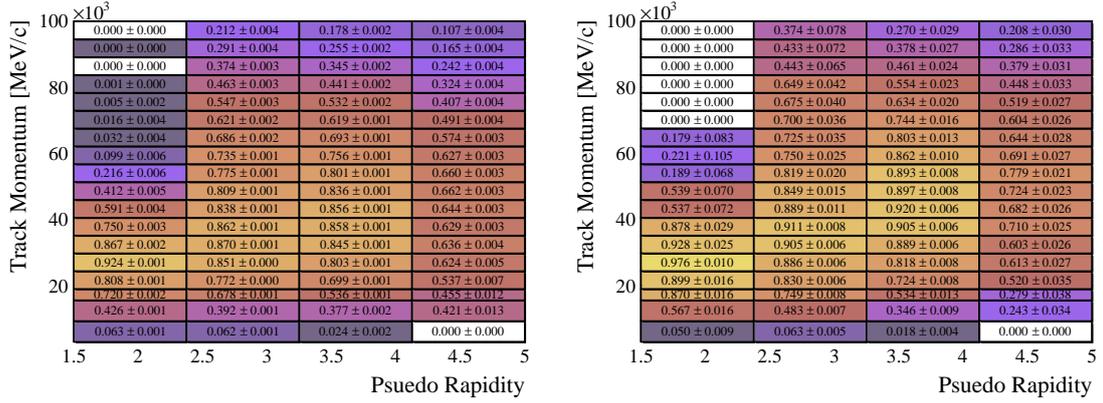


Figure 73: A 2D projection of the PID efficiency lookup table for kaons determined from $D^{*+} \rightarrow (D^0 \rightarrow K^-\pi^+)\pi^+$ decays in data (left) and simulation (right) [124].

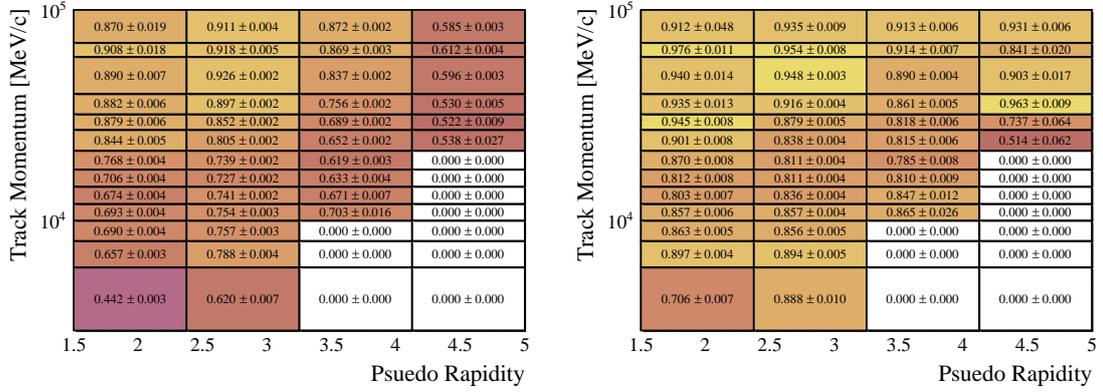


Figure 74: A 2D projection of the PID efficiency lookup table for muons determined from $J/\psi \rightarrow \mu^+\mu^-$ decays in data (left) and simulation (right) [124].

	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	Ratio
Full q^2	0.850 ± 0.006	0.798 ± 0.007	1.067 ± 0.008
$q^2 < 7 \text{ GeV}^2/c^4$	0.848 ± 0.006		1.062 ± 0.008
$q^2 > 7 \text{ GeV}^2/c^4$	0.854 ± 0.006		1.074 ± 0.008

Table 31: Averaged PID correction factors applied to Monte Carlo samples to correct for simulation and data differences.

lookup tables in MC and data. The corrected Monte Carlo yield is taken as the product sum of the correction weights of each B daughter track. Systematic uncertainties are evaluated by performing 1000 pseudo-experiments, each time varying the contents of the lookup tables within its quoted uncertainties. This is done separately for the different q^2 ranges considered. The corresponding PID correction ratios between the signal and normalisation channel are shown in Figure 75 as an example for the full q^2 range. Additional systematic uncertainties are due to the binning scheme in track momenta, pseudo rapidity and multiplicity variables in the PIDCalib lookup table. Varying the binning scheme leads to an additional systematic uncertainty of 0.4%. The final PID corrections averaged over all tracks and all events for each q^2 bin used in the analysis are given in Table 31.

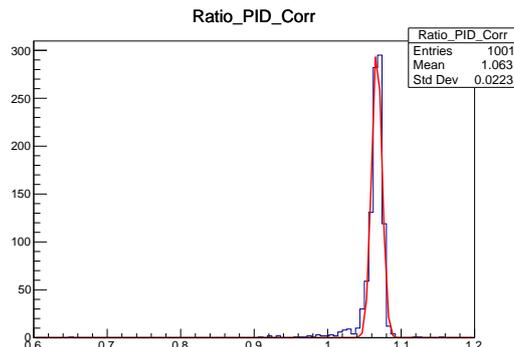


Figure 75: Spread of the PID correction ratio between the signal and normalisation channel by varying the efficiency corrections in the different bins of the lookup tables within their uncertainties. This spread is used as a systematic uncertainty related to the limited statistics of the calibration samples used to derive the correction. Here the PID correction for the full q^2 range is shown as an example. Figure taken from [9].

10.5 Tracking correction

In order to perform a branching fraction measurement it is necessary to determine the track reconstruction efficiency. This is measured using a data-driven tag-and-probe method from $J/\psi \rightarrow \mu^+ \mu^-$ decays, where the tag muon track is fully reconstructed and identified as a muon and the probe track is only partially reconstructed using only specific subdetectors to probe the remaining ones [126]. The tracking efficiency is defined to be the fraction of $J/\psi \rightarrow \mu^+ \mu^-$ where the probe track can

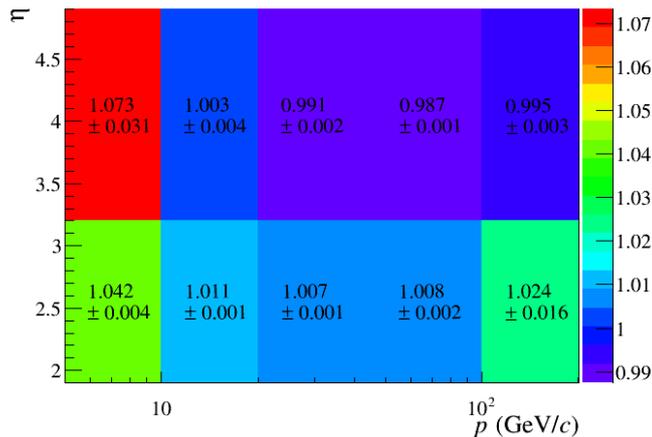


Figure 76: The tracking efficiency look-up table used to correct charged tracks, binned in momentum and pseudorapidity. Figure taken from [127].

be matched to a fully reconstructed track. The ratio of the resulting tracking efficiency between data and simulation is then used to weight the simulation sample according to pseudorapidity and momentum to obtain an efficiency correction in each (p_T, y) bin. Such a lookup table is provided by the collaboration and shown in Figure 76. Here the correction factors range from 0.987-1.073. These correction factors are applied as a weight on each track depending on its kinematic and the efficiencies are corrected by taking the product of the weights for each track. Before this weight can be applied the simulated sample needs to be weighted to agree in event multiplicity and other variables with the data. This is done using a BDT as explained in Section 7. As the signal channel $B_s^- \rightarrow K^- \mu^+ \nu_\mu$ is a two body decay and the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ is a four body decay, the corrections partially cancel when taking the ratio of the efficiencies.

The systematic uncertainties on the correction factor are determined by performing 1000 pseudo-experiments, each time varying the efficiencies in the lookup tables within their uncertainties. The weight per event is then obtained by multiplying the efficiency ratios for all signal tracks which is shown in Figure 77. The RMS of the distribution of the tracking correction ratio gives the statistical uncertainty, which is 0.1% for signal, 0.4% for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ and 0.4% for their ratio. In addition systematic uncertainties from hadronic interactions with the detector material have to be taken into account. They partially cancel in the ratio between the two decays, but due to the two extra tracks in the normalisation channel from $D_s^- \rightarrow K^+ K^- \pi^-$ there is an additional uncertainty for the kaon of 1.1% and the pion track of 1.4%. The sum of both gives an uncertainty of around 3% including also the uncertainty of the method. The tracking corrections to the efficiency calculations are summarised in Table 32 for each of the q^2 bins used in the fit.

10.6 Migration in between q^2 bins

The q^2 variable is reconstructed from the neutrino solution selected by the linear regression algorithm described in Section 5.2. Due to the non-negligible resolution

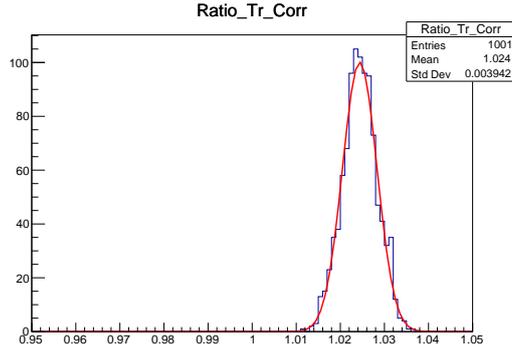


Figure 77: Distribution of tracking correction ratio from varying the individual efficiency corrections from the lookup table within their uncertainties. Figure taken from [9].

$K^- \mu^+ q^2$ Sel.	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	Ratio
No Sel.	0.999 ± 0.012	0.975 ± 0.035	1.025 ± 0.03
$q^2_{K^- \mu^+} < 7 \text{ GeV}^2/c^4$	0.999 ± 0.012		1.025 ± 0.03
$q^2_{K^- \mu^+} > 7 \text{ GeV}^2/c^4$	0.999 ± 0.012		1.025 ± 0.03

Table 32: Tracking efficiency corrections applied to Monte Carlo events.

of the reconstructed q^2 , the migration of events between the selected $q^2 > 7 \text{ GeV}^2/c^4$ and $q^2 < 7 \text{ GeV}^2/c^4$ regions have to be determined. This leads to rejected events which should have been selected with a truly reconstructed q^2 in the requested region but are reconstructed only outside (outward migration). Furthermore there are events outside the region of interest based on the true q^2 which are selected inside due to their resolution (inward migration). The migration of events is determined by comparing the true q^2 distribution against the reconstructed one as shown in Figure 78. Here the regions containing events migrating in and out are marked in green and pink, whereas the correctly selected events are marked in white for the low q^2 bin and in olive for the high q^2 bin. A correction factor is calculated from these simulated events by taking the ratio of true versus reconstructed events in the different q^2 regions. The resulting correction factors are:

$$\begin{aligned}
 \text{migration corr. } q^2 > 7 \text{ GeV}^2/c^4 &= 1.022 \pm 0.018 \\
 \text{migration corr. } q^2 < 7 \text{ GeV}^2/c^4 &= 0.990 \pm 0.017.
 \end{aligned}
 \tag{81}$$

These results are obtained from simulated $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ events after a full selection is applied. The systematic uncertainty from the q^2 migration is assigned as the uncertainty of the correction factor.

10.7 Final corrected relative efficiency

The calculated efficiencies and correction factors applied to determine the full corrected efficiency of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ as explained in the previous section are summarised in Table 33. The final corrected ratio of efficiencies for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ divided by $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ is given in Table 34.

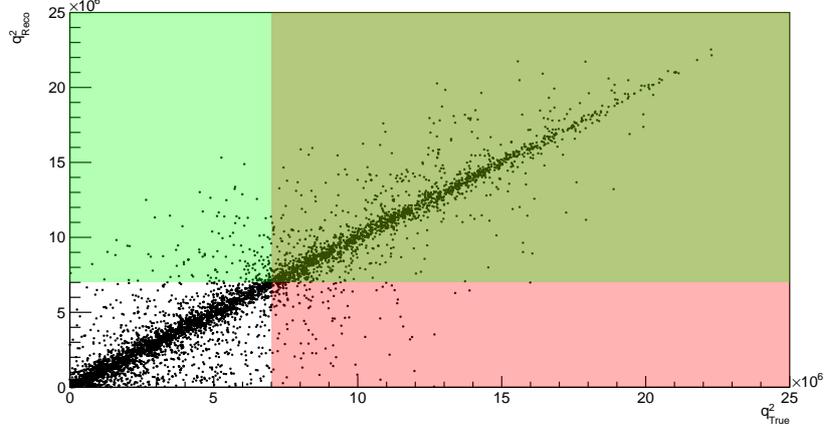


Figure 78: The reconstructed q^2 solution using the regression model plotted against the true q^2 for simulated $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ events. The different migration regions are marked in green and pink when selecting the q^2 region of interest above and below $7 \text{ GeV}^2/c^4$.

Source	Efficiency [%] $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$		
			$q^2 < 7 \text{ GeV}^2/c^4$	$q^2 > 7 \text{ GeV}^2/c^4$
Generator	17.87 ± 0.08	20.84 ± 0.08	19.67 ± 0.14	21.13 ± 0.09
Selection	0.7817 ± 0.02404	0.4162 ± 0.0138	0.7790 ± 0.0209	0.3242 ± 0.0121
Source	Correction factor			
Tracking	0.975 ± 0.035	0.999 ± 0.012	0.999 ± 0.012	0.999 ± 0.012
Trigger	0.661 ± 0.007	0.687 ± 0.007	0.692 ± 0.007	0.684 ± 0.009
PID	0.789 ± 0.007	0.850 ± 0.006	0.848 ± 0.006	0.854 ± 0.006
σ_{mcorr}		1.036 ± 0.005	1.036 ± 0.005	1.036 ± 0.005
Isolation	0.988 ± 0.002	0.993 ± 0.001	0.993 ± 0.001	0.993 ± 0.001
Charged BDT		0.989 ± 0.007	0.989 ± 0.007	0.989 ± 0.007
Same sign BDT		0.959 ± 0.011	0.959 ± 0.011	0.959 ± 0.011
q^2 migration			0.990 ± 0.017	1.022 ± 0.018
Total Corrected Efficiency [%]				
	0.07097 ± 0.0030	0.0498 ± 0.00021	0.0860 ± 0.0035	0.0395 ± 0.0018

Table 33: Summary of efficiencies and corrections entering into the combined efficiency for the signal and normalisation modes.

Source	Efficiency Ratio		
	Full q^2	$q^2 < 7\text{GeV}^2/c^4$	$q^2 > 7\text{GeV}^2/c^4$
Generator	1.166 ± 0.006843	1.101 ± 0.008975	1.182 ± 0.007311
Selection	0.5265 ± 0.02704	0.9876 ± 0.04052	0.4094 ± 0.02333
Source	Correction factor		
Tracking	1.025 ± 0.03	1.025 ± 0.03	1.025 ± 0.03
PID	1.067 ± 0.008	1.062 ± 0.008	1.074 ± 0.008
Trigger	1.04 ± 0.008	1.046 ± 0.009	1.033 ± 0.009
$\sigma_{m\text{corr}}$	1.036 ± 0.005	1.036 ± 0.005	1.036 ± 0.005
Isolation	1.005 ± 0.002	1.005 ± 0.002	1.005 ± 0.002
Charged BDT	0.989 ± 0.007	0.989 ± 0.007	0.989 ± 0.007
Same sign BDT	0.959 ± 0.011	0.959 ± 0.011	0.959 ± 0.011
q^2 migration		0.990 ± 0.017	1.022 ± 0.018
Total Corrected Efficiency Ratio			
	0.6897 ± 0.04315	1.2100 ± 0.07016	0.5556 ± 0.0390

Table 34: Summary of efficiency and correction ratios entering into the combined efficiency.

10.8 Systematic uncertainties

The uncertainties on the correction factors calculated before to correct for MC and data differences are taken as systematic uncertainties when calculating the final ratio of branching fractions.

The corrected efficiency for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ is plotted against the true corrected mass and true q^2 in Figure 79 which are built from truth matched quantities. One can see that the selection is q^2 dependent such that events at high q^2 are suppressed. This bias together with the lack of knowledge on the shape of true q^2 distribution leads to a large uncertainty on the efficiency. As described earlier in Section 2.5, the different form factor models predict a very different q^2 distribution. Their different distributions after all applied selection cuts are compared in Figure 80. One form factor prediction has to be chosen for the different q^2 bins. Here LCSR [52] prediction are used for the low q^2 bin as it is valid in that kinematic region. The latest Lattice QCD calculations from the MILC collaboration [51] are used in the high q^2 bin as well as for the full q^2 bin.

Since the generator and selection efficiency are determined from MC samples reweighted to the corresponding form factor models, there is a systematic uncertainty related to the finite MC statistic of these samples. The systematic uncertainty due to the generator level efficiency is small due to its large generated sample size. The selection efficiency is calculated before and after applying all selection cuts explained in detail in Section 6. Since some of them are very tight in order to reject as much background as possible the statistics of the sample is limited which leads to the largest systematic uncertainty.

The finite MC statistics also leads to an uncertainty in the template shape for the templates used in the signal and normalisation fit. By using the Beeston-Barlow light method as described in Section 8.2 for the maximum likelihood

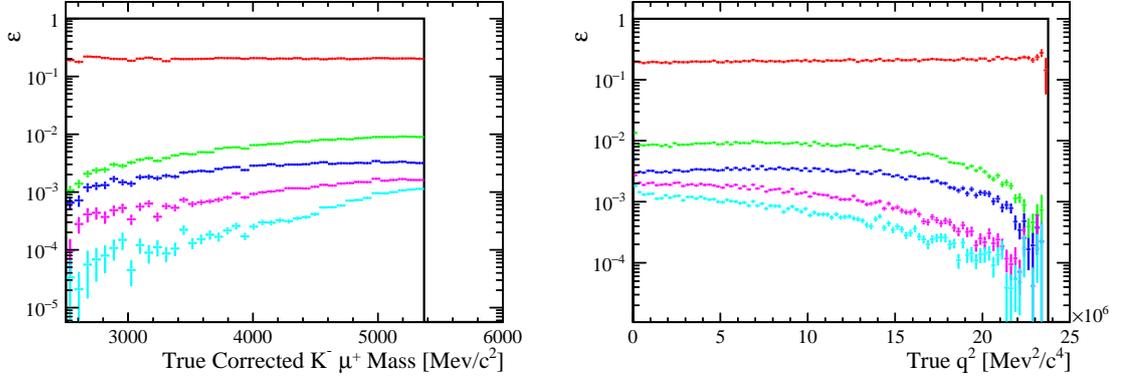


Figure 79: The corrected efficiencies for selections on $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ candidates are plotted against the true corrected mass (left) and against the true q^2 (right). The following sections cuts are applied successively: no selections (black), generator level cuts (red), stripping and trigger (green), corrected mass error (blue), Isolation and Charged BDT (magenta), SS BDT (light blue). Figure modified from [9].

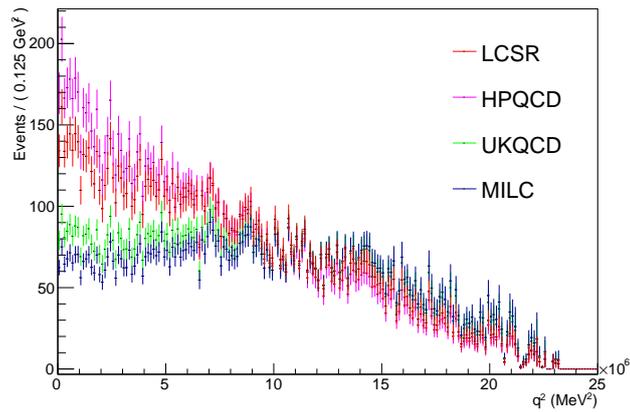


Figure 80: q^2 distribution of the different FF models.

template fits, the effect of this systematic uncertainty is already folded into the fit likelihood. Therefore the uncertainties on the yields presented in Tables 17 and 23 contain both the statistical uncertainty of the data sample and the systematic uncertainty due to the limited MC sample size. To separate out the latter component from the total fit uncertainty, both fits are repeated without the use of the Beeston-Barlow method. The quadrature difference is separated from the total fit uncertainty as the systematic uncertainty from the limited statistic of the fit templates, which is 3.5% for the full q^2 range and 4.3% for the different q^2 bins as a relative uncertainty for the ratio of branching fractions. By doing this the reported statistical uncertainty refers only to the statistical uncertainty of the data, not including the MC statistical uncertainty.

The D_s^+ of the normalisation channel is reconstructed via the $D_s^- \rightarrow K^- K^+ \pi^-$ decay mode. Therefore the external branching fraction is needed to normalise the decay $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. The latest PDG [12] gives:

$$\mathcal{B}(D_s^- \rightarrow K^- K^+ \pi^-) = 5.45 \pm 0.17\% \quad (82)$$

which gives a relative uncertainty of 3.1% for the ratio of branching fractions.

Systematic uncertainties on the signal fit come from variations of the corrected mass due to the different form factor models. This is tested by reweighting the default Isgur-Wise model with LQCD and LCSR predictions. The effect on the corrected mass is shown in Figure 81 for the different q^2 regions used in the analysis, which is only barely visible. Therefore this systematic uncertainty is negligible.

Table 35 summarizes the total systematic uncertainties for both signal and normalisation channel. Their relative systematic uncertainties associated to the branching fraction ratio are given in Table 36.

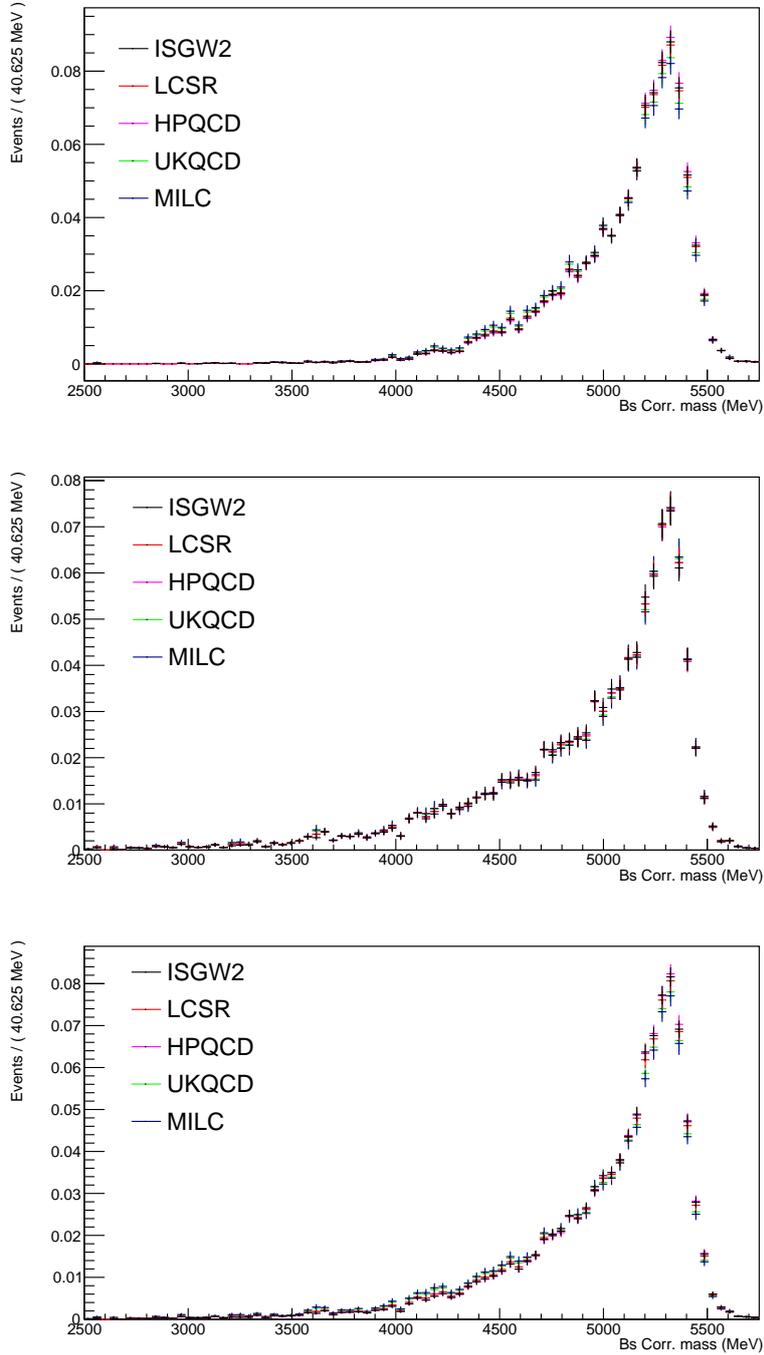


Figure 81: Reweighting the corrected mass with different FF predictions for different q^2 regions: $q^2 < 7$ (top), $q^2 > 7$ (middle), full q^2 (bottom)

Uncertainty [%]	$B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$		
		No q^2 sel.	$q^2 < 7$	$q^2 > 7$
Generator efficiency	0.5	0.4	0.7	0.4
Selection efficiency	3.1	3.3	2.7	3.7
Tracking	3.5	1.2	1.2	1.2
Trigger	0.7	0.7	0.7	0.9
PID	0.7	0.6	0.6	0.6
m_{corr} error		0.5	0.5	0.5
Isolation	0.2	0.1	0.1	0.1
Charged BDT		0.7	0.7	0.7
Same Sign BDT		1.1	1.1	1.1
q^2 migration			1.7	1.8
Fit MC statistics	2.3	2.5	3.6	3.6
$\mathcal{B}(D_s^- \rightarrow K^- K^+ \pi^-)$	3.1			

Table 35: Summary of systematic uncertainties for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$.

Variable	Relative uncertainty		
	No q^2 sel.	$q^2 < 7$	$q^2 > 7$
Selection efficiency	5.1	4.1	5.7
Fit MC statistics	3.5	4.3	4.3
$\mathcal{B}(D_s^- \rightarrow K^- K^+ \pi^-)$	3.1	3.1	3.1
Tracking	3.0	3.0	3.0
q^2 migration		1.7	1.8
Same Sign BDT	1.1	1.1	1.1
Trigger	0.8	0.9	0.9
Particle Identification	0.8	0.8	0.8
Charged BDT	0.7	0.7	0.7
Generator efficiency	0.6	0.8	0.6
m_{corr} error	0.5	0.5	0.5
Isolation	0.2	0.2	0.2
Total	7.8	7.9	8.8

Table 36: Summary of systematic uncertainties associated with the branching fraction ratio of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$.

11 Results

This chapter combines all previously obtained quantities, such as the signal and normalisation yields determined in Sections 8, 9 and the relative efficiency calculated in Section 10, and determines the ratio of branching fractions of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ with respect to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. This branching fraction ratio is then used together with theoretical predictions to extract the ratio of CKM matrix elements $|V_{ub}|/|V_{cb}|$ from these decays. In addition the total $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ branching fraction is determined. The obtained results are compared to already existing measurements at the end of the section and discussed.

11.1 Measurement of the branching fraction ratio

The ratio of branching fractions for the signal and normalisation channel, $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ respectively, is determined from the ratio of yields at production which is obtained by correcting the fitted yields by their efficiencies. Since the D_s^- of the normalisation channel is reconstructed through the decay channel $D_s^- \rightarrow K^+ K^- \pi^-$ this branching fraction has also to be taken into account. The branching fraction ratio is calculated from the following formula

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} = \frac{N(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{N(B_s^0 \rightarrow D_s^- (\rightarrow K^+ K^- \pi^-) \mu^+ \nu_\mu)} \times \frac{\epsilon(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{\epsilon(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)} \times \mathcal{B}(D_s^- \rightarrow K^+ K^- \pi^-). \quad (83)$$

The input values are summarised in Table 37. The fitted normalisation and signal yields come from Section 8 and 9, respectively, where the given uncertainty corresponds to the statistical uncertainty of the data as the MC statistic uncertainty is taken into account as a systematic uncertainty. The relative efficiency is determined in the previous section. Together with the external branching fraction taken from the PDG [12] they can be used to calculate the branching fraction ratio.

This results in the ratios of branching fractions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decays in the different q^2 bins considered:

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \Big|_{\text{full } q^2} = (6.0791 \pm 0.2006(\text{stat}) \pm 0.4759(\text{syst})) \times 10^{-3} \quad (84)$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \Big|_{\text{low } q^2} = (1.7163 \pm 0.0714(\text{stat}) \pm 0.1350(\text{syst})) \times 10^{-3} \quad (85)$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \Big|_{\text{high } q^2} = (3.8086 \pm 0.1426(\text{stat}) \pm 0.2903(\text{syst})) \times 10^{-3} \quad (86)$$

where the first uncertainty is statistical coming from the size of the fitted data sample and the second uncertainty is systematic. The uncertainty on the branching fraction ratio is dominated by the systematic uncertainty, which is twice as big as the statistical uncertainty.

The various systematic uncertainties entering into these ratios were summarised in

Input Variable	Value	Reference
$N(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)$	14693.8 ± 353.4	Table 23
low q^2	7277.9 ± 252.0	
high q^2	7415.9 ± 253.8	
$N(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)$	191000 ± 4393	Table 17
$\epsilon(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)/\epsilon(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)$	0.6897 ± 0.04315	Table 34
low q^2	1.2100 ± 0.07016	
high q^2	0.5556 ± 0.0390	
$\mathcal{B}(D_s^- \rightarrow K^+ K^- \pi^-)$	0.0545 ± 0.0017	[12]

Table 37: Summary of quantities entering the calculation of the ratio of branching fractions.

the previous section in Table 36. The largest systematic uncertainty comes from the selection efficiency which is calculated from MC samples and is therefore related to the statistics of the MC sample. The second largest systematic uncertainty is also related to the MC statistics, but this time due to the template statistical uncertainty used in the signal and normalisation fit. Both of them can be reduced by generating more MC samples. Especially for the signal fit a larger statistic of the $|V_{cb}|$ template would help to reduce the uncertainty as this component has the largest uncertainty in the template shape coming from the small MC statistics as shown in Figure 63 of the Section 9.5. After all selection cuts only 10k events left in the full q^2 range whereas the fit pushes this component up by a factor of 10 requiring around 100k events.

The branching fraction $\mathcal{B}(D_s^- \rightarrow K^+ K^- \pi^-)$ taken from the PDG [12] is the third largest systematic uncertainty, but it is only marginally bigger than the uncertainty from the tracking. The latter one is irreducible since it is dominated by the uncertainty of the material interaction of the two additional tracks of the normalisation decay. Systematic uncertainties due to the q^2 migration can be improved by a more precise form factor prediction for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. A more realistic simulation of the same sign, charge BDT response as well as isolation and corrected mass variables could help to improve their corresponding systematic uncertainty. Uncertainties due to the trigger and PID can be improved due to a higher statistics of their control modes in data.

11.2 Measurement of CKM elements $|V_{ub}|/|V_{cb}|$

The ratio of CKM matrix elements $|V_{ub}|/|V_{cb}|$ can be determined from the branching fraction ratio calculated in the previous section using the formula:

$$\frac{|V_{ub}|}{|V_{cb}|} = \left(\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \times R_{FF} \right)^{\frac{1}{2}} \quad (87)$$

where R_{FF} is a ratio determined from the form factor (FF) models described in more detail in Section 2.4 defined as

$$R_{FF} = \frac{\int_0^{q_{max}^2} \frac{d\Gamma(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{dq^2} / |V_{cb}|^2 dq^2}{\int_0^{q_{max}^2} \frac{d\Gamma(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{dq^2} / |V_{ub}|^2 dq^2}. \quad (88)$$

The corresponding predicted decay widths were discussed in Table 3 of Section 2.5 for the different available FF models. Using the prediction from the MILC collaboration [51] for the integrated $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay width and the latest $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ FF predictions from the HPQCD collaboration [47] leads to the following FF ratio:

$$R_{FF} = \frac{9.155 \pm 0373}{4.263 \pm 0.922} = 2.147 \pm 0.473. \quad (89)$$

Using this ratio of form factors together with the branching fraction ratio for the total q^2 range from Equation 84 gives

$$\frac{|V_{ub}|}{|V_{cb}|} = 0.11425 \pm 0.00485(\text{exp.}) \pm 0.01257(\text{theo.}) \quad (90)$$

where the first uncertainty is the combined experimental uncertainty from the statistical and systematic uncertainties on the branching fraction ratio given in Equation 84 and the second uncertainty is theoretical from the uncertainties on the form factor predictions.

A comparison of this new result with the existing results of inclusive and exclusive $|V_{ub}|$ and $|V_{cb}|$ results from HFLAV [60] as well as the recent $|V_{ub}|/|V_{cb}|$ determination from Λ_b -decays [5] by LHCb is shown in Figure 82. The world average indicated in red is a combined fit done by HFLAV to the exclusive $|V_{ub}|$ and $|V_{cb}|$ measurements coming from $\bar{B} \rightarrow \pi l \bar{\nu}_\ell$, $\bar{B} \rightarrow D^* l \bar{\nu}_\ell$ and $\bar{B} \rightarrow D l \bar{\nu}_\ell$ decays respectively, as well as the constraint from Λ_b decays, excluding the measurement performed in this thesis. The fit gives a value of $P(\chi^2) = 7.7\%$, the dashed ellipse around it corresponds to a 1σ 2-dimensional contour of 68% confidence level. The inclusive $|V_{ub}|$ and $|V_{cb}|$ values are shown as a point with error bars, they are extracted from different calculations where the GGOU calculation is used for $|V_{ub}|$ and the kinematic scheme for $|V_{cb}|$. Details for these can be found in the recent HFLAV publication [60]. The light-blue band corresponds to the result of $|V_{ub}|/|V_{cb}|$ determined from $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ with respect to $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decays in this thesis. The bands displayed in the Figure correspond to $\Delta\chi^2 = 1$ contours of the different values.

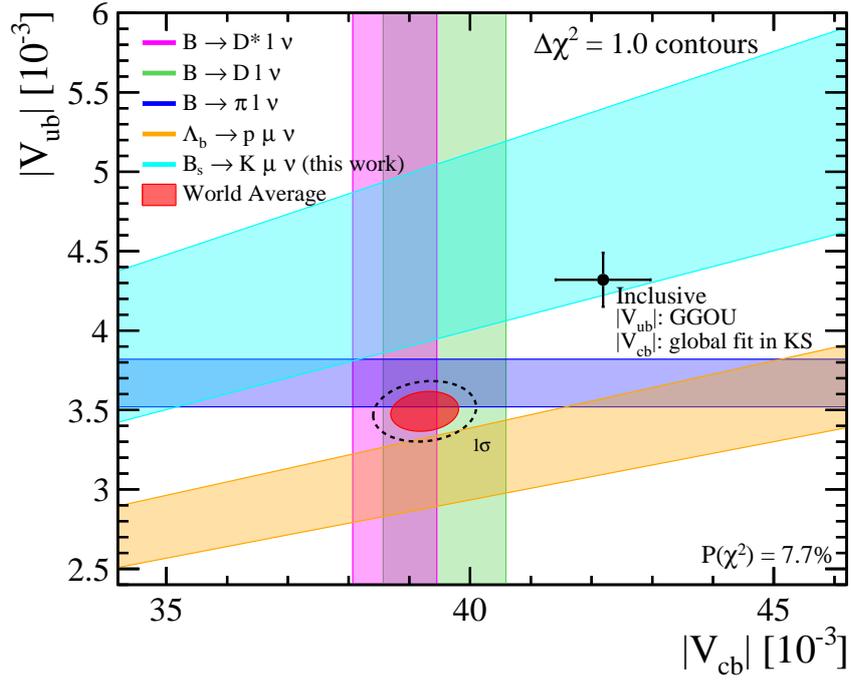


Figure 82: Comparison of exclusive and inclusive $|V_{ub}|$ and $|V_{cb}|$ measurements. Exclusive $|V_{ub}|$ comes from $\bar{B} \rightarrow \pi l \bar{\nu}_\ell$, exclusive $|V_{cb}|$ from $\bar{B} \rightarrow D l \bar{\nu}_\ell$ and $\bar{B} \rightarrow D^* l \bar{\nu}_\ell$ decays and $|V_{ub}|/|V_{cb}|$ from Λ_b decays. Those measurements are combined in a fit which is displayed as the world average in red. The point with error bars gives the inclusive $|V_{ub}|$ and $|V_{cb}|$ values from different calculations. The light blue band shows the result of this thesis. This plot was modified from the official HFLAV publication [60] to include the result of this thesis.

11.3 Extraction of $\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)$

The total $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ branching fraction can be determined from the measured ratio of branching fractions from Equation 84 according to the formula

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = & \tau_{B_s^0} \times \frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \times |V_{cb}|^2 \\ & \times \int_0^{q_{max}^2} \frac{d\Gamma(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{dq^2} / |V_{cb}|^2 dq^2, \end{aligned} \quad (91)$$

where $\tau_{B_s^0} = 1.509 \pm 0.004$ ps is the world average of the B_s^0 lifetime taken from the Heavy Flavor Averaging Group (HFLAV) [60]. For the $|V_{cb}|$ value either the inclusive or the exclusive average from the 2019 update of the PDG [12] is used, given as

$$|V_{cb}|_{incl.} = (42.2 \pm 0.8) \times 10^{-3}, \quad |V_{cb}|_{excl.} = (39.5 \pm 0.9) \times 10^{-3}. \quad (92)$$

The total decay width for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ was calculated in Section 2.5.3 for different form factor predictions. Using the latest lattice QCD prediction from the HPQCD collaboration [47] which is valid for the full q^2 range gives

$$\int_0^{q_{max}^2} \frac{d\Gamma(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)}{dq^2} / |V_{cb}|^2 dq^2 = 9.155 \pm 0.373 \text{ ps}^{-1}. \quad (93)$$

From that the branching fraction of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ using the inclusive $|V_{cb}|$ is determined as

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = & (1.496 \pm 0.049(\text{stat.}) \pm 0.117(\text{syst.}) \pm 0.061(\text{theo.}) \\ & \pm 0.057(|V_{cb}|_{incl}) \pm 0.004(\tau_{B_s^0})) \times 10^{-4}, \end{aligned} \quad (94)$$

using instead the exclusive $|V_{cb}|$ value leads to

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = & (1.310 \pm 0.043(\text{stat.}) \pm 0.103(\text{syst.}) \pm 0.053(\text{theo.}) \\ & \pm 0.060(|V_{cb}|_{excl}) \pm 0.003(\tau_{B_s^0})) \times 10^{-4}, \end{aligned} \quad (95)$$

where the first uncertainty is statistical, the second uncertainty is systematic, the third arises from the theoretical form factor predictions related to the decay width, the fourth is from the uncertainty from the used $|V_{cb}|$ value and the fifth from the B_s^0 lifetime.

Those two branching fractions are compatible with each other within 1.2σ . They can be compared to the predicted $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ branching fractions from the different form factor models given in Table 3 of Section 2.5 which were calculated using the exclusive $|V_{cb}|$ value. The measured branching fractions are in good agreement with the predictions from the HPQCD [49] collaboration giving $(1.60 \pm 0.32) \times 10^{-4}$. The predictions from the UKQCD [50] and MILC [51] collaboration are slightly lower at $(0.94 \pm 0.27) \times 10^{-5}$ and $(0.88 \pm 0.20) \times 10^{-5}$, respectively, which are in agreement at 1.4σ and 2.2σ with the extracted total branching fraction. The predicted $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ branching fractions from LCSR [52] is predicted to be much higher as $(2.29 \pm 0.24) \times 10^{-4}$ which differs by 4σ from the measured branching fraction.

11.4 Implications of the measurement

In order to put the measurement of $|V_{ub}|/|V_{cb}|$ from this thesis into context, it has to be compared to previous measurements. The ratio of partial branching fractions at high q^2 from $\Lambda_b \rightarrow p\mu^-\nu_\mu$ and $\Lambda_b \rightarrow \Lambda_c^+\mu^-\nu_\mu$ decays [5] was measured by LHCb in 2015. Using the recent average of $\mathcal{B}(\Lambda_c^+ \rightarrow pK^-\pi^+) = 6.28 \pm 0.32$ [12] and the LQCD predictions of the form factors [128] leads to the ratio of CKM matrix elements calculated by HFLAV [60]:

$$|V_{ub}|/|V_{cb}|_{\Lambda_b} = 0.079 \pm 0.004(\text{exp.}) \pm 0.004(\text{theo.}) \quad (96)$$

The PDG [12] also quotes a ratio of $|V_{ub}|/|V_{cb}|$ which is evaluated separately for inclusive and exclusive B -decays.

$$|V_{ub}|/|V_{cb}| = 0.101 \pm 0.007 \quad (\text{inclusive}), \quad (97)$$

$$|V_{ub}|/|V_{cb}| = 0.094 \pm 0.005 \quad (\text{exclusive}). \quad (98)$$

These ratios are an average of different measurements. For exclusive $|V_{cb}|$ measurements from $\bar{B} \rightarrow D^*l\bar{\nu}_\ell$ and $\bar{B} \rightarrow Dl\bar{\nu}_\ell$ decays are combined with theoretical predictions, whereas for exclusive $|V_{ub}|$ various $\bar{B} \rightarrow \pi l\bar{\nu}_\ell$ measurements from BaBar and Belle are combined with form factor predictions from Lattice QCD.

The result for $|V_{ub}|/|V_{cb}|$ obtained in this thesis from semileptonic B_s^0 decays together with form factor predictions from the MILC collaboration [51] for $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ and from the HPQCD collaboration [47] for $B_s^0 \rightarrow D_s^-\mu^+\nu_\mu$ gives

$$\frac{|V_{ub}|}{|V_{cb}|} = 0.114 \pm 0.005(\text{exp.}) \pm 0.013(\text{theo.}) \quad (99)$$

This result is significantly larger than the previous LHCb result using Λ_b decays [5] and compatible with it only at 2.5σ . The result of this thesis has much larger uncertainties from the form factor predictions since the whole available q^2 range was used instead of the only the very limited q^2 range used in the Λ_b analysis. The result of this thesis agrees within 1.4σ with the exclusive $|V_{ub}|/|V_{cb}|$ average from the PDG, but it is slightly higher than that. The result is in good agreement with the inclusive $|V_{ub}|/|V_{cb}|$ PDG average at the level of 0.9σ .

The $|V_{ub}|/|V_{cb}|$ result obtained in this thesis from B_s^0 decays are in agreement with previous measurements based on B^+ and B^0 decays and gives an additional constraint to the length of the unitarity triangle opposite to the angle β .

11.5 Discussion and outlook

The $|V_{ub}|/|V_{cb}|$ measurement from semileptonic B_s^0 decays presented in this thesis is limited by theoretical uncertainties which are much larger than the experimental ones. For lattice calculations the uncertainties are dominated by the chiral continuum extrapolation, light and heavy-quark discretization, and renormalization [51] which are planned to be further reduced by improvements in lattice techniques. Also further improvements on LCSR are planned such as higher loop corrections

which will reduce the theoretical uncertainties further [52]. In addition the form factor predictions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ are in tension at low q^2 with as much as around 2.3σ at $q^2 = 0$ as estimated by the MILC collaboration [51]. This tension is still up for debate and being discussed in the theory community, it might come from a different z-expansion between the collaborations. This assumption could be verified by updated form factor predictions from the UKQCD collaboration as well as from the MILC collaboration using different ensembles, which are expected very soon. Also an updated LQCD average by the Flavour Lattice Averaging Group (FLAG) would help to understand this tension, since by the time of their last average in 2019 [129] the latest results from the MILC collaboration [51] were not yet included.

Experimentally the uncertainty on $|V_{ub}|/|V_{cb}|$ is dominated by systematic uncertainties coming from the limited MC statistic used to calculate the selection efficiency as well as the template MC statistic used in the signal fit, which is very small for the large inclusive $b \rightarrow c$ background component. Both of these components can be reduced by generating more MC samples.

In addition for the signal fit there is a large uncertainty and correlation between the B2CC and higher excited kaon components due to their similar shape in corrected mass. The latter component could be further constrained by theoretical predictions of the $B_s^0 \rightarrow K^*$ and higher excited kaon resonance form factors and decay rates. Those could be compared to the fixed fractions between the different K^* components as used in the signal fit given in Equation 73. Another approach was followed in the official analysis note which is in internal working group review of the semileptonic group within the collaboration at the moment. Here the B2CC yield is fixed to the yield obtained from a method to reconstruct the B^+ peak from $J/\psi K^+$ which is similar to the q^2 reconstruction described in Section 5.2, but it has large uncertainties from choosing the correct solution as well as from correcting for the inefficiency of the method and extrapolating it to the full B2CC yield. Nevertheless leads this approach to very similar signal yields.

This analysis was performed on $2fb^{-1}$ of data collected in 2012 by LHCb which represents only a small fraction of the total data collected so far of about $9fb^{-1}$. Therefore the measurement could be extended to measure the branching fraction ratio in finer bins of q^2 . This could further constrain the q^2 dependence of the form factors or even fully measure the form factors in data. The latter would be challenging due to the limited resolution of the reconstructed q^2 , as discussed in Section 5.2 and would require a complicated unfolding. In addition for this kind of measurement a magnitude higher amount of simulated MC would be needed, in order to not be systematically limited by MC statistics as the current measurement and to correctly model the different kind of backgrounds contributing to the measurement. The generation of simulated data already presents a challenge today due to the large computing resources needed as well as the disk space in order to store these events. An order of magnitude more MC needs new developments such as ReDecay [130], RICHless [131] or tracker-only [132] simulations. ReDecay splits up the events in two groups: one containing the signal process and the other

containing the rest of the event, the latter one is then re-used multiple times whereas the signal process is regenerated every time. RICHless MC generation inherits the idea to reconstruct events without to use of the two RICH detectors which are mainly used for hadronic particle identification and take up about 30% of the simulation time. Another approach is to use the tracker-only simulation which goes one step further and does not simulate the RICH, calorimeter and muon system response. Those developments speed-up the MC simulation by large factors (between 8-20) and thus result in a reduced computation time. The latter two approaches even reduce the event size and such the disk space needed to store those. Those techniques will be of crucial importance for possible updates and extensions of this analysis but also in general for the larger data samples which will be collected in Run 3 of the LHC and beyond by the upgraded LHCb detector.

12 Conclusions

This thesis reports the measurement of the CKM matrix elements $|V_{ub}|/|V_{cb}|$ from semileptonic B_s^0 -meson decays using $2fb^{-1}$ of pp data collected in 2012 at $\sqrt{s} = 8$ TeV by the LHCb experiment. The B_s^0 -meson is reconstructed through $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decays and their relative branching fraction is measured. The measurement is performed in two bins of the four-momentum squared of the muon and neutrino q^2 , as well as for the full kinematically allowed q^2 range. The ratio of branching fraction are measured to be

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \Big|_{\text{full } q^2} = (6.0791 \pm 0.2006(\text{stat}) \pm 0.4759(\text{syst})) \times 10^{-3}$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \Big|_{\text{low } q^2} = (1.7163 \pm 0.0714(\text{stat}) \pm 0.1350(\text{syst})) \times 10^{-3}$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu)}{\mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu)} \Big|_{\text{high } q^2} = (3.8086 \pm 0.1426(\text{stat}) \pm 0.2903(\text{syst})) \times 10^{-3},$$

where the first error is statistical and the second systematic. The uncertainty on these measurements is dominated by the systematic uncertainty, where the biggest uncertainty systematic comes from the limited statistic of the Monte Carlo samples. The ratio of CKM matrix elements is extracted from this experimentally measured branching fraction ratio together with theoretical form factor predictions. Using Lattice QCD predictions from the MILC [51] for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and from the HPQCD [47] collaborations for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$, leads to the following ratio of CKM matrix elements

$$|V_{ub}|/|V_{cb}| = 0.114 \pm 0.005(\text{exp.}) \pm 0.013(\text{theo.}),$$

where the first error is the combined experimental uncertainty from the branching fraction ratio and the second one is the theoretical uncertainty from form factor predictions. This result is higher than the first LHCb result using Λ_b decays but it also has much larger theoretical uncertainties. It is also bigger than the exclusive $|V_{ub}|/|V_{cb}|$ average from the PDG, but compatible with it at 1.4σ . The result is in good agreement with the inclusive PDG average.

From the experimental measurement of the branching fraction ratio the total branching fraction of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ can be determined for the first time to be

$$\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = (1.31 \pm 0.14) \times 10^{-4}$$

where the uncertainty is the combined one from the experimental and theoretical uncertainties as well from external inputs. The uncertainty is dominated by the experimental systematic uncertainty.

Experimentally the different measurements performed in this thesis are dominated by their systematic uncertainty, where the largest source is coming from the

limited Monte Carlo statistics used to calculate the efficiencies as well as for the simulated templates used in the signal fit. These limitations can be overcome by using recent developments of fast simulation techniques for Monte Carlo samples in the future.

In general these measurements presented are limited by their theoretical uncertainties coming from the form factor predictions for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. Currently the predicted differential decay rate differs especially at low q^2 and the total integrated decay rate varies by a factor of 2.6 dependent on the used form factor prediction. This also means that the presented result of $|V_{ub}|/|V_{cb}|$ is very dependent on the theoretical predictions used to obtain it, whereas the branching fraction ratio is much less dependent on theory predictions. Therefore the measured branching fraction ratio provides an important constraint for the theory community as well as the extracted total branching fraction of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. In addition several improvements are planned from Lattice QCD predictions for the form factors in the near future which are expected to give percent level precision at the high q^2 region and might also clarify the observed difference at low q^2 . New calculations from the UKQCD and MILC collaboration are in progress and are therefore eagerly awaited.

Appendices

collaboration [ref.]	ensemble	heavy b-quarks	valence s-quark
HPQCD [49]	$N_f = (2 + 1)$ MILC-asqtad	NRQCD b quarks	HISQ action
RBC-UKQCD [50]	$N_f = (2 + 1)$ DWF	RHQ action for b quark	DWF light
Fermilab/MILC [51]	$N_f = (2 + 1)$ MILC-asqtad	Fermilab b quark	asqtad light

Table 38: Lattice simulation details of the papers providing form factor results for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$.

A Form factor comparisons

This Section gives a detailed summary of the used $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ form factor calculations. It lists the parameter values from the papers as well as their correlation matrix. Also plots from the published papers are compared to those generated by the analysis software using these results to check that they are correctly reproduced.

A.1 $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ publications

For the signal channel $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ 4 different form factor calculation from References [49–52] are used. The first three are from Lattice QCD whereas the last one comes from Light-Cone sum rule predictions. The Lattice calculations differ in several ways from each other which will be briefly compared. References [50] and [49] use the $B_s^0 K$ threshold for the z-parametrisation whereas Reference [51] uses the $B\pi$ threshold. Therefore the z-parameter is different and their form factor dependence on z can not be directly compared. In addition Reference [49] uses a modified z-expansion, where the chiral continuum extrapolation is combined with the z expansion into one fit function. References [50] and [51] use the same procedure by first doing a chiral-continuum extrapolation at high q^2 , followed by a z-expansion extrapolation to $q^2 = 0$. Also the setup for the simulations is different and is summarised in Table 38. Here DWF stands for domain wall fermions, RHQ for relativistic heavy quark, NRQCD for non-relativistic QCD and HISQ for an improved light quark action (Highly Improved Staggered Quark action). More details can be found in the latest FLAG report [53].

A.1.1 HPQCD parameters

Table 39 gives the z-expansion fit parameters for the BCL parametrisation from the HPQCD collaboration. They use $K=3$ for the expansion which leads to 3 coefficients in $b_i^{(0)}$ and $b_i^{(+)}$. Figure 83– 87 show plots from their paper along with the reproduced plots using the coefficients from Table 39 and generated with the analysis software.

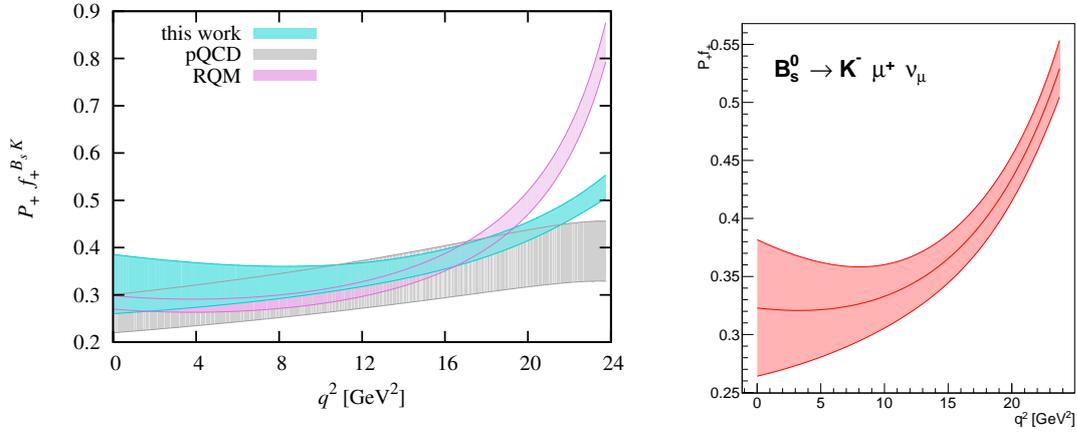


Figure 83: Form factors from HPQCD plotted against z . Left Figure taken from [49] and on the right generated using fit parameters taken from this paper. The light blue band on the left should be compared to the generated distribution on the right. Figure modified from [9].

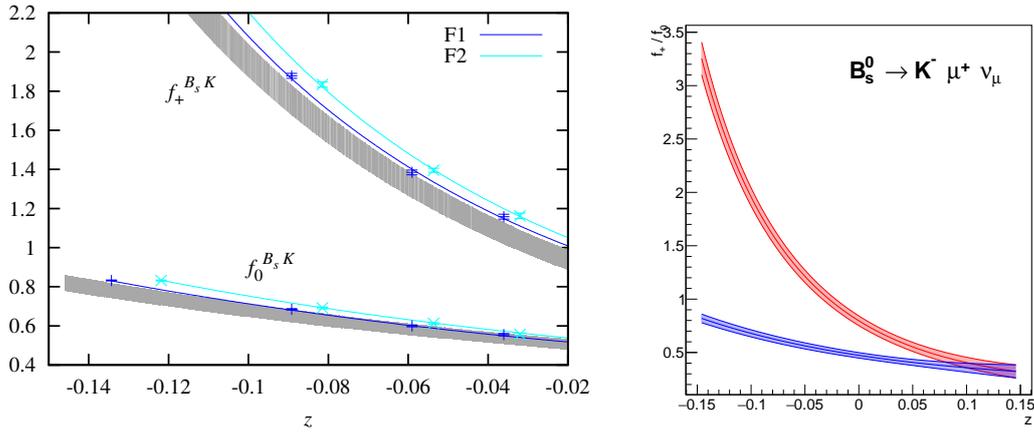


Figure 84: Form factors from HPQCD plotted against z . Left Figure taken from [49] and on the right generated using fit parameters taken from this paper. On the right Figure the red band corresponds to the vector form factor f_+ and the blue to the scalar form factor f_0 . Figure modified from [9].

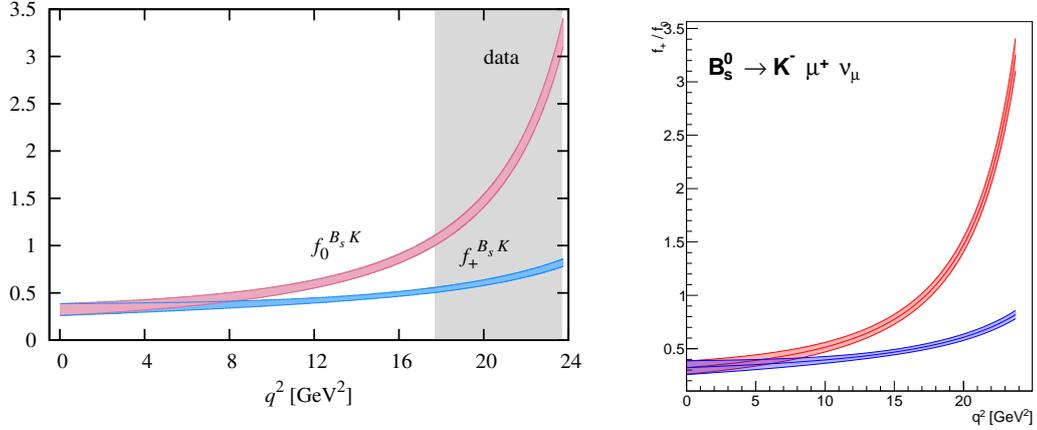


Figure 85: Form factors from HPQCD plotted against q^2 . Left Figure taken from [49] and on the right generated using fit parameters taken from this paper. On the right Figure the red band corresponds to the scalar form factor f_0 and the blue to the vector form factor f_+ . Figure modified from [9].

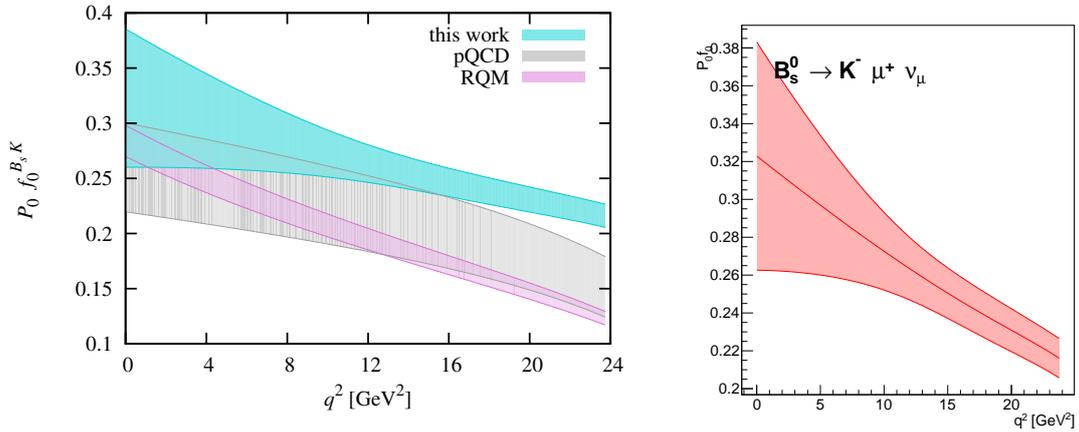


Figure 86: Form factors from HPQCD plotted against q^2 . Left Figure taken from [49] and on the right generated using fit parameters taken from this paper. The light blue band on the left should be compared to the generated distribution on the right. Figure modified from [9].

	$b_1^{(0)}$	$b_2^{(0)}$	$b_3^{(0)}$	$b_1^{(+)}$	$b_2^{(+)}$	$b_3^{(+)}$
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 39: Coefficients and corresponding errors of the z -expansion for the BCL parametrisation of the HPQCD form factors with the associated covariance matrix. Parameters taken from [49].

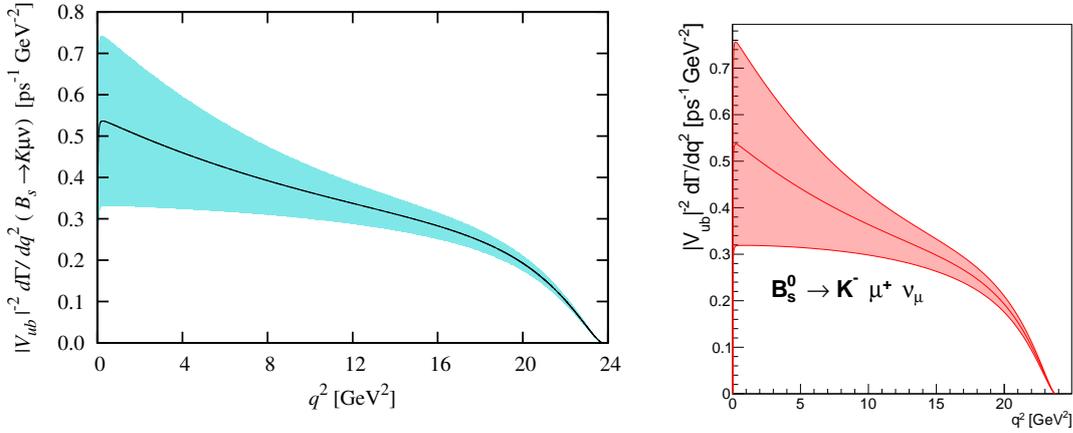


Figure 87: The differential $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay rate from HPQCD plotted against q^2 . Left Figure taken from [49] and on the right generated using fit parameters taken from this paper. Figure modified from [9].

A.1.2 RBC-UKQCD parameters

Table 40 gives the z -expansion parameters fit parameters for the BCL parametrisation from the RBC-UKQCD collaboration. They also use $K=3$ for the expansion which leads to 3 coefficients in $b_i^{(0)}$ and $b_i^{(+)}$. Figure 88- 90 show plots from their paper along with the reproduced plots using the coefficients from Table 40 and generated with the analysis software.

	$b_{(+)}^0$	$b_{(+)}^1$	$b_{(+)}^2$	$b_{(0)}^0$	$b_{(0)}^1$	$b_{(0)}^2$
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 40: Central values, errors, and correlation matrix for the BCL z -parametrisations of form factors from RBC-UKQCD. Values taken from [50].

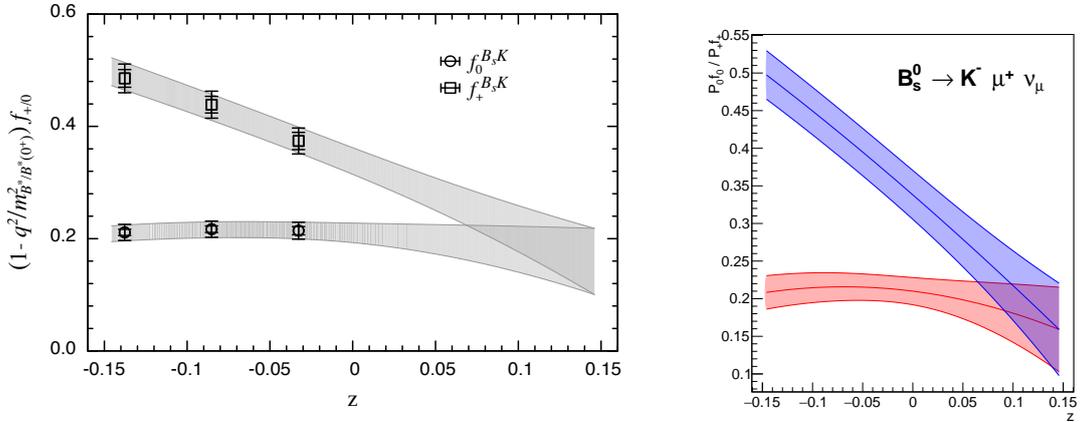


Figure 88: Form factors from RBC-UKQCD plotted against z . Left Figure taken from [50] and on the right generated using fit parameters taken from this paper. On the right Figure the red band corresponds to the scalar form factor f_0 and the blue to the vector form factor f_+ . Figure modified from [9].

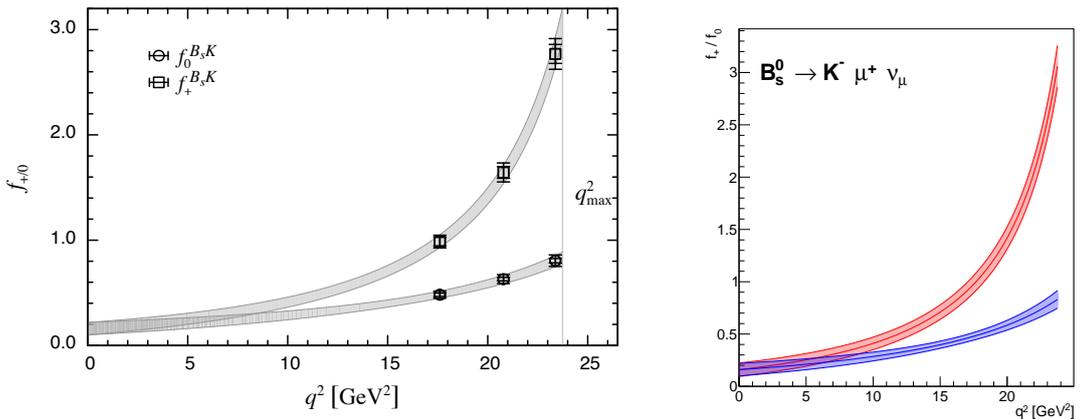


Figure 89: Form factors from RBC-UKQCD plotted against q^2 . Left Figure taken from [50] and on the right generated using fit parameters taken from this paper. On the right Figure the blue band corresponds to the scalar form factor f_0 and the red to the vector form factor f_+ . Figure modified from [9].

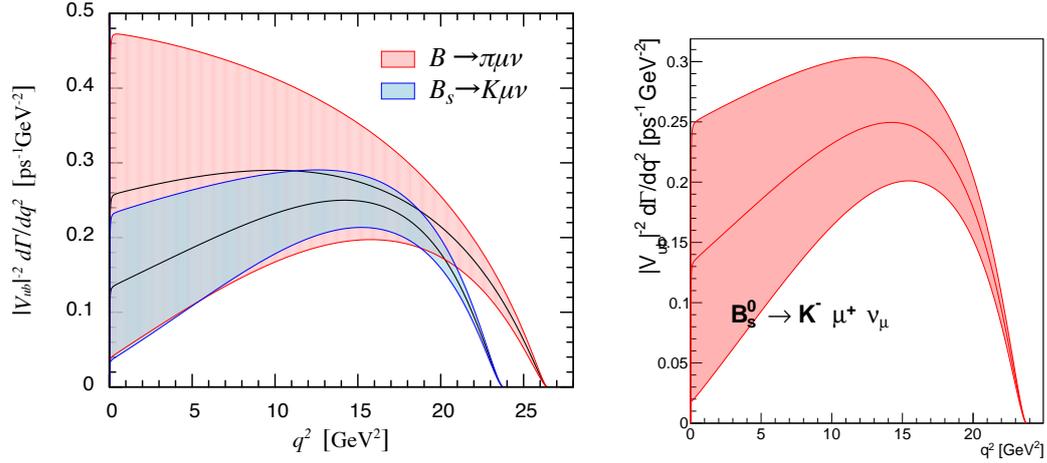


Figure 90: The differential $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay rate from RBC-UKQCD plotted against q^2 . Left Figure taken from [50] and on the right generated using fit parameters taken from this paper. The dark blue band in the left Figure should be compared to the distribution on the right. Figure modified from [9].

A.1.3 Fermilab/MILC parameters

Table 41 gives the z-expansion fit parameters for the BCL parametrisation from the Fermilab/MILC collaboration. They use $K=4$ for the expansion which leads to 4 coefficients in $b_i^{(0)}$ and $b_i^{(+)}$ for the z-parametrisation of the form factors. Figure 91 and 92 show plots from their paper along with the reproduced plots using the coefficients from Table 41 and generated with the analysis software.

	$b_{(+)}^0$	$b_{(+)}^1$	$b_{(+)}^2$	$b_{(+)}^3$	$b_{(0)}^0$	$b_{(0)}^1$	$b_{(0)}^2$	$b_{(0)}^3$
Value	0.3623	-0.9559	-0.8525	0.2785	0.1981	-0.1661	-0.6430	-0.3754
Error	0.0178	0.1307	0.4783	0.6892	0.0101	0.1130	0.4385	0.4535
$b_{(+)}^0$	1.0000	0.6023	0.0326	-0.1288	0.7122	0.6035	0.5659	0.5516
$b_{(+)}^1$	0.6023	1.0000	0.4735	0.2677	0.7518	0.9086	0.9009	0.8903
$b_{(+)}^2$	0.0326	0.4735	1.0000	0.9187	0.5833	0.7367	0.7340	0.7005
$b_{(+)}^3$	-0.1288	0.2677	0.9187	1.0000	0.4355	0.5553	0.5633	0.5461
$b_{(0)}^0$	0.7122	0.7518	0.5833	0.4355	1.0000	0.8667	0.7742	0.7337
$b_{(0)}^1$	0.6035	0.9086	0.7367	0.5553	0.8667	1.0000	0.9687	0.9359
$b_{(0)}^2$	0.5659	0.9009	0.7340	0.5633	0.7742	0.9687	1.0000	0.9899
$b_{(0)}^3$	0.5515	0.8903	0.7005	0.5461	0.7337	0.9359	0.9899	1.0000

Table 41: Central values, errors, and correlation matrix for the BCL z -parametrisations of Fermilab/MILC form factors, values are taken from [51].

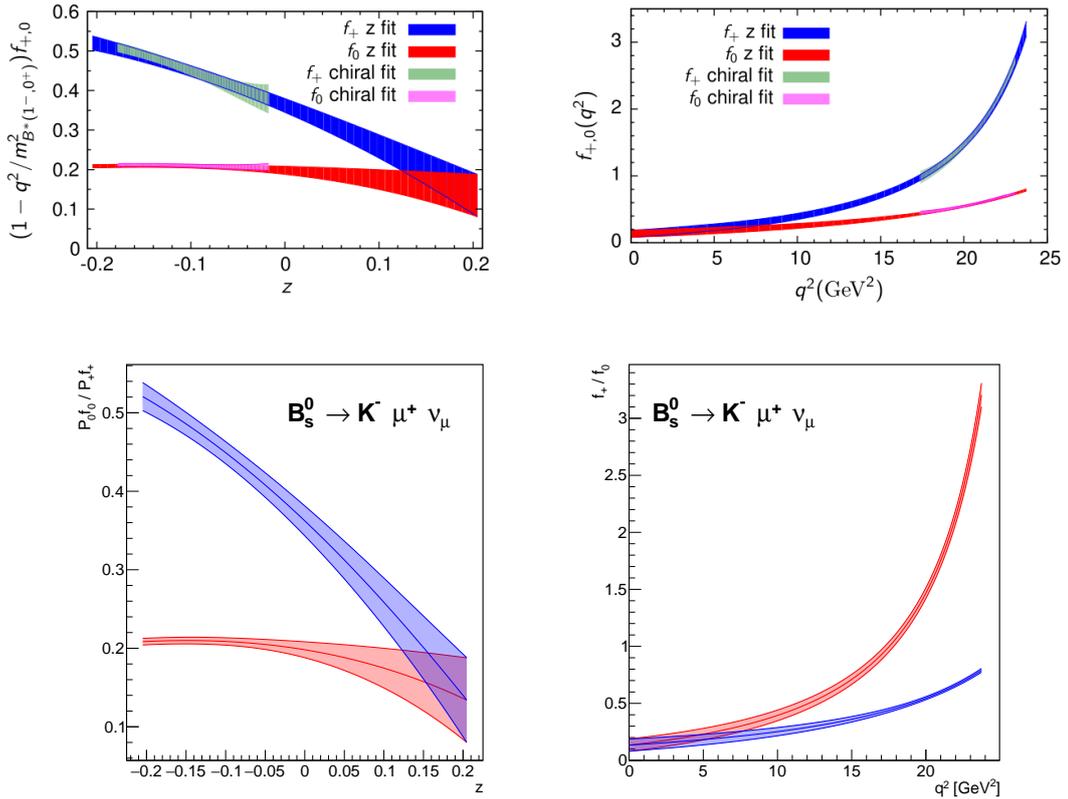


Figure 91: Form factors from Fermilab/MILC plotted against z on the left and on the right plotted against q^2 . The top two images are taken from [51] whereas the bottom two plots are generated using fit parameters taken from this paper. In the bottom left plot the scalar form factor f_0 is drawn in red whereas in the bottom right plot the colors between the form factors are reversed.

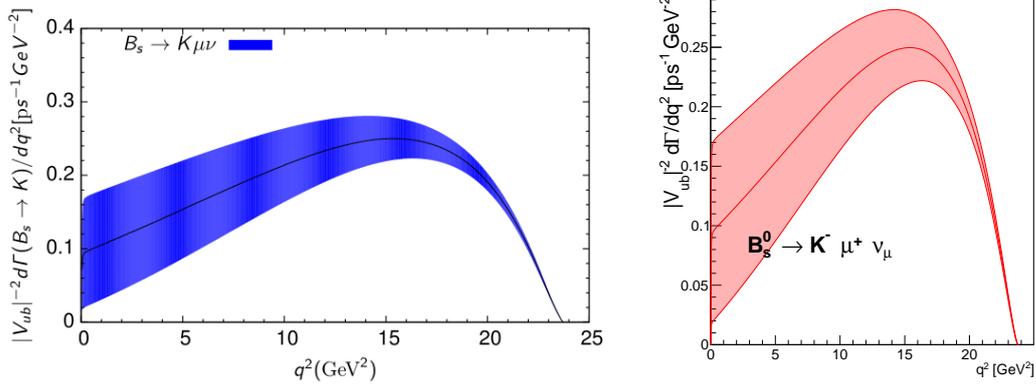


Figure 92: The differential $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay rate from Fermilab/MILC plotted against q^2 on the left from [51] and on the right generated using fit parameters taken from this paper.

	$f^+(0)$	b_1^+	$f^T(0)$	b_1^T
Value	0.336	-2.53	0.320	-1.08
Error	0.023	1.17	0.019	1.53
$f^+(0)$	1.00	0.79	0.000	0.00
b_1^+	0.79	1.00	0.000	0.00
$f^T(0)$	0.00	0.00	1.000	0.74
b_1^T	0.00	0.00	0.740	1.00

Table 42: Central values, errors, and correlations for the BCL z -parametrisations of f^+ and f^T form factors from LCSR predictions. Parameters taken from [52].

A.1.4 LCSR parameters

Light-cone sum rules parameters are provided by Reference [52]. They use a modified version of the z -expansion which is a rescaling of the coefficients to keep the form factor at zero momentum transfer $f^{+,T}(0)$ as one of the independent parameters. The general form is given by

$$f^{+,T}(q^2) = \frac{f^{+,T}(0)}{1 - q^2/M^2} \left(1 + \sum_{k=0}^{K-1} b_k^{+,T} \left[z(q^2)^k - z(0)^k - (-1)^{k-K} \frac{k}{K} (z(q^2)^K - z(0)^K) \right] \right), \quad (100)$$

where in Reference [52] $K=2$ is used for the BCL-parametrisation. Table 42 gives the z -expansion parameters used in their paper. Figure 93 and 94 show plots from their paper along with the reproduced plots using the coefficients from Table 42 and generated with the analysis software.

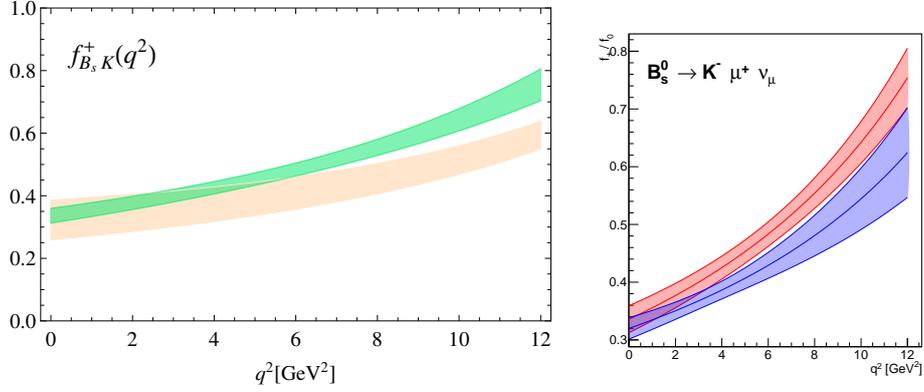


Figure 93: Form factors from LCSR plotted against q^2 . Left Figure taken from [52] and on the right generated using fit parameters taken from this paper. The green band in the left Figure should be compared to the red distribution on the right. Figure modified from [9].

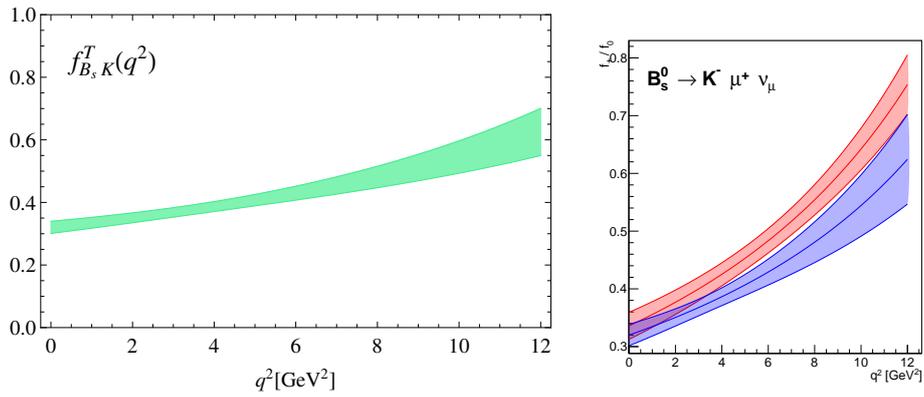


Figure 94: Form factors from LCSR plotted against q^2 . Left Figure taken from [52] and on the right generated using fit parameters taken from this paper. The green band in the left Figure should be compared to the blue distribution on the right. Figure modified from [9].

collaboration [ref.]	ensemble	heavy b-quarks	valence c-quark
HPQCD [59]	$N_f = (2 + 1)$ MILC-asqtad	NRQCD b quarks	HISQ action
Fermilab/MILC [58]	$N_f = (2 + 1)$ MILC-asqtad	Fermilab b quark	Fermilab c-quarks
HPQCD 2019 [47]	$N_f = (2 + 1 + 1)$ MILC-HISQ	heavy-HISQ approach for b quark	HISQ action

Table 43: Lattice simulation details of the papers providing form factor results for $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$.

	$a_0^{(0)}$	$a_1^{(0)}$	$a_2^{(0)}$	$a_0^{(+)}$	$a_1^{(+)}$	$a_2^{(+)}$
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 44: Central values, errors, and covariance matrix for the BCL z -parametrisations of HPQCD form factors from for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. Parameters are taken from [59].

A.2 $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ publications

For the normalisation channel $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ three different for factor calculations from LQCD predictions are used corresponding to References [47, 58, 59]. The Lattice calculations differ in their simulation set-up, which is compared in Table 43. It should be noted that Reference [58] extracts the fit parameters of $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ together with $B \rightarrow D$ parameters as a ratio of form factors, which is why they are correlated. In Reference [51] they reconstruct the $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ form factors from the ratio provided in [58] together with $B \rightarrow D$ form factors from [133]. Therefore the form factors taken from Reference [58] alone should be treated with caution.

A.2.1 HPQCD parameters

Table 44 gives the z -expansion fit parameters for the BCL parametrisation with $K=3$ from the HPQCD collaboration [59]. Figure 95- 97 show plots from their older paper along with the reproduced plots using the coefficients from Table 44 which are generated with the analysis software.

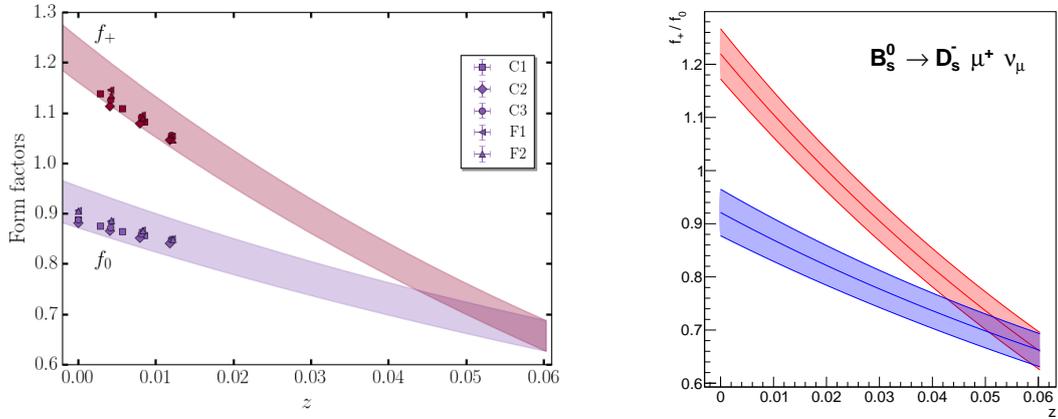


Figure 95: Form factors from HPQCD plotted against z . Left Figure taken from [59] and on the right generated using fit parameters taken from this paper. The red band in the right Figure corresponds to the vector form factor f_+ , while the blue band shows the scalar form factor f_0 . Figure modified from [9].

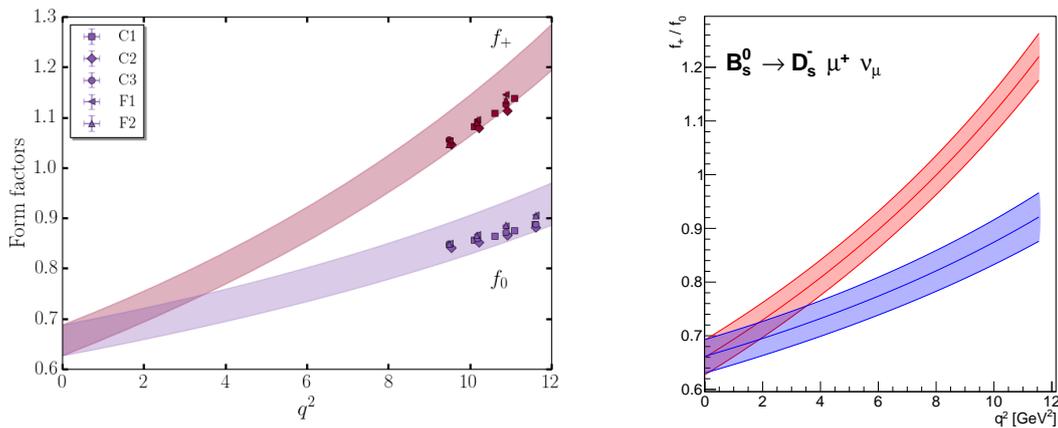


Figure 96: Form factors from HPQCD plotted against q^2 . Left Figure taken from [59] and on the right generated using fit parameters taken from this paper. The red band in the right Figure corresponds to the vector form factor f_+ , while the blue band shows the scalar form factor f_0 . Figure modified from [9].

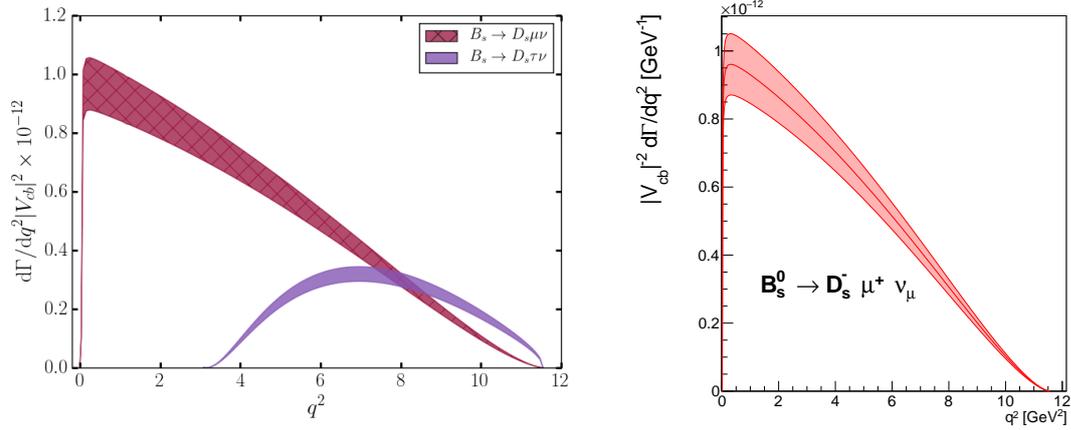


Figure 97: The differential $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decay rate from HPQCD plotted against q^2 . Left Figure taken from [59] and on the right generated using fit parameters taken from this paper. The dark red distribution in the left Figure should be compared to the distribution in the right Figure. Figure modified from [9].

	$a_0^{(+)}$	$a_1^{(+)}$	$a_2^{(+)}$	$a_0^{(0)}$	$a_1^{(0)}$	$a_2^{(0)}$
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 45: Central values, errors, and correlation matrix for the BCL z -parametrisations of Fermilab/MILC form factors for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. Values are taken from [58].

A.2.2 Fermilab/MILC parameters

Table 45 gives the z -expansion fit parameters for the BCL parametrisation with $K=3$ from the Fermilab/MILC collaboration [58]. Figure 98 shows the plot from their paper along with the reproduced plots using the coefficients from Table 45 and generated with the analysis software.

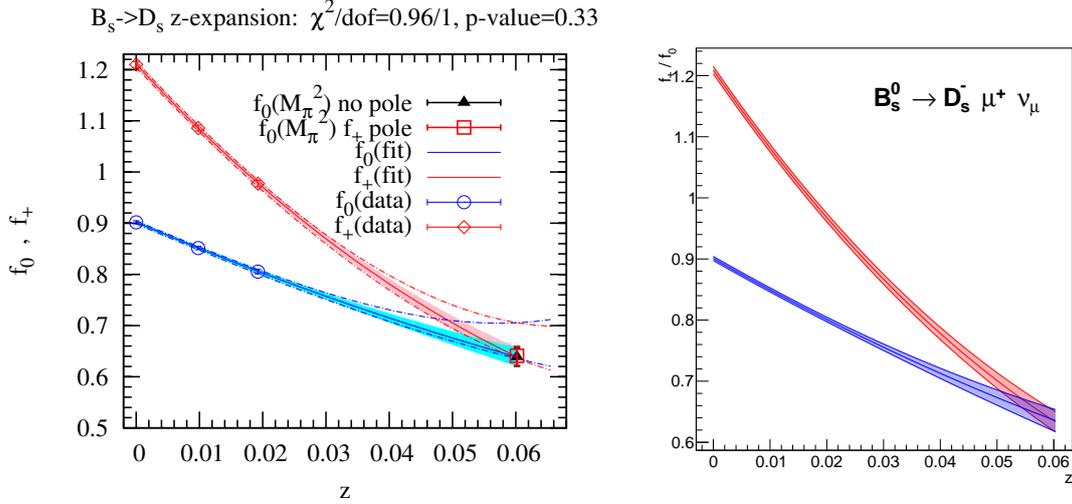


Figure 98: Form factors from Fermilab/MILC plotted against z . Left Figure taken from [58] and on the right generated using fit parameters taken from this paper. The red distribution in the right Figure corresponds the vector form factor f_+ while the blue shows the scalar form factor f_0 . Figure modified from [9].

	$a_0^{(+)}$	$a_1^{(+)}$	$a_2^{(+)}$	$a_0^{(0)}$	$a_1^{(0)}$	$a_2^{(0)}$
Value	0.66574	-0.25944	-0.10636	0.66574	-3.23599	-0.07478
$a_0^{(+)}$	0.00015	0.00188	0.00070	0.00015	0.00022	0.00003
$a_1^{(+)}$	0.00188	0.06129	0.16556	0.00188	0.01449	0.00001
$a_2^{(+)}$	0.00070	0.16556	3.29493	0.00070	0.18757	-0.00614
$a_0^{(-)}$	0.00015	0.00188	0.00070	0.00015	0.00022	0.00003
$a_1^{(-)}$	0.00022	0.01449	0.18757	0.00022	0.20443	0.10080
$a_2^{(-)}$	0.00003	0.00001	-0.00614	0.00003	0.10080	4.04413

Table 46: Central values, errors, and correlation matrix for the BCL z -parametrisations of the latest HPQCD form factors for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$. Coefficients are taken from [47].

A.2.3 HPQCD 2019 parameters

Table 46 gives the z -expansion fit parameters for the BCL parametrisation with $K=3$ from the latest paper from the HPQCD collaboration [47]. Figure 99- 101 show plots from their paper along with the reproduced plots using the coefficients from Table 46 and generated with the analysis software.

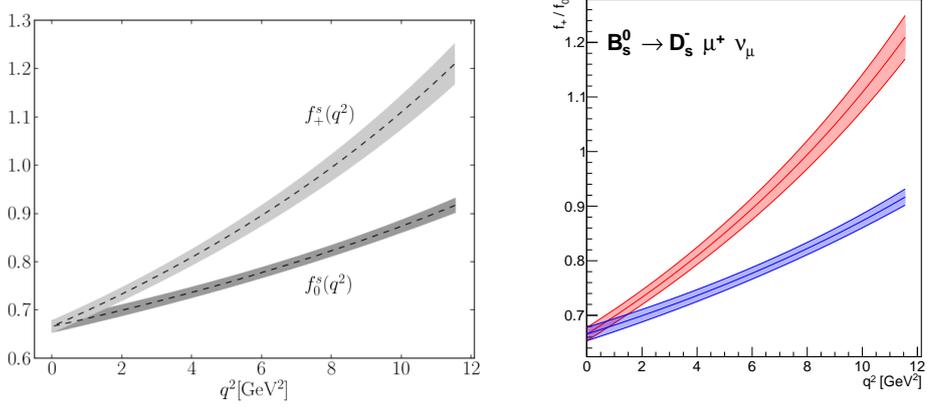


Figure 99: Latest Form factors from HPQCD plotted against q^2 . Left Figure taken from [47] and on the right generated using fit parameters taken from this paper. The red distribution in the right Figure corresponds the vector form factor f_+ while the blue shows the scalar form factor f_0 .

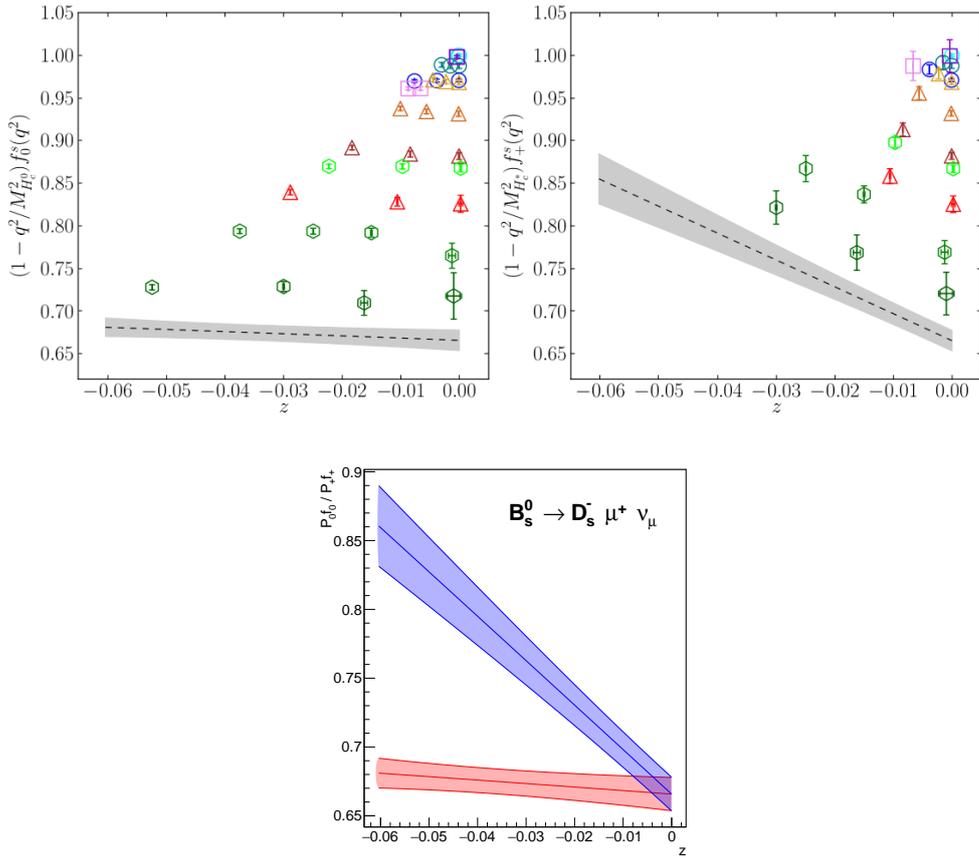


Figure 100: Latest Form factors from HPQCD plotted against z . The top two plots are Figures taken from [47] while the bottom one is generated using fit parameters taken from this paper. The blue distribution in the bottom Figure corresponds to the vector form factor f_+ shown in top right while the red band of the bottom Figure shows the scalar form factor f_0 displayed in the top left Figure.

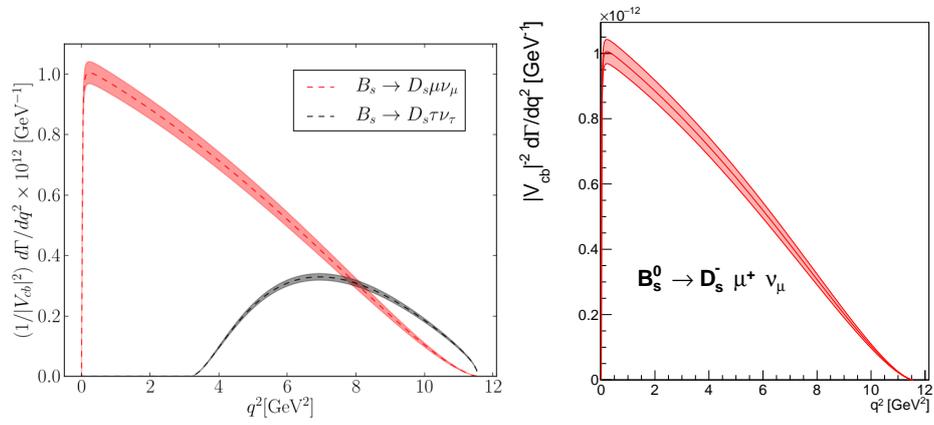


Figure 101: The differential $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu$ decay rate from the latest HPQCD paper plotted against q^2 . The left Figure is taken from [47] and the right is generated using fit parameters taken from this paper. The red distribution in the left Figure should be compared to the distribution in the right Figure.

B Validation of selection BDT

As explained in the end of Section 6 the two BDT used in the selection have to be validated using the control channel $B^+ \rightarrow J/\psi K^+$. This section shows the input variables used in the BDTs, Monte Carlo simulation is compared to background subtracted data in the following plots. The channel is fully reconstructed as $B^+ \rightarrow J/\psi K^+$ as this reconstruction is used to evaluate also the BDT efficiency. Figure 102 and 103 show the input distribution for the charge BDT, Figure 104 shows the distribution for the input variables for the SS BDT.

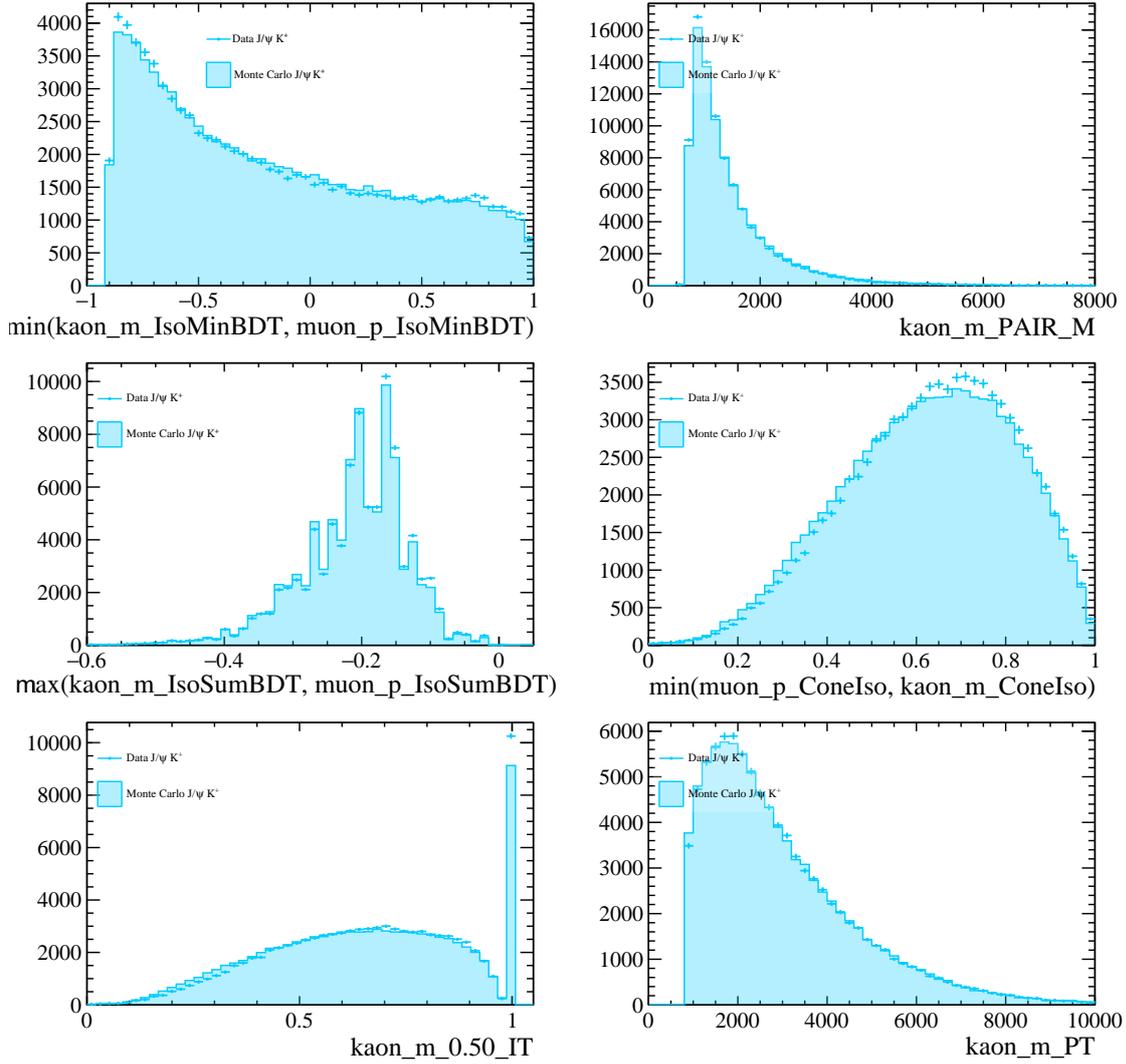


Figure 102: Distributions of mixed events before (blue) and after (red) kinematic reweighting compared to the data in the region above the B_s^0 mass $m_{K-\mu^+} > m_{B_s^0}$ (black) for the corrected mass error (top left), B_s^0 p_T (top right), missing p_T (bottom left) and the different B_s^0 flight distance coordinates.

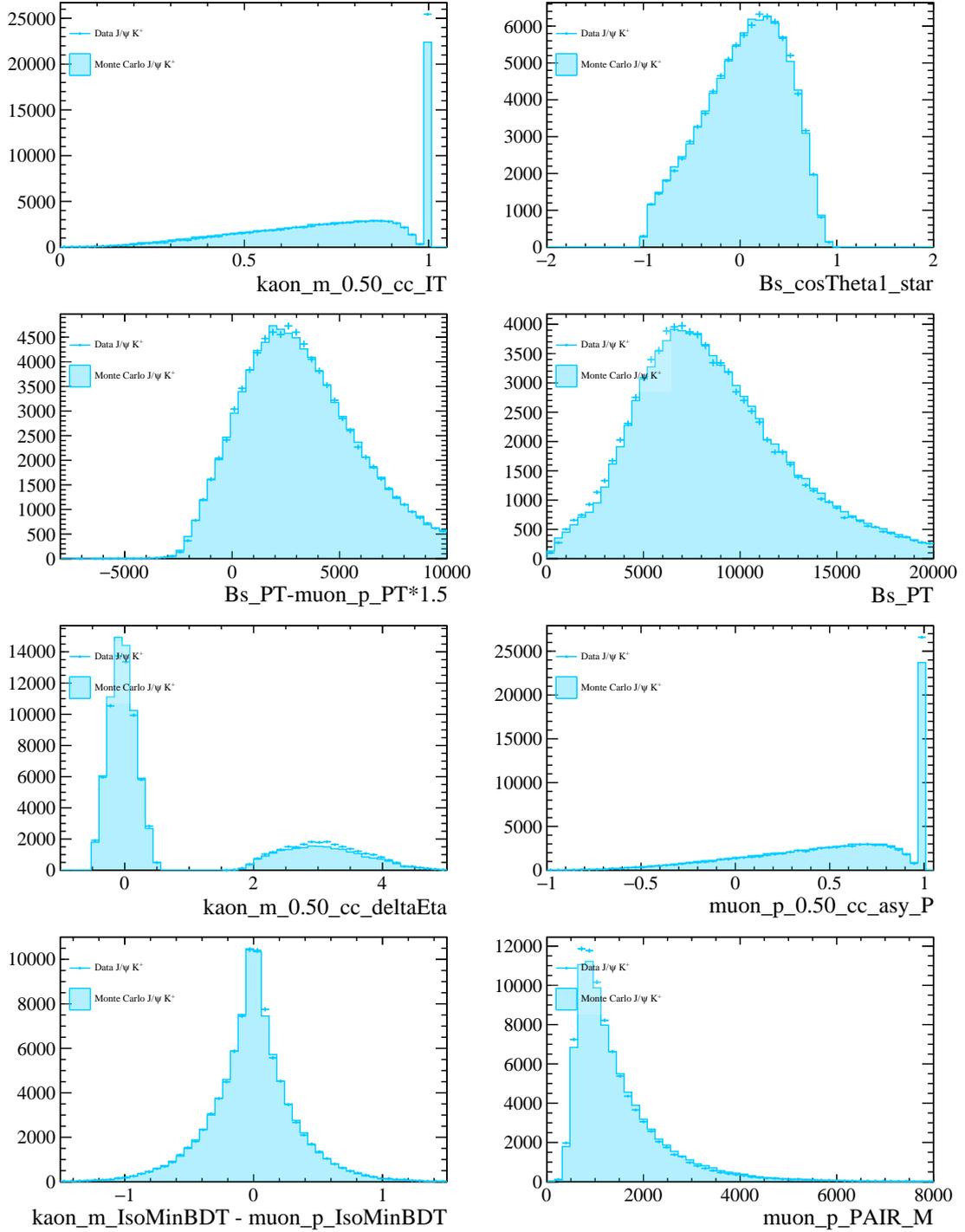


Figure 103: Distributions of mixed events before (blue) and after (red) kinematic reweighting compared to the data in the region above the B_s^0 mass $m_{K-\mu^+} > m_{B_s^0}$ (black) for the corrected mass error (top left), B_s^0 p_T (top right), missing p_T (bottom left) and the different B_s^0 flight distance coordinates.

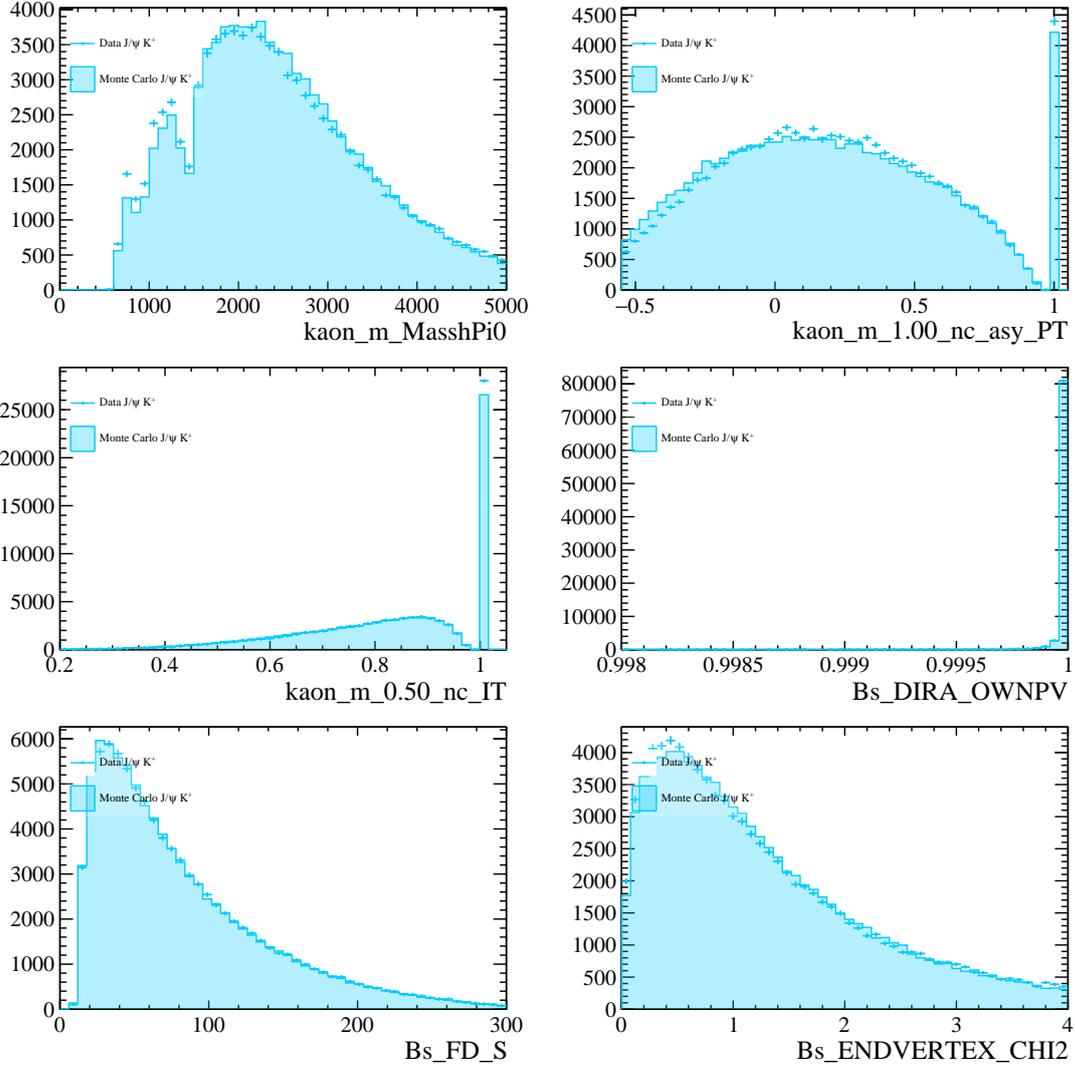


Figure 104: Distributions of mixed events before (blue) and after (red) kinematic reweighting compared to the data in the region above the B_s^0 mass $m_{K^-\mu^+} > m_{B_s^0}$ (black) for the corrected mass error (top left), B_s^0 p_T (top right), missing p_T (bottom left) and the different B_s^0 flight distance coordinates.

C Validation of reweighting BDT

As explained in detail in Section 7 a separate BDT is trained to correct for MC and data differences. In addition to the validation plots given in Section 7, here all reweighted kinematic distributions are shown for $B^+ \rightarrow J/\psi K^+$ in Figure 105 and for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$ in Figure 106.

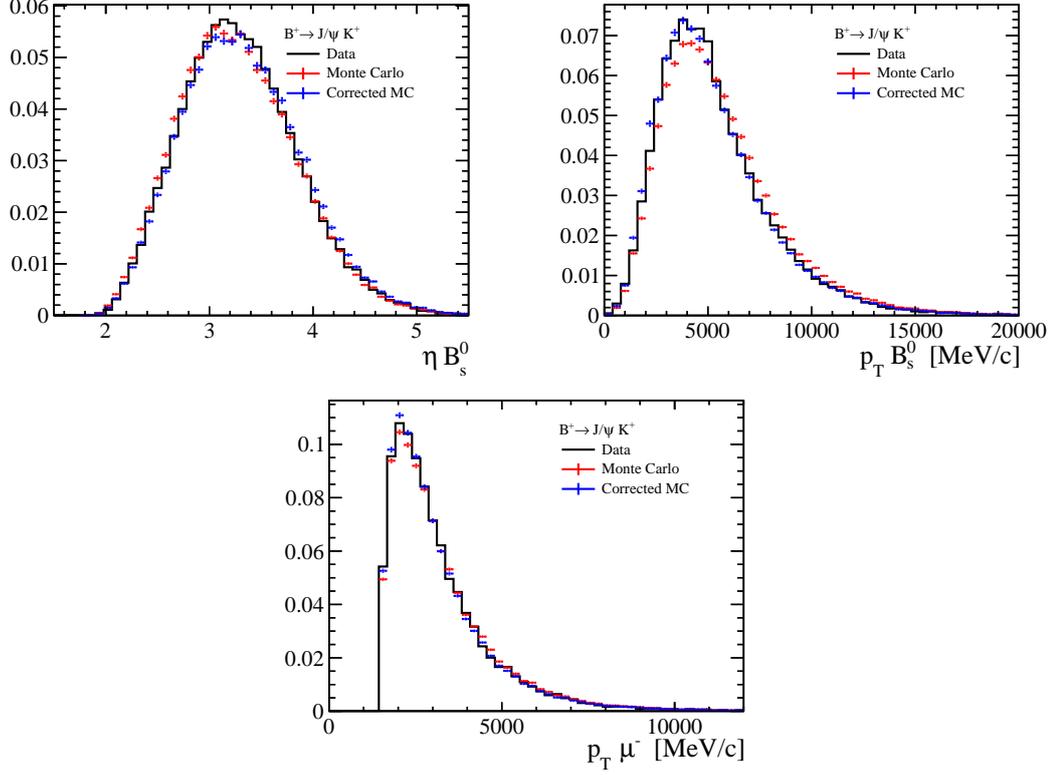


Figure 105: Kinematic distributions for $B^+ \rightarrow J/\psi K^+$ reconstructed using the $K^+ \mu^-$ pair for simulation before and after correction, plotted against background subtracted data. Figure taken from [9].

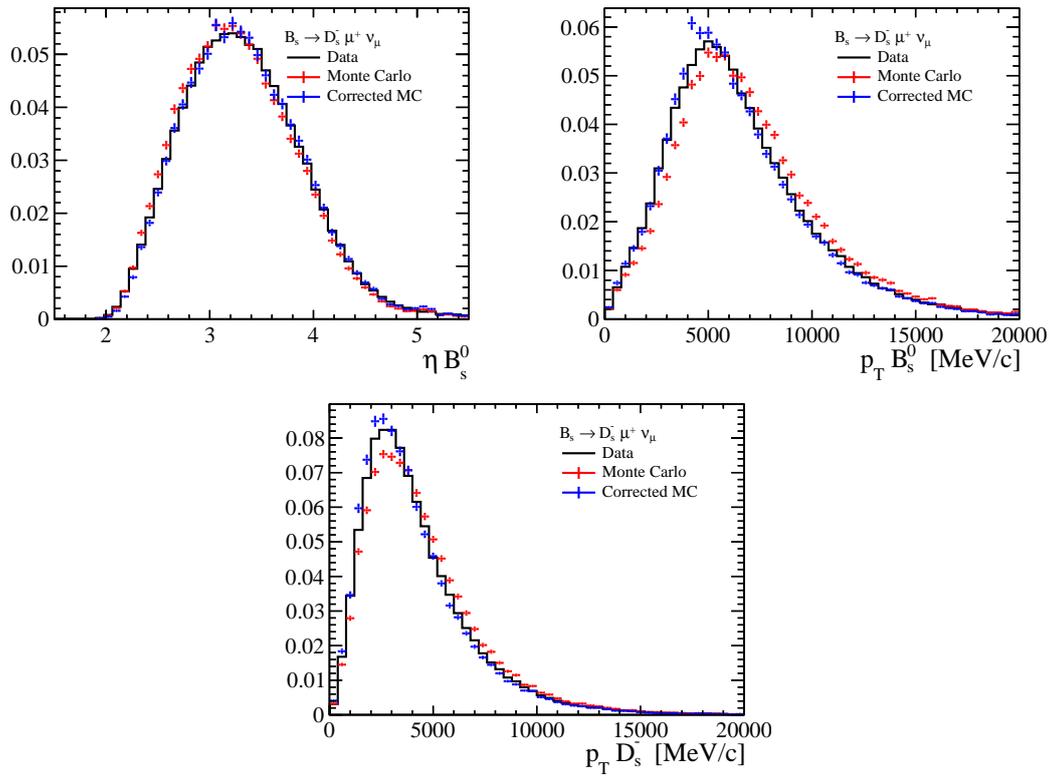


Figure 106: Kinematic distributions for $B_s^0 \rightarrow D_s^- \mu^+ \nu_\mu X$ before and after correction using the BDT reweighting compared to data. Figure taken from [9].

D Validation of event mixing

This appendix shows additional plots for the validation of the event mixing which is used to model the combinatoric background for the signal channel $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$. The event mixing method itself was explained already in Section 9. It needs to be reweighted in order to reproduce the B_s^0 meson kinematics in the region above the B_s^0 mass, this is done in the momentum and transverse momentum of the B_s^0 candidate. Additional kinematic distributions for modelled events before (in blue) and after kinematic reweighting (red points) are compared with data (black line) above the B_s^0 mass in Figure 107.

The kinematical distribution of the event mixed $\mu^+ K^-$ pair in the full accessible range is shown on Figure 108.

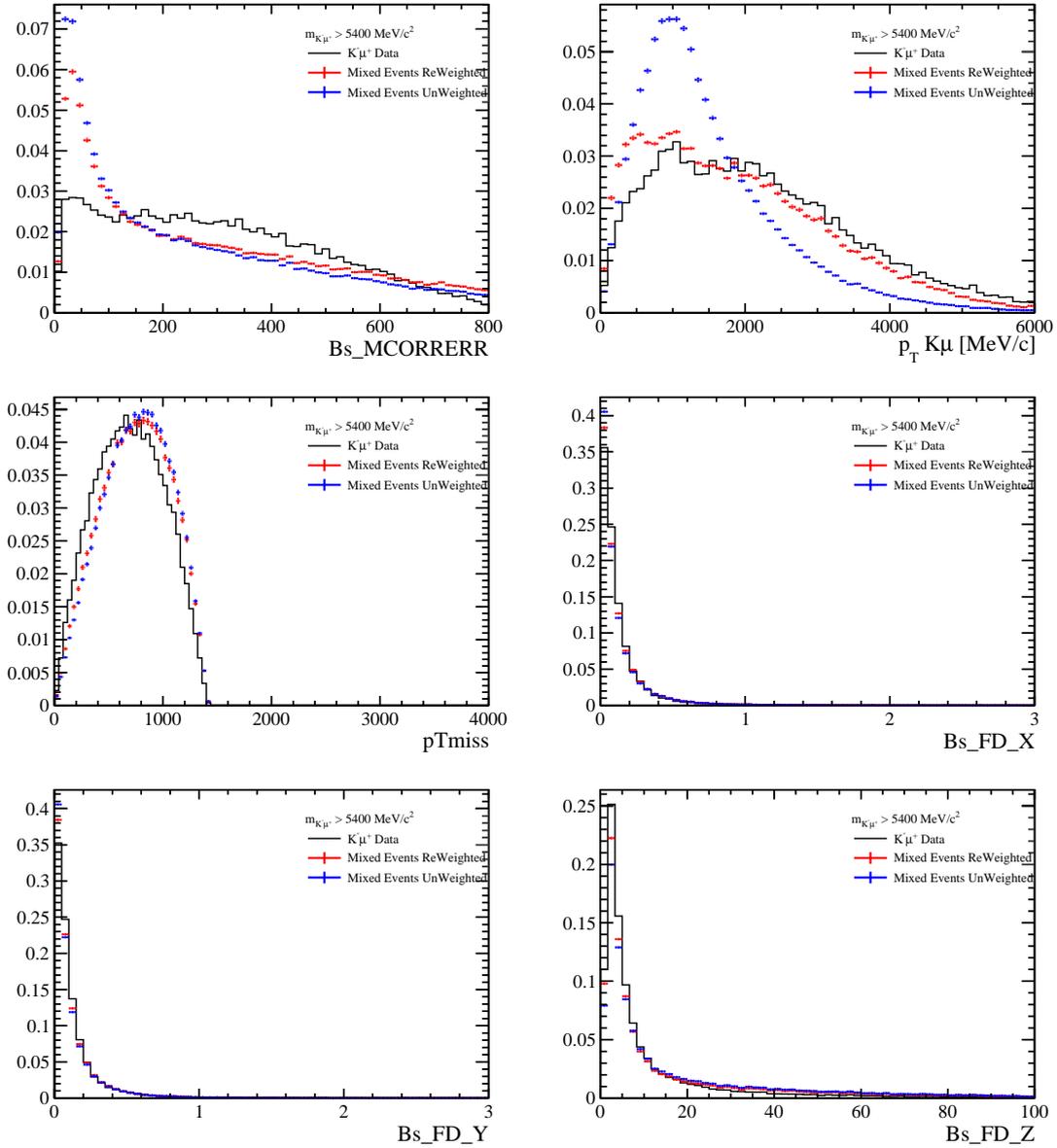


Figure 107: Distributions of mixed events before (blue) and after (red) kinematic reweighting compared to the data in the region above the B_s^0 mass $m_{K\mu^+} > m_{B_s^0}$ (black) for the corrected mass error (top left), B_s^0 p_T (top right), missing p_T (bottom left) and the different B_s^0 flight distance coordinates. Figure taken from [9].

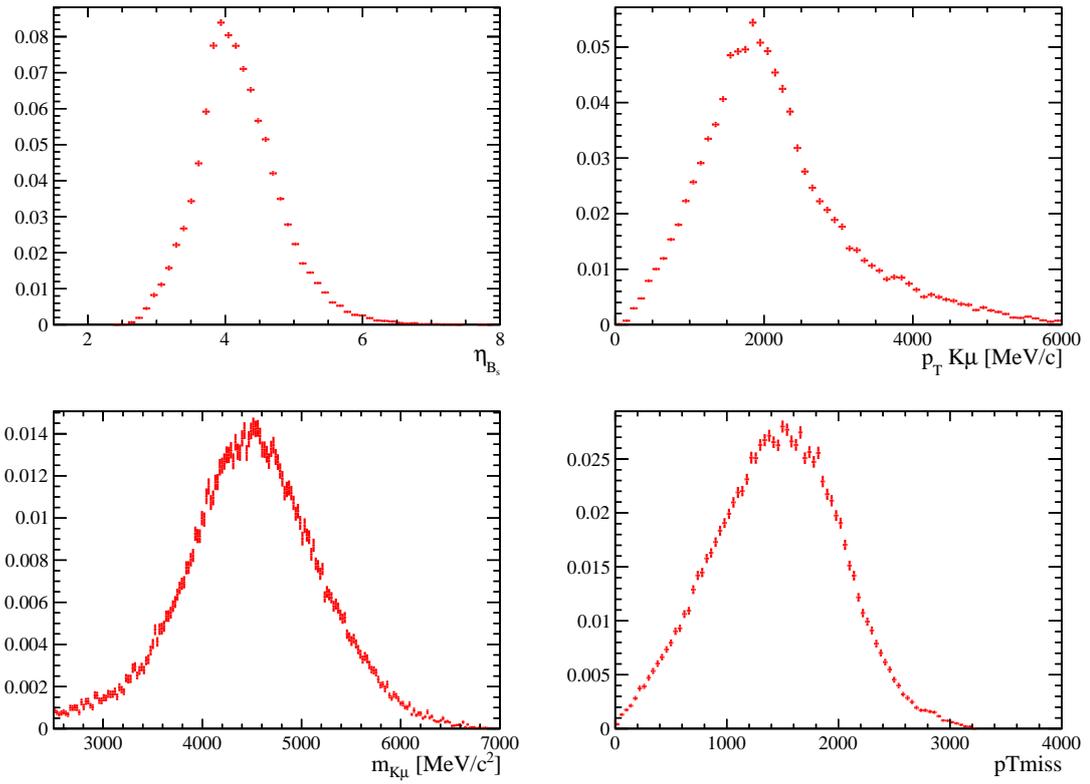


Figure 108: Distributions of mixed events for the B_s^0 pseudorapidity (top left), B_s^0 p_T (top right), invariant $K^- \mu^+$ mass (bottom right), missing p_T (bottom left). Figure taken from [9].

Trigger Line	Simulation	data	correction (N. σ w.r.t 1.)
L0Muon (Down)	94.09 ± 0.27	94.41 ± 0.16	0.996 ± 0.003 (1.3)
L0Muon (Up)	92.90 ± 0.30	93.65 ± 0.18	0.992 ± 0.004 (2.0)
Hlt2SingleMuon (Down)	54.79 ± 2.91	55.73 ± 1.24	0.983 ± 0.057 (0.3)
Hlt2SingleMuon (Up)	53.51 ± 3.03	57.58 ± 1.27	0.929 ± 0.057 (1.3)
TopoMu2BodyBBDT(Down)	93.95 ± 1.42	94.41 ± 0.80	0.995 ± 0.017 (0.3)
TopoMu2BodyBBDT(Up)	91.42 ± 1.61	93.59 ± 0.85	0.976 ± 0.019 (1.3)

Table 47: TISTOS efficiency for each trigger line used for the signal decay in simulation and data. The third row shows the correction factor data/simulation.

E TisTos efficiency checks

In order to crosscheck the trigger efficiency, the TISTOS efficiency is also evaluated in different 2-D binnings by using the IP or the IPChi2 instead of the B_s^0 corrected mass. Both variables were studied following the procedure outlined in Section 10. The correction factors for each variable change are comparable with the ones given in Table 28 with only minor changes. Those changes are fully covered by the uncertainty of the correction factor and therefore no further systematic uncertainty is assigned.

Another cross check of the trigger efficiency is performed by estimating the TISTOS efficiencies for the individual trigger lines used in this thesis: L0Muon, HLT2SingleMuon and TopoMu2BodyBBDT. Those lines are considered separately and the check is performed using again the $B^+ \rightarrow J/\psi K^+$ channel. Table 47 shows the TISTOS ratio for each trigger line used along with the correction factor split in magnet polarity. No significant correction is observed following this method.

The variation of the TISTOS efficiency is further studied in bins of B_s^0 η and p_T distributions in Figs. 109,110 as well as in bins of daughters kinematics for the momentum and transverse momentum which are shown in Figure 111 to 114. The distributions between data and MC agree very well.

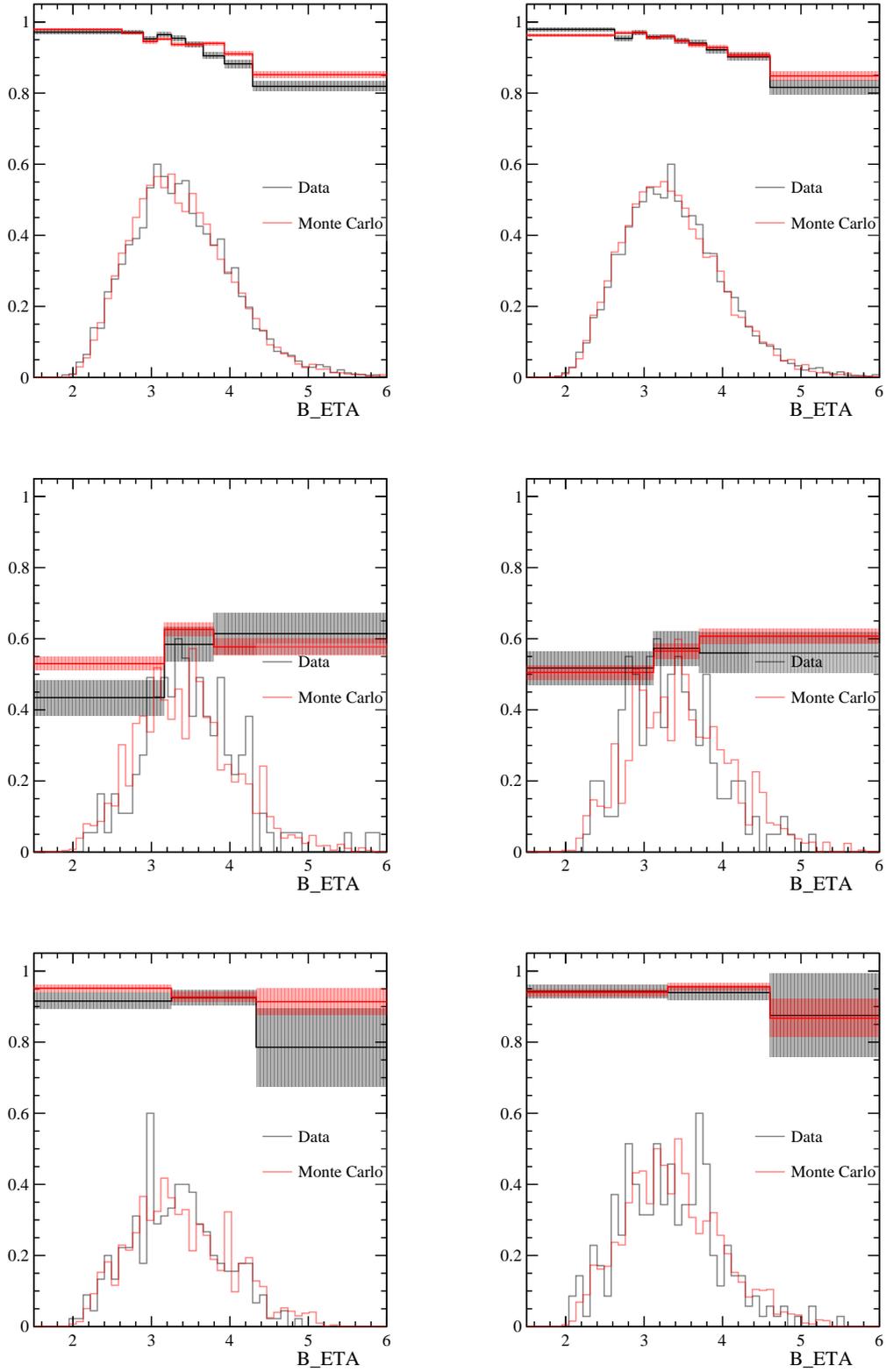


Figure 109: TISTOS efficiency variations in bins of $B_s^0 \eta$ for L0Muon trigger on top, Hlt2SingleMuon in the middle and TopoMu2BodyBBDT on the bottom. Those are determined per magnet polarity, on the left for up-polarity and on the right for down polarity. Figure taken from [9].

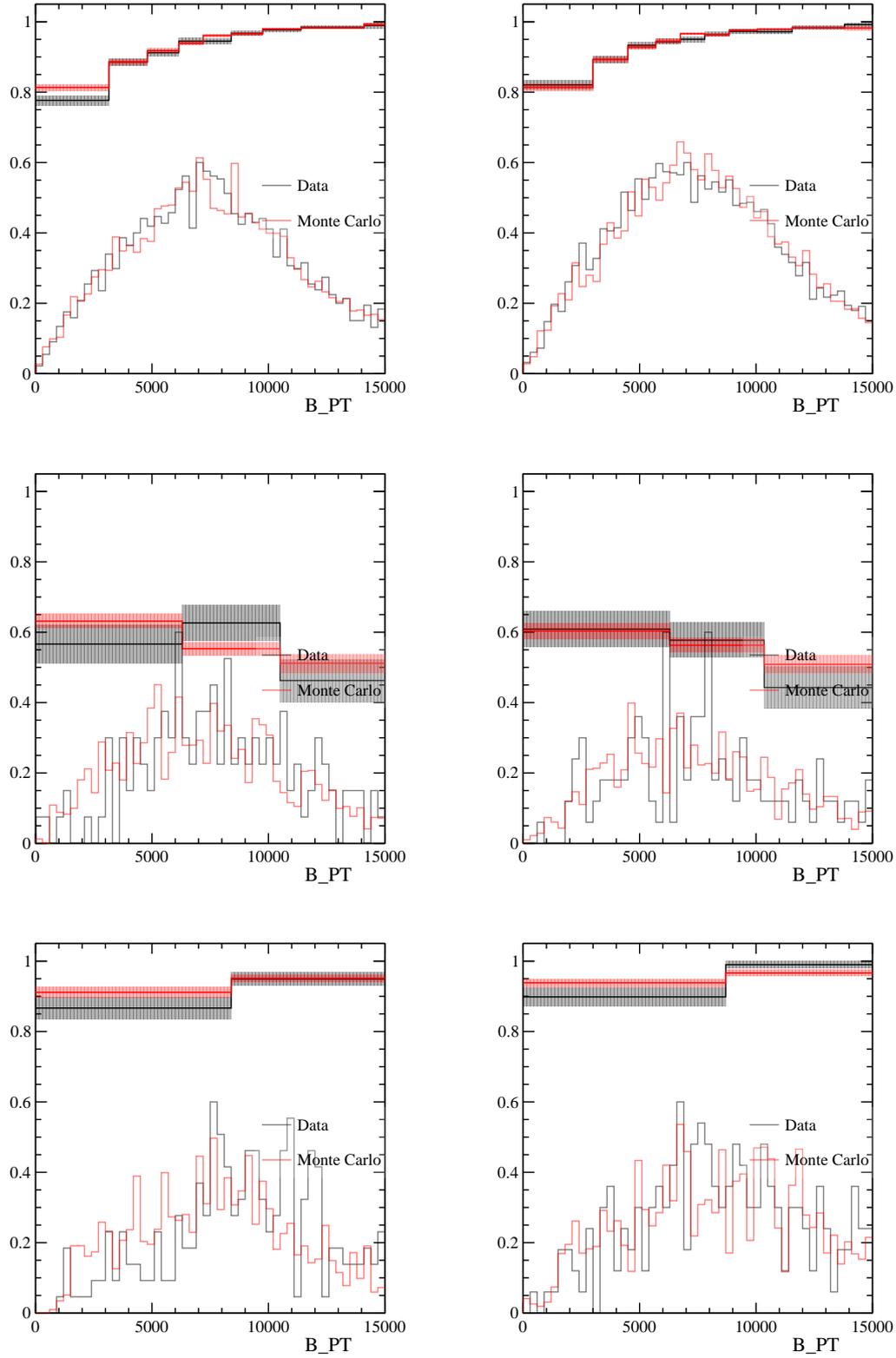


Figure 110: TISTOS efficiency variations in bins of $B_s^0 p_T$ for the L0Muon trigger on top, Hlt2SingleMuon in the middle and TopoMu2BodyBBDT on the bottom. Those are determined per magnet polarity, on the left for up-polarity and on the right for down polarity. Figure taken from [9].

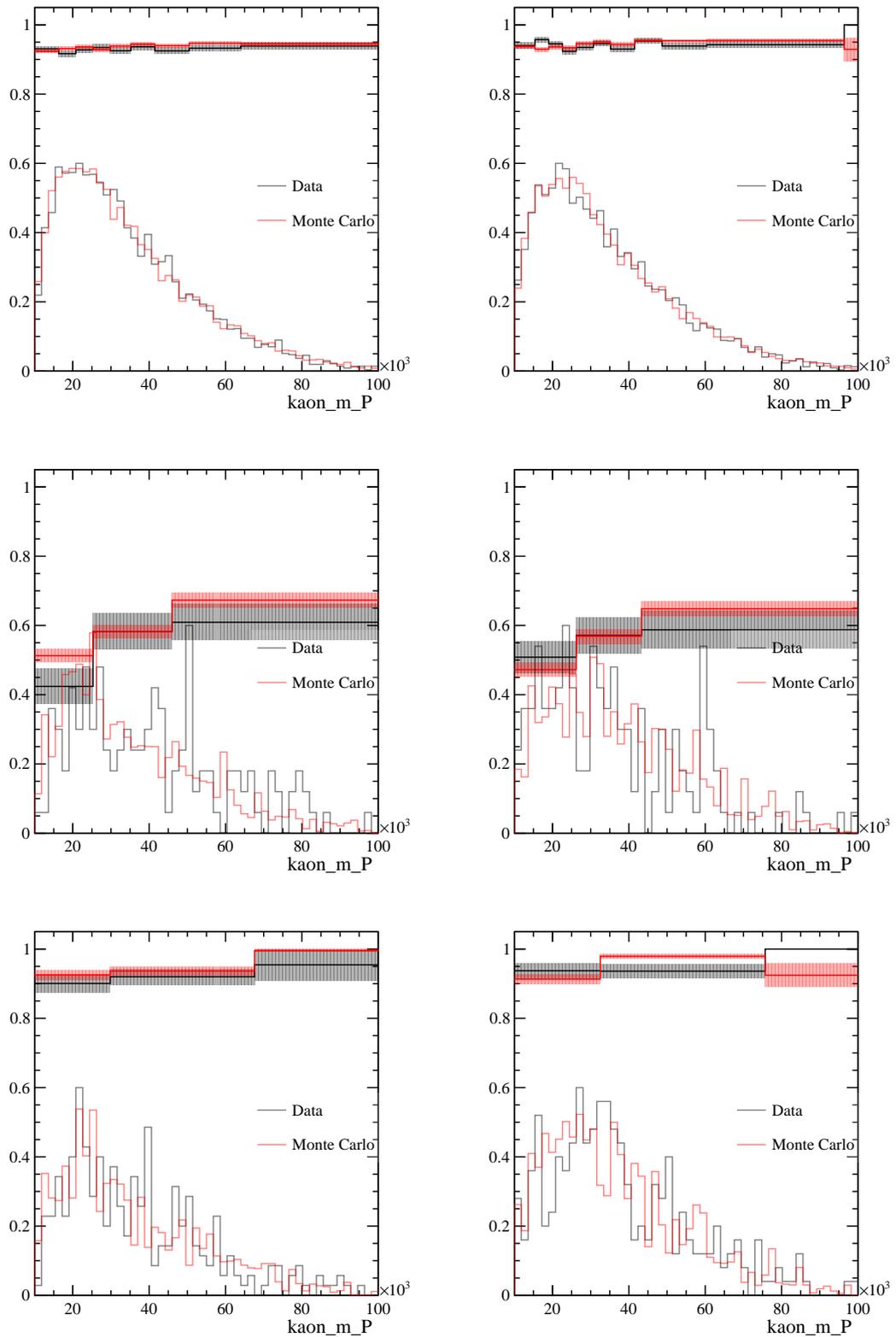


Figure 111: TISTOS efficiency variations in bins of Kaon momenta for the L0Muon trigger on top, Hlt2SingleMuon in the middle and TopoMu2BodyBBDT on the bottom. Those are determined per magnet polarity, on the left for up-polarity and on the right for down polarity. Figure taken from [9].

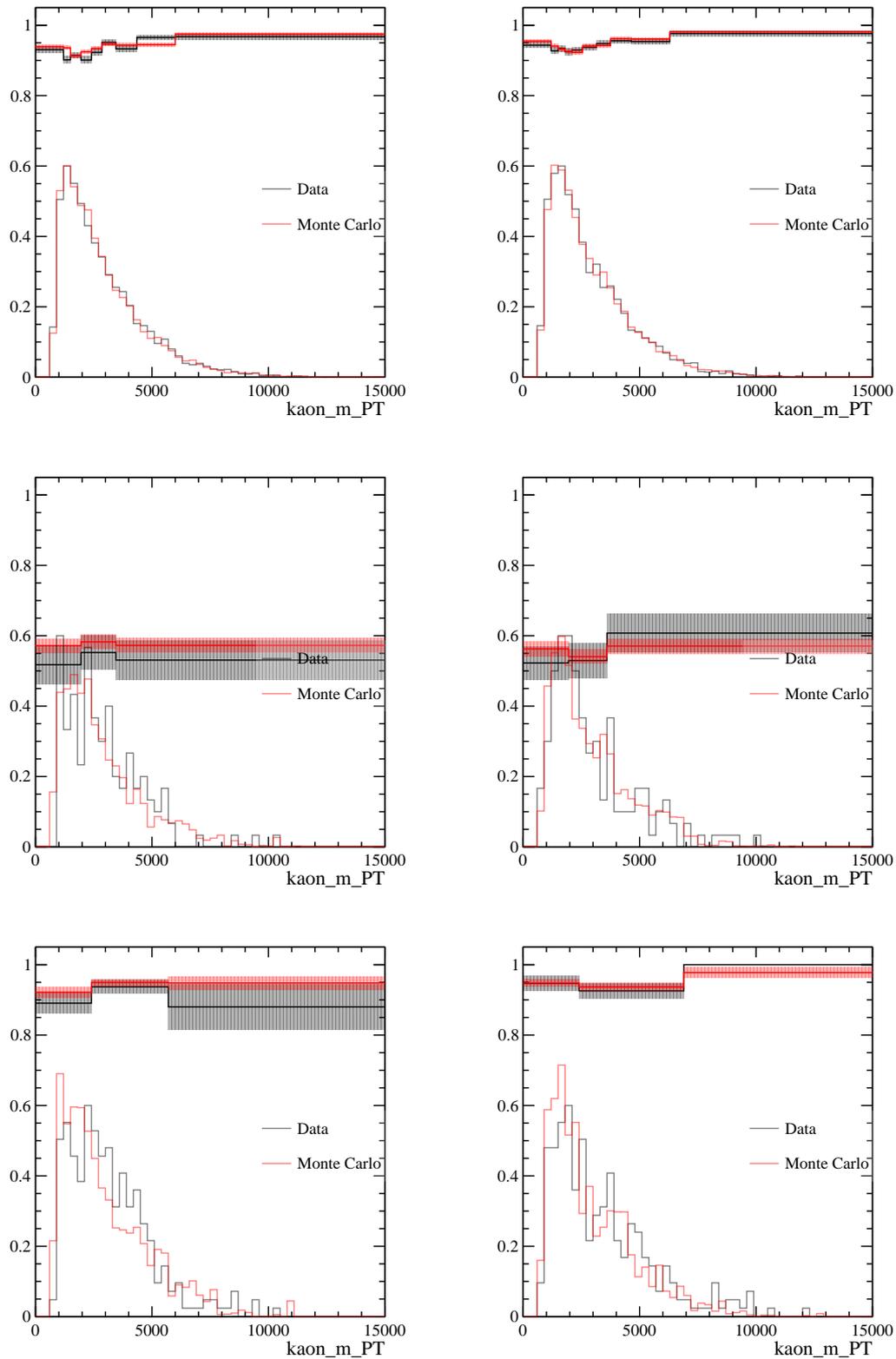


Figure 112: TISTOS efficiency variations in bins of Kaon p_T for the L0Muon trigger on top, Hlt2SingleMuon in the middle and TopoMu2BodyBBDT on the bottom. Those are determined per magnet polarity, on the left for up-polarity and on the right for down polarity. Figure taken from [9].

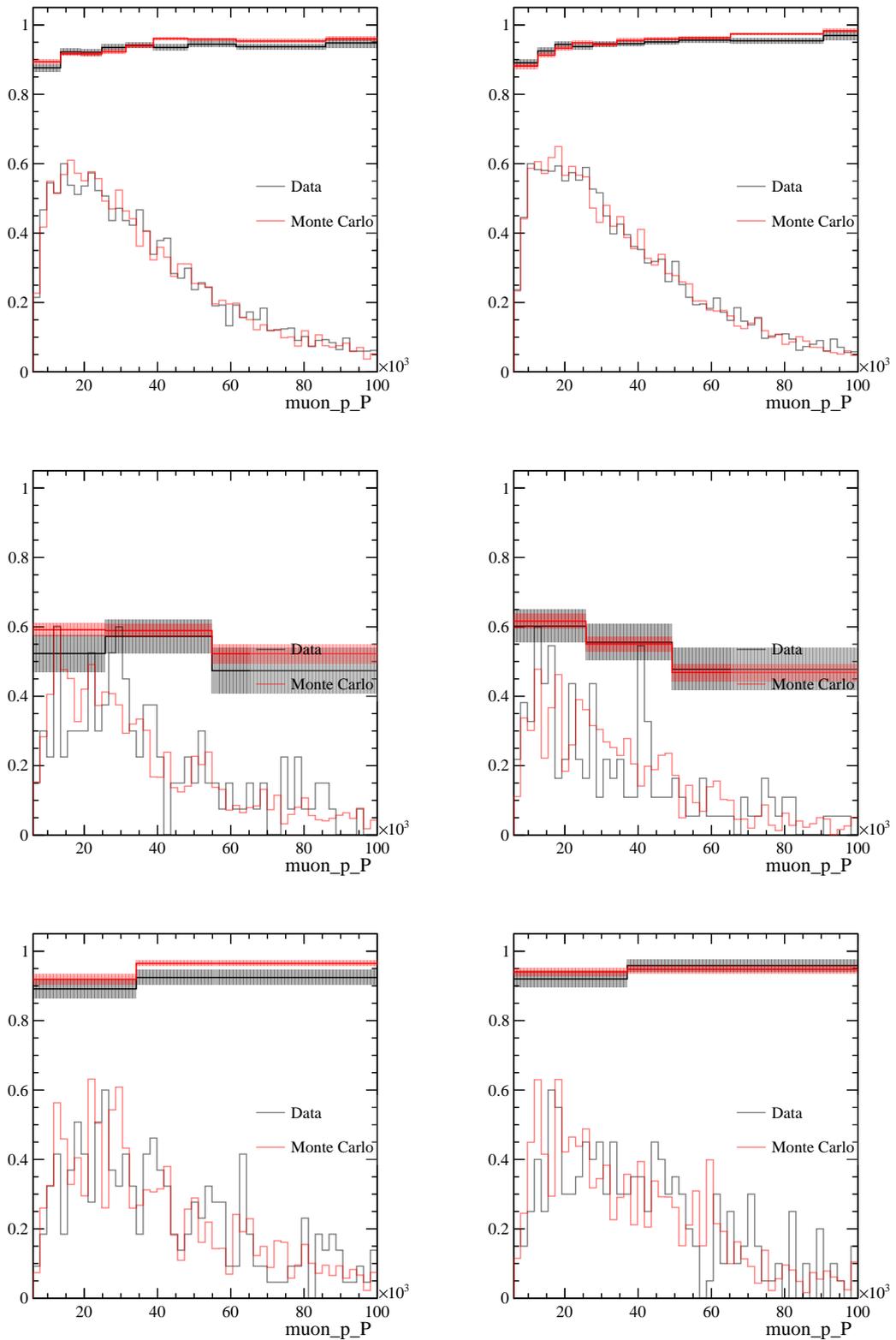


Figure 113: TISTOS efficiency variations in bins of muon momenta for the L0Muon trigger on top, Hlt2SingleMuon in the middle and TopoMu2BodyBBDT on the bottom. Those are determined per magnet polarity, on the left for up-polarity and on the right for down polarity. Figure taken from [9].

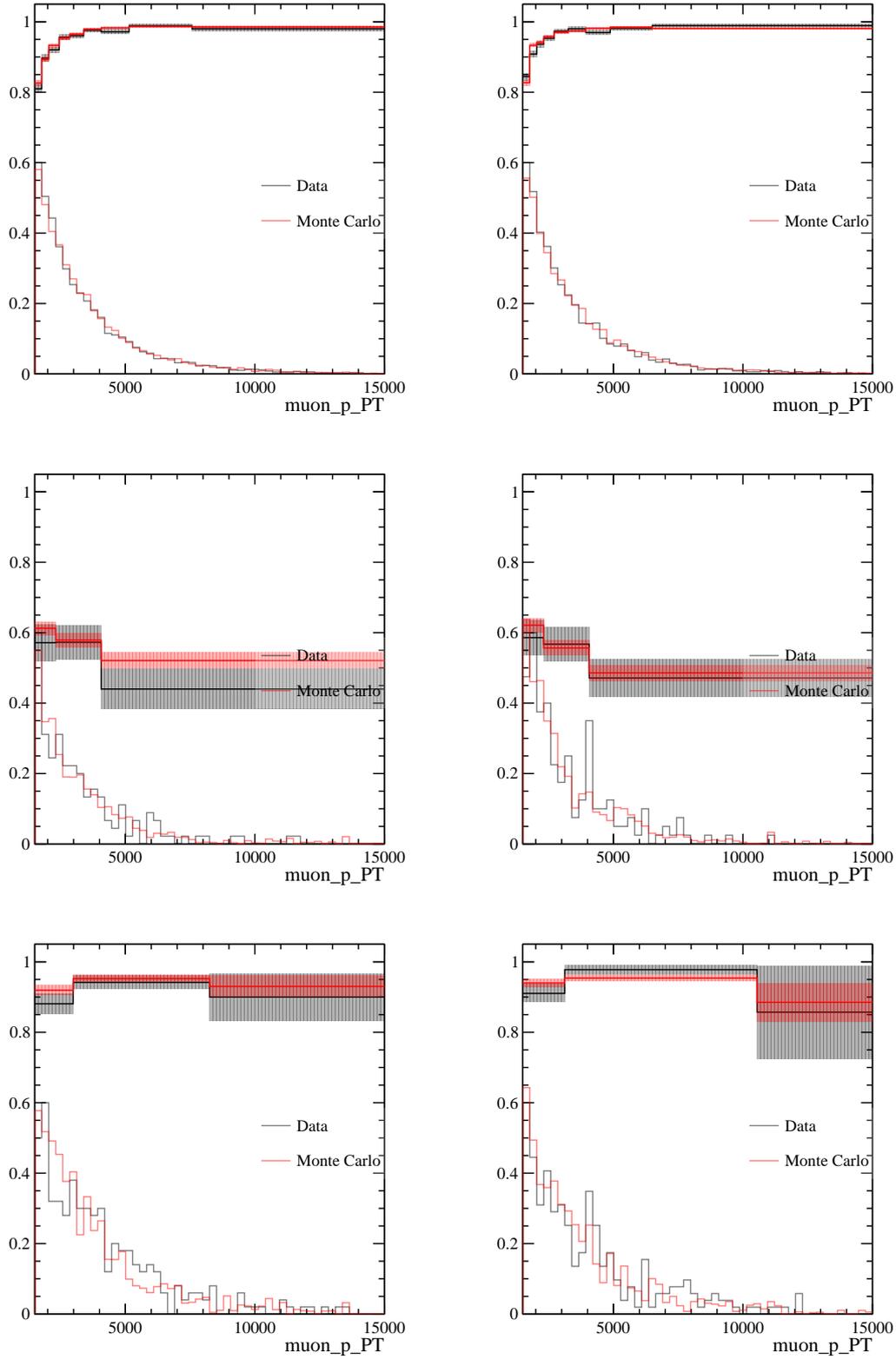


Figure 114: TISTOS efficiency variations in bins of muon p_T for the L0Muon trigger on top, Hlt2SingleMuon in the middle and TopoMu2BodyBBDT on the bottom. Those are determined per magnet polarity, on the left for up-polarity and on the right for down polarity. Figure taken from [9].

References

- [1] 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](#).
- [2] 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](#).
- [3] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531, [648(1963)].
- [4] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.
- [5] LHCb, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, *Nature Phys.* **11** (2015) 743, [arXiv:1504.01568](#).
- [6] H. J. Rothe, *Lattice gauge theories: An Introduction*, World Sci. Lect. Notes Phys. **43** (1992) 1, [World Sci. Lect. Notes Phys.82,1(2012)].
- [7] V. M. Braun, *Light cone sum rules*, in *Progress in heavy quark physics. Proceedings, 4th International Workshop, Rostock, Germany, September 20-22, 1997*, pp. 105–118, 1997. [arXiv:hep-ph/9801222](#).
- [8] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **10** (2015) 172, [arXiv:1509.00771](#).
- [9] S. Braun *et al.*, *Measurement of the ratio $\mathcal{B}(B_s^0 \rightarrow K^- \mu^+ \nu) / \mathcal{B}(B_s^0 \rightarrow D_s^- \mu^+ \nu)$ and $|V_{ub}| / |V_{cb}|$* , URL: https://twiki.cern.ch/twiki/pub/LHCbPhysics/VubFromBs/Bs2KMuNu_anaNote_v1.pdf.
- [10] I. T. Smith, *Determining the weak couplings $|V_{ub}|$ and $|V_{cb}|$ using semileptonic decays*, PhD thesis, Edinburgh U., 2019.
- [11] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.
- [12] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, *Phys. Rev.* **D98** (2018) 030001.
- [13] E. Noether, *Invariant Variation Problems*, *Gott. Nachr.* **1918** (1918) 235, [arXiv:physics/0503066](#).
- [14] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays*, *Phys. Rev. Lett.* **115** (2015) 072001, [arXiv:1507.03414](#).

- [15] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [16] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [17] Super-Kamiokande, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562, [arXiv:hep-ex/9807003](#).
- [18] SNO, Q. R. Ahmad *et al.*, *Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001) 071301, [arXiv:nucl-ex/0106015](#).
- [19] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, *Sov. Phys. JETP* **7** (1958) 172.
- [20] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870.
- [21] M. Drewes, *The Phenomenology of Right Handed Neutrinos*, *Int. J. Mod. Phys. E* **22** (2013) 1330019, [arXiv:1303.6912](#).
- [22] A. N. Taylor *et al.*, *Gravitational lens magnification and the mass of abell 1689*, *Astrophys. J.* **501** (1998) 539, [arXiv:astro-ph/9801158](#).
- [23] V. C. Rubin, N. Thonnard, and J. Ford, W.Ā. *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/*, *Astrophys. J.* **238** (1980) 471.
- [24] K. C. Freeman, *On the disks of spiral and SO Galaxies*, *Astrophys. J.* **160** (1970) 811.
- [25] J. R. Ellis *et al.*, *Supersymmetric Relics from the Big Bang*, *Nucl. Phys. B* **238** (1984) 453.
- [26] G. R. Farrar and M. E. Shaposhnikov, *Baryon asymmetry of the universe in the standard electroweak theory*, *Phys. Rev.* **D50** (1994) 774, [arXiv:hep-ph/9305275](#).
- [27] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, *Phys. Rev.* **D51** (1995) 379, [arXiv:hep-ph/9404302](#).
- [28] S. P. Martin, *A Supersymmetry primer*, [arXiv:hep-ph/9709356](#)[arXiv:hep-ph/9709356](#), [Adv. Ser. Direct. High Energy Phys.21,1(2010); Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [29] M. Schmaltz and D. Tucker-Smith, *Little Higgs review*, *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229, [arXiv:hep-ph/0502182](#).
- [30] H. Georgi and S. L. Glashow, *Unity of all elementary-particle forces*, *Phys. Rev. Lett.* **32** (1974) 438.

- [31] XENON, E. Aprile *et al.*, *Physics reach of the XENON1T dark matter experiment*, *JCAP* **1604** (2016) 027, [arXiv:1512.07501](#).
- [32] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [33] CMS, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [34] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531.
- [35] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.
- [36] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **53** (1984) 1802.
- [37] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51** (1983) 1945.
- [38] M. Gupta and G. Ahuja, *Flavor mixings and textures of the fermion mass matrices*, *Int. J. Mod. Phys. A* **27** (2012) 1230033, [arXiv:1302.4823](#).
- [39] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation*, *Phys. Rev. Lett.* **55** (1985) 1039.
- [40] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, *Eur. Phys. J.* **C41** (2005), no. 11, [arXiv:hep-ph/0406184](#), updated results and plots available at <http://ckmfitter.in2p3.fr/>.
- [41] 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 Δm_s and $BR(B \rightarrow \tau\nu_\tau)$* , *JHEP* **10** (2006) 081, [arXiv:hep-ph/0606167](#), updated results and plots available at <http://www.utfit.org/>.
- [42] C. A. Dominguez, *Introduction to QCD sum rules*, *Mod. Phys. Lett.* **A28** (2013) 1360002, [arXiv:1305.7047](#).
- [43] K. G. Wilson, *Confinement of quarks*, *Phys. Rev. D* **10** (1974) 2445.
- [44] C. Gattringer and C. B. Lang, *Quantum chromodynamics on the lattice*, vol. 788, Springer, Berlin, 2010. doi: 10.1007/978-3-642-01850-3.
- [45] P. B. Mackenzie, *An Improved Hybrid Monte Carlo Method*, *Phys. Lett.* **B226** (1989) 369.

- [46] J. Dingfelder and T. Mannel, *Leptonic and semileptonic decays of B mesons*, *Rev. Mod. Phys.* **88** (2016), no. 3 035008.
- [47] E. McLean, C. T. H. Davies, J. Koponen, and A. T. Lytle, *$B_s \rightarrow D_s \ell \nu$ Form Factors for the full q^2 range from Lattice QCD with non-perturbatively normalized currents*, [arXiv:1906.00701](https://arxiv.org/abs/1906.00701).
- [48] C. Bourrely, L. Lellouch, and I. Caprini, *Model-independent description of $b \rightarrow \pi \ell \nu$ decays and a determination of $|V_{ub}|$* , *Phys. Rev. D* **79** (2009) 013008.
- [49] C. M. Bouchard *et al.*, *$B_s \rightarrow K \ell \nu$ form factors from lattice QCD*, *Phys. Rev. D* **90** (2014) 054506, [arXiv:1406.2279](https://arxiv.org/abs/1406.2279).
- [50] J. M. Flynn *et al.*, *$B \rightarrow \pi \ell \nu$ and $B_s \rightarrow K \ell \nu$ form factors and $|V_{ub}|$ from 2+1-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks*, *Phys. Rev. D* **91** (2015), no. 7 074510, [arXiv:1501.05373](https://arxiv.org/abs/1501.05373).
- [51] Fermilab Lattice, MILC, A. Bazavov *et al.*, *$B_s \rightarrow K \ell \nu$ decay from lattice QCD*, *Phys. Rev. D* **100** (2019), no. 3 034501, [arXiv:1901.02561](https://arxiv.org/abs/1901.02561).
- [52] A. Khodjamirian and A. V. Rusov, *$B_s \rightarrow K \ell \nu_\ell$ and $B_{(s)} \rightarrow \pi(K) \ell^+ \ell^-$ decays at large recoil and CKM matrix elements*, *JHEP* **08** (2017) 112, [arXiv:1703.04765](https://arxiv.org/abs/1703.04765).
- [53] Flavour Lattice Averaging Group, S. Aoki *et al.*, *FLAG Review 2019*, *Eur. Phys. J. C* **80** (2020), no. 2 113, [arXiv:1902.08191](https://arxiv.org/abs/1902.08191).
- [54] ALPHA, F. Bahr *et al.*, *Continuum limit of the leading-order HQET form factor in $B_s \rightarrow K \ell \nu$ decays*, *Phys. Lett. B* **757** (2016) 473, [arXiv:1601.04277](https://arxiv.org/abs/1601.04277).
- [55] G. Duplancić and B. c. v. Melić, *$b, B_s \rightarrow k$ form factors: An update of light-cone sum rule results*, *Phys. Rev. D* **78** (2008) 054015.
- [56] W.-F. Wang and Z.-J. Xiao, *Semileptonic decays $b/B_s \rightarrow (\pi, k)(l^+ l^-, \ell \nu, \nu \bar{\nu})$ in the perturbative qcd approach beyond the leading order*, *Phys. Rev. D* **86** (2012) 114025.
- [57] R. N. Faustov and V. O. Galkin, *Charmless weak B_s decays in the relativistic quark model*, *Phys. Rev. D* **87** (2013) 094028.
- [58] J. A. Bailey *et al.*, *$B_s \rightarrow D_s/B \rightarrow D$ Semileptonic Form-Factor Ratios and Their Application to $BR(B_s^0 \rightarrow \mu^+ \mu^-)$* , *Phys. Rev. D* **85** (2012) 114502, [arXiv:1202.6346](https://arxiv.org/abs/1202.6346), [Erratum: *Phys. Rev. D* **86**, 039904(2012)].
- [59] C. J. Monahan *et al.*, *$B_s \rightarrow D_s \ell \nu$ Form Factors and the Fragmentation Fraction Ratio f_s/f_d* , *Phys. Rev. D* **95** (2017), no. 11 114506, [arXiv:1703.09728](https://arxiv.org/abs/1703.09728).
- [60] HFLAV, Y. S. Amhis *et al.*, *Averages of b -hadron, c -hadron, and τ -lepton properties as of 2018*, [arXiv:1909.12524](https://arxiv.org/abs/1909.12524), updated results and plots available at <https://hflav.web.cern.ch/>.

- [61] LHCb Collaboration, R. Aaij *et al.*, *Measurement of $|V_{cb}|$ with $B_s^0 \rightarrow D_s^{(*)-} \mu^+ \nu_\mu$ decays*, *Phys. Rev. D* **101** (2020) 072004, [arXiv:2001.03225](#).
- [62] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004. doi: [10.5170/CERN-2004-003-V-1](#).
- [63] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [64] LHCb, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [65] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.
- [66] R. Lindner, *LHCb layout 2/ LHCb schema 2*, LHCb Collection., URL: <http://cds.cern.ch/record/1087860>, Feb, 2008.
- [67] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, URL: https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html, last visited on 16/03/2020.
- [68] LHCb collaboration, *LHCb webpage*, URL: https://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html, last visited on 17/03/2020.
- [69] LHCb collaboration, *LHCb VELO (VERTex LOcator): Technical Design Report*, CERN-LHCC-2001-011. LHCb-TDR-005.
- [70] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, *JINST* **9** (2014) P09007, [arXiv:1405.7808](#).
- [71] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. LHCb-TDR-006.
- [72] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. LHCb-TDR-008.
- [73] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. LHCb-TDR-001.
- [74] LHCb collaboration, *LHCb public webpage*, URL: <https://lhcb-public.web.cern.ch/lhcb-public/en/Detector/Trackers2-en.html>, last visited on 18/03/2020.
- [75] LHCb collaboration, *LHCb Silicon Tracker webpage*, URL: <https://lhcb.physik.uzh.ch/ST/public/material/index.php>, last visited on 18/03/2020.
- [76] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, *JINST* **9** (2014) P01002, [arXiv:1311.3893](#).

- [77] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, *Int. J. Mod. Phys. A* **30** (2015) 1530022, [arXiv:1412.6352](#).
- [78] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J. C* **73** (2013) 2431, [arXiv:1211.6759](#).
- [79] LHCb RICH, A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, *Nucl. Instrum. Meth. A* **876** (2017) 221, [arXiv:1703.08152](#).
- [80] J. R. Harrison, *Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays*, PhD thesis, Manchester U., 2014.
- [81] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, *JINST* **8** (2013) P02022, [arXiv:1211.1346](#).
- [82] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036. LHCb-TDR-002.
- [83] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCb-TDR-004.
- [84] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, *JINST* **8** (2013) P10020, [arXiv:1306.0249](#).
- [85] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031. LHCb-TDR-010.
- [86] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, *JINST* **8** (2013) P04022, [arXiv:1211.3055](#).
- [87] A. Puig, *The LHCb trigger in 2011 and 2012*, LHCb-PUB-2014-046.
- [88] LHCb collaboration, *LHCb Trigger twiki pages*, URL: <https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbTrigger>, last visited on 27/03/2020.
- [89] LHCb collaboration, *The Brunel project*, URL: <http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/>.
- [90] LHCb collaboration, *The Gaudi project*, URL: <http://gaudi.web.cern.ch/gaudi/>.
- [91] LHCb collaboration, *The DaVinci project*, URL: <http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/>.
- [92] LHCb collaboration, *The Gauss project*, URL: <http://lhcbdoc.web.cern.ch/lhcbdoc/gauss/>.

- [93] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [94] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [95] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [96] 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](#).
- [97] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [98] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [99] LHCb collaboration, *The Boole project*, URL: <http://lhcbdoc.web.cern.ch/lhcbdoc/boole/>.
- [100] LHCb collaboration, *The Moore project*, URL: <http://lhcbdoc.web.cern.ch/lhcbdoc/moore/>.
- [101] LHCb collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, [arXiv:1111.2357](#).
- [102] W. L. Sutcliffe, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, PhD thesis, Imperial Coll., London, 2016-09-20.
- [103] E. McLean, C. T. H. Davies, A. T. Lytle, and J. Koponen, *Lattice QCD form factor for $B_s \rightarrow D_s^* l \nu$ at zero recoil with non-perturbative current renormalisation*, Phys. Rev. **D99** (2019), no. 11 114512, [arXiv:1904.02046](#).
- [104] M. Atoui, V. Morénas, D. Bećirevic, and F. Sanfilippo, *$B_s \rightarrow D_s l \nu_\ell$ near zero recoil in and beyond the Standard Model*, Eur. Phys. J. **C74** (2014), no. 5 2861, [arXiv:1310.5238](#).
- [105] 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 .
- [106] 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](#).

- [107] 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](#).
- [108] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in python*, *CoRR abs/1201.0490* (2012) [arXiv:1201.0490](#).
- [109] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, *Nucl. Instrum. Meth.* **A555** (2005) 356, [arXiv:physics/0402083](#).
- [110] 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.
- [111] R. Kohavi, *A study of cross-validation and bootstrap for accuracy estimation and model selection*, 1995.
- [112] O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, .
- [113] J. Albrecht *et al.*, *Search for Lepton Flavor Violation and Baryon Number Violation in τ decays*, [LHCb-ANA-2012-095](#).
- [114] C. Bozzi *et al.*, *Measurement of the mixing frequency Δm_d using semileptonic B^0 decays*, .
- [115] F. Archilli *et al.*, *Search for the $B^0 \rightarrow \mu^+ \mu^-$ decay and measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction and effective lifetime*, .
- [116] LHCb collaboration, R. Aaij *et al.*, *A precise measurement of the B^0 meson oscillation frequency*, *Eur. Phys. J.* **C76** (2016) 412, [arXiv:1604.03475](#).
- [117] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, *J. Phys. Conf. Ser.* **762** (2016), no. 1 012036, [arXiv:1608.05806](#).
- [118] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, *Comput. Phys. Commun.* **77** (1993) 219.
- [119] J. S. Conway, *Incorporating nuisance parameters in likelihoods for multisource spectra*, 2011.
- [120] F. James and M. Roos, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, *Computer Physics Communications* **10** (1975), no. 6 343–367.
- [121] 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.
- [122] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.* **A389** (1997) 81.

- [123] LHCb collaboration, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, *Nature Physics* **11** (2015) 743, [arXiv:1504.01568](#).
- [124] 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.
- [125] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.
- [126] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, *JINST* **10** (2015), no. 02 P02007, [arXiv:1408.1251](#).
- [127] LHCb collaboration, *LHCb Tracking twiki page*, URL: <https://twiki.cern.ch/twiki/bin/view/LHCbInternal/TrackingEffStatus2012S20>, last visited on 27/03/2020.
- [128] W. Detmold, C. Lehner, and S. Meinel, $\Lambda_b \rightarrow p\ell^-\bar{\nu}_\ell$ and $\Lambda_b \rightarrow \Lambda_c\ell^-\bar{\nu}_\ell$ form factors from lattice QCD with relativistic heavy quarks, *Phys. Rev.* **D92** (2015), no. 3 034503, [arXiv:1503.01421](#).
- [129] Flavour Lattice Averaging Group, S. Aoki *et al.*, *FLAG Review 2019*, [arXiv:1902.08191](#).
- [130] D. Müller, M. Clemencic, G. Corti, and M. Gersabeck, *ReDecay: A novel approach to speed up the simulation at LHCb*, *Eur. Phys. J.* **C78** (2018), no. 12 1009, [arXiv:1810.10362](#).
- [131] LHCb, M. P. Whitehead, *A palette of fast simulations in LHCb*, *PoS ICHEP2018* (2019) 271.
- [132] M. Calvi *et al.*, *Emulation of the hadronic trigger for tracker-only simulation*, Tech. Rep. LHCb-INT-2019-025. CERN-LHCb-INT-2019-025, CERN, Geneva, Dec, 2019.
- [133] MILC, J. A. Bailey *et al.*, $B \rightarrow l \nu$ form factors at nonzero recoil and $-V_{cb}$ from 2+1-flavor lattice QCD, *Phys. Rev. D* **92** (2015), no. 3 034506, [arXiv:1503.07237](#).