

# Angular Analysis of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ Decays with the LHCb Experiment

Von der Fakultät für Mathematik, Informatik und  
Naturwissenschaften der RWTH Aachen University zur  
Erlangung des akademischen Grades eines Doktors der  
Naturwissenschaften genehmigte Dissertation

vorgelegt von

MARCEL MATEROK, M. SC. RWTH AACHEN UNIVERSITY

aus

RECKLINGHAUSEN

Berichter: DR. CHRISTOPH LANGENBRUCH  
PROF. DR. MICHAEL KRÄMER

Tag der mündlichen Prüfung: 05.12.2022

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.





## Abstract

This thesis presents an angular analysis of the rare decay  $B_s^0 \rightarrow (\phi \rightarrow K^+ K^-) \mu^+ \mu^-$ . The full data set taken at the LHCb experiment during Runs 1 and 2 of the LHC is used to perform this analysis. The angular observables, consisting of four  $CP$ -averages and four  $CP$ -asymmetries, are extracted in four-dimensional maximum likelihood fits to the three decay angles and the reconstructed mass of the  $B_s^0$  candidate in regions of the squared invariant dimuon mass ( $q^2$ ). The angular observables are locally consistent with the Standard Model (SM) prediction at the level of 2 standard deviations ( $\sigma$ ) in each of the  $q^2$  regions. This measurement constitutes the most precise determination of the angular observables to date and supersedes the previous measurement performed by the LHCb experiment.

The consistency of the measured  $CP$ -averages with the theory prediction is tested by determining a shift in the muonic vector coupling strength ( $\Delta\mathcal{R}e(\mathcal{C}_9)$ ) from its SM value with a global fit. The data are found to be globally consistent with the SM prediction at the level of  $1.9\sigma$  but prefer a shift of the muonic vector coupling strength of  $\Delta\mathcal{R}e(\mathcal{C}_9) = -1.3^{+0.7}_{-0.6}$ . This deviation is consistent with shifts in other  $b \rightarrow s\ell^+\ell^-$  transitions.

## Kurzdarstellung

In dieser Arbeit wird die Winkelanalyse des seltenen Zerfalls  $B_s^0 \rightarrow (\phi \rightarrow K^+ K^-) \mu^+ \mu^-$  präsentiert. Für die Analyse wurde das vom LHCb Experiment während der Läufe 1 und 2 des LHC aufgezeichnete Datenset benutzt. Die Winkelobservablen, bestehend aus vier  $CP$ -Mittelwerten und vier  $CP$ -Asymmetrien, wurden in einer vier dimensionalen Maximum Likelihood Anpassung an die drei Zerfallswinkel und die rekonstruierte Masse des  $B_s^0$  Kandidaten in Regionen der quadrierten invarianten Masse des Muonenpaares ( $q^2$ ) durchgeführt. Die Winkelobservablen stimmen lokal mit der Standardmodell (SM) Vorhersage auf dem Level von 2 Standard Abweichungen ( $\sigma$ ) in jeder der  $q^2$  Regionen überein. Diese Messung ist die genaueste Bestimmung der Winkelobservablen bisher und ersetzt die vorherige, am LHCb Experiment durchgeführte Messung.

Die Übereinstimmung der gemessenen  $CP$ -Mittelwerte mit der Theorie Vorhersage wird mit der Bestimmung einer Verschiebung der muonischen Vektor-Kopplungsstärke ( $\Delta\mathcal{R}e(\mathcal{C}_9)$ ) von seinem SM Wert mittels einer globalen Anpassung getestet. Die Daten stimmen global mit der SM Vorhersage auf dem Level von  $1.9\sigma$  überein, bevorzugen aber eine Verschiebung der muonischen Vektor-Kopplungsstärke von  $\Delta\mathcal{R}e(\mathcal{C}_9) = -1.3^{+0.7}_{-0.6}$ . Diese Abweichung ist konsistent mit Diskrepanzen in anderen  $b \rightarrow s\ell^+\ell^-$  Übergängen.





# Table of Contents

|          |                                                                                        |           |
|----------|----------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                                                    | <b>1</b>  |
| <b>2</b> | <b>Theory</b>                                                                          | <b>3</b>  |
| 2.1      | The Standard Model of particle physics . . . . .                                       | 3         |
| 2.1.1    | Gauge theories . . . . .                                                               | 3         |
| 2.1.2    | Elementary particle content of the SM . . . . .                                        | 4         |
| 2.1.3    | Quantum chromodynamics . . . . .                                                       | 5         |
| 2.1.4    | Electroweak interaction . . . . .                                                      | 6         |
| 2.1.5    | Spontaneous symmetry breaking and Yukawa coupling . . . . .                            | 8         |
| 2.1.6    | The CKM mechanism . . . . .                                                            | 9         |
| 2.2      | Physics of the $B_s^0$ meson . . . . .                                                 | 10        |
| 2.2.1    | Production of $b\bar{b}$ pairs at LHCb . . . . .                                       | 10        |
| 2.2.2    | $B_s^0$ meson mixing and time-evolution . . . . .                                      | 11        |
| 2.3      | The rare decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ . . . . .                          | 15        |
| 2.3.1    | Effective field theory for $b \rightarrow s \ell^+ \ell^-$ transitions . . . . .       | 15        |
| 2.3.2    | Form factors for the $B_s^0 \rightarrow \phi$ transition . . . . .                     | 17        |
| 2.3.3    | Transversity amplitudes . . . . .                                                      | 17        |
| 2.3.4    | Differential decay rate . . . . .                                                      | 19        |
| <b>3</b> | <b>The LHCb experiment</b>                                                             | <b>23</b> |
| 3.1      | The Large Hadron Collider . . . . .                                                    | 23        |
| 3.2      | The LHCb detector . . . . .                                                            | 25        |
| 3.2.1    | Luminosity at LHCb . . . . .                                                           | 26        |
| 3.2.2    | Tracking at LHCb . . . . .                                                             | 27        |
| 3.2.3    | Ring Imaging Cherenkov Detectors . . . . .                                             | 33        |
| 3.2.4    | Calorimeters . . . . .                                                                 | 34        |
| 3.2.5    | Muon chambers . . . . .                                                                | 36        |
| 3.2.6    | Trigger . . . . .                                                                      | 37        |
| <b>4</b> | <b>Analysis tools</b>                                                                  | <b>40</b> |
| 4.1      | Multivariate analysis . . . . .                                                        | 40        |
| 4.2      | Maximum likelihood technique for parameter estimation . . . . .                        | 42        |
| 4.3      | Statistically subtracting backgrounds with the $_s\mathcal{P}lot$ -technique . . . . . | 43        |
| <b>5</b> | <b>Selection of data and simulation samples</b>                                        | <b>45</b> |
| 5.1      | Data samples . . . . .                                                                 | 45        |
| 5.2      | Samples of simulated events . . . . .                                                  | 46        |

|          |                                                                                                           |           |
|----------|-----------------------------------------------------------------------------------------------------------|-----------|
| 5.3      | Partitioning in $q^2$ regions . . . . .                                                                   | 48        |
| 5.4      | Event selection . . . . .                                                                                 | 48        |
| 5.4.1    | Definition of important quantities of particle candidates . . . . .                                       | 49        |
| 5.4.2    | Trigger . . . . .                                                                                         | 51        |
| 5.4.3    | Truth-matching of simulated events . . . . .                                                              | 52        |
| 5.4.4    | LHCb wide preselection . . . . .                                                                          | 52        |
| 5.4.5    | Analysis specific preselection . . . . .                                                                  | 54        |
| 5.4.6    | BDT . . . . .                                                                                             | 54        |
| 5.4.7    | Multiple candidates . . . . .                                                                             | 56        |
| 5.5      | Corrections to simulation . . . . .                                                                       | 60        |
| 5.5.1    | Reweighting of the track multiplicity . . . . .                                                           | 60        |
| 5.5.2    | Corrections to the $B_s^0$ kinematics . . . . .                                                           | 61        |
| 5.5.3    | Correction to the L0 trigger . . . . .                                                                    | 62        |
| 5.5.4    | Correction of the PID response using MEERKAT . . . . .                                                    | 67        |
| 5.5.5    | Comparison of control mode data and simulation . . . . .                                                  | 68        |
| <b>6</b> | <b>Signal and background yields</b>                                                                       | <b>70</b> |
| 6.1      | Considered decay modes . . . . .                                                                          | 71        |
| 6.2      | Determination of selection efficiencies . . . . .                                                         | 72        |
| 6.3      | Expected signal and background yields . . . . .                                                           | 74        |
| <b>7</b> | <b>Angular analysis of the decay <math>B_s^0 \rightarrow \phi \mu^+ \mu^-</math></b>                      | <b>75</b> |
| 7.1      | Angular acceptance . . . . .                                                                              | 75        |
| 7.2      | Fit model . . . . .                                                                                       | 79        |
| 7.2.1    | Fit model parameter determination . . . . .                                                               | 80        |
| 7.2.2    | Implementation of the angular acceptance in the likelihood . . . . .                                      | 81        |
| 7.3      | Fit validation . . . . .                                                                                  | 82        |
| 7.3.1    | Fit validation using simulation . . . . .                                                                 | 83        |
| 7.3.2    | Fit validation using control mode data . . . . .                                                          | 85        |
| 7.3.2.1  | Results from control model fits . . . . .                                                                 | 85        |
| 7.3.2.2  | Comparison with dedicated $B_s^0 \rightarrow J/\psi \phi$ analyses . . . . .                              | 88        |
| 7.3.3    | Fit validation using pseudoexperiments . . . . .                                                          | 91        |
| <b>8</b> | <b>Systematic uncertainties</b>                                                                           | <b>94</b> |
| 8.1      | Approach to the determination of systematic uncertainties . . . . .                                       | 94        |
| 8.2      | Peaking backgrounds . . . . .                                                                             | 95        |
| 8.3      | Acceptance orders . . . . .                                                                               | 96        |
| 8.4      | Effect of neglecting time-dependent efficiencies . . . . .                                                | 96        |
| 8.5      | Effect of using simulated $B_s^0 \rightarrow \phi \mu^+ \mu^-$ events with $\Delta\Gamma_s = 0$ . . . . . | 97        |
| 8.6      | Corrections to simulation based on weights . . . . .                                                      | 98        |
| 8.7      | Correction of the PID response . . . . .                                                                  | 99        |
| 8.8      | Limited size of simulated samples for acceptance . . . . .                                                | 100       |
| 8.9      | Signal mass model . . . . .                                                                               | 100       |
| 8.10     | Background modeling . . . . .                                                                             | 101       |
| 8.11     | Shared angular background parameters . . . . .                                                            | 101       |
| 8.12     | Acceptance evaluation point in $q^2$ . . . . .                                                            | 102       |
| 8.13     | Neglecting $S$ -wave contribution . . . . .                                                               | 103       |

|           |                                                                      |            |
|-----------|----------------------------------------------------------------------|------------|
| 8.14      | Remaining fit bias . . . . .                                         | 103        |
| 8.15      | Combination of systematic uncertainties . . . . .                    | 104        |
| <b>9</b>  | <b>Results for the angular observables</b>                           | <b>109</b> |
| 9.1       | Results for angular observables . . . . .                            | 109        |
| 9.2       | Correlation matrices . . . . .                                       | 112        |
| 9.3       | Compatibility of the $CP$ -averages with the SM . . . . .            | 115        |
| <b>10</b> | <b>Summary and outlook</b>                                           | <b>117</b> |
| <b>A</b>  | <b>Definitions for additional time-dependent angular observables</b> | <b>120</b> |
| <b>B</b>  | <b>Comparison of control mode data and simulation for all years</b>  | <b>122</b> |
| <b>C</b>  | <b>Determining amplitudes from angular observables</b>               | <b>131</b> |
| <b>D</b>  | <b>Distributions of the pulls from the fit validation</b>            | <b>133</b> |
| <b>E</b>  | <b>Alternative tables for the systematic uncertainties</b>           | <b>137</b> |
| <b>F</b>  | <b>Projections of fits to rare mode data</b>                         | <b>141</b> |



# Chapter 1

## Introduction

Particle physics aims to describe the fundamental building blocks of matter, the elementary particles, and the interaction between particles. The most successful description is a quantum field theory commonly referred to as the Standard Model of particle physics (SM) [1–3] which was developed during the 1960’s and finalised in the 1970’s. The SM withstood nearly all tests and is arguably one of the most successful theories to date. However, in its current form the SM fails to describe or give an explanation for astronomical observations such as the matter-antimatter asymmetry in the universe [4] or the measurements of the cosmic microwave background [5–8], the latter of which could be explained with a cold dark matter cosmology model [9, 10]. Therefore, New Physics (NP), an extension beyond the SM, is needed to explain these observations. At least one new particle would be required to explain dark matter in the context of particle physics and the matter-antimatter asymmetry requires the simultaneous violation of the charge and parity ( $CP$ ) operators. While  $CP$ -violation is part of the SM, it is not sizeable enough to account the the measured matter-antimatter asymmetry.

During the last decade tensions have arisen in the field of  $B$  hadron physics, for example in transitions of a  $b$  quark to an  $s$  quark and a pair of leptons ( $b \rightarrow s\ell^+\ell^-$ ), where measurements do not align perfectly with the SM prediction [11–32]. These tensions in  $b \rightarrow s\ell^+\ell^-$  transitions are known as the flavour anomalies. The  $b \rightarrow s\ell^+\ell^-$  transitions can only occur at the loop level in the SM and as a consequence they are potentially sensitive to additional, heavy mediators not described in the SM. In these loop level transitions, any NP effect can be probed indirectly as the new mediator and be heavier than the decaying  $b$  quark and, therefore, would not need to be produced. The indirect approach of probing NP has the advantage of not being limited by the available centre-of-mass energy and can, therefore, reach higher energy scales for NP scenarios than direct searches. The discrepancies in  $b \rightarrow s\ell^+\ell^-$  transitions have been seen at different experiments [33–38] across various decay modes and analysis methods, such as lepton flavour universality tests [12, 18–23, 25, 29–32], angular analyses [11, 14–16, 21, 22, 26–29] and branching fraction measurements [11, 13–15, 17, 21, 22, 24]. The various measurements seem to imply a deviation of the coupling structure of the muon with respect to the SM prediction [39–41] which is interesting in the context of tensions of the muon anomalous magnetic moment ( $g-2$ ) [42–45]. One of the most significant deviations in  $b \rightarrow s\ell^+\ell^-$  transitions from the SM expectation has been observed in the branching fraction measurement of  $B_s^0 \rightarrow \phi\mu^+\mu^-$  decays with 3.6 standard deviations ( $\sigma$ ) [24]. An angular analysis of this decay mode provides complementary information to the branching fraction measurement and is able to probe the muonic coupling structure and, therefore, the operator structure of potential NP effects in detail. At the LHCb experiment,

which was designed to measure rare heavy flavour decays and to probe  $CP$ -violation, the Run 1 data has already been used to perform an angular analysis which was compatible with the SM [11,15]. So far, the LHCb experiment is the only experiment to perform an angular analysis of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays.

In this thesis, an angular analysis of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays is performed with data taken during Run 1 and Run 2 of the Large Hadron Collider by the LHCb experiment. The addition of the Run 2 data set to the Run 1 data set represents a data set approximately four times the size of the Run 1 sample corresponding to an expected increase of a factor two in the statistical power. Therefore, it is expected to be the most precise measurement of the angular observables in the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$ .

This thesis is structured as follows: In Chap. 2 the theoretical framework is introduced, while the experimental framework is introduced in Chap. 3. The analysis tools used throughout this thesis are described in Chap. 4. The selection of data and simulated samples are described in Chap. 5 and the expected signal and background yields are estimated in Chap. 6. The details on the angular analysis are then given in Chap. 7 and in Chap. 8 the systematic uncertainties relevant to this analysis are described. The results of the measurement of the angular observables are given in Chap. 9, and this thesis is summarised in Chap. 10 and an outlook for future angular analyses of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays at LHCb is given.

The work presented in this thesis has been published by the LHCb collaboration in Ref. [46] and is also presented as a pre-print on arXiv under the identifier 2107.13428. The results have also been shown virtually to the public for the first time by the author of this thesis at the FPCP conference in 2021. Furthermore, the selection and correction to the simulation have been developed in cooperation with Sophie Kretzschmar and is, therefore, identical to that used in Ref. [24].

# Chapter 2

## Theory

The aim of this Chapter is to give an overview of the Standard Model of particle physics (SM), the physics of a  $B_s^0$  meson and of the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$ . The Standard Model is introduced in Sec. 2.1 with gauge theories being introduced based on Refs. [47–49] in Sec. 2.1.1. The mathematical formalism used in the SM is given in Secs. 2.1.3–2.1.5 which is largely based on Ref. [50]. In Sec. 2.2 the physics of the  $B_s^0$  meson is discussed. The production of  $B_s^0$  mesons at LHCb is covered in Sec. 2.2.1, whereas the  $B_s^0$  meson mixing and its time evolution is outlined in Sec. 2.2.2. The latter Section is based on Refs. [51, 52] and the formulae derived in this Section are generally also true for other mesons such as the  $K^0$ ,  $D^0$  and  $B^0$  mesons. The theory of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay is discussed in Sec. 2.3. There, the effective field theory description for the quark-level  $b \rightarrow s \ell^+ \ell^-$  transition, form factors describing the hadronic transition of a  $B_s^0$  meson to a  $\phi$  meson, transversity amplitudes, and the differential decay rate are covered. This Section is largely based on Refs. [53–57].

### 2.1 The Standard Model of particle physics

There are four known fundamental forces in physics, namely the electromagnetic, the weak nuclear, the strong nuclear and the gravitational force. The Standard Model of particle physics (SM) [1–3] is a theory to describe all of these fundamental forces except gravity, and additionally all elementary particles and their properties. It is a quantum field theory describing the forces as gauge fields and the elementary particles as excitation of fields. The mathematical description of the SM was finalised in the 1970s and with the discovery of the Higgs boson [58–60] in a joint effort of the ATLAS and CMS collaborations in 2012 [61], all predicted elementary particles of the SM have been confirmed. An overview of how the SM was developed is given in Ref. [62].

#### 2.1.1 Gauge theories

The SM is a gauge theory that is locally invariant under transformation involving the Lie groups [63]  $U(1)$ ,  $SU(2)$  and  $SU(3)$ . Furthermore, it describes the gauge fields as relativistic, spin-dependent and quantised solutions to the Lagrange density (Lagrangian,  $\mathcal{L}$ ). The Lagrangian of the SM is a function of all fields that exist in the theory and the equations of

motion for a given field,  $\Psi$ , are given by the Euler-Lagrange equation, which has the form

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} = \frac{\partial \mathcal{L}}{\partial \Psi}.$$

While the complete Lagrangian needs to be invariant under any transformation under which the SM is invariant, any single field is not required to be invariant. If the field  $\Psi$  is not invariant under a local U(1) transformation such as

$$\Psi \mapsto \Psi' = e^{i\theta(x)}\Psi,$$

where  $\theta(x)$  is a phase that depends on the space-time coordinates, its derivative  $\partial_\mu \Psi$  is also not invariant under this transformation and is given as

$$\partial_\mu \Psi' = \partial_\mu \Psi + (i\partial_\mu \theta(x))\Psi.$$

However, as a consequence of the SM being invariant under this type of transformation, the complete Lagrangian has to be invariant under this transformation and another field  $\phi$  has to exist and absorb the variant part of the field  $\Psi$ . To achieve this, the gauge field  $\Phi$  has to transform as

$$\Phi \mapsto \Phi' = \Phi - \partial_\mu \theta(x).$$

That way, the term  $(\partial_\mu + i\Phi)$  can be used as the gauge covariant derivative  $D_\mu$  which is invariant under the local transformation. In different parts of the SM Lagrangian different gauge covariant derivatives will be used.

### 2.1.2 Elementary particle content of the SM

In the SM there are two types of particles, bosons and fermions. Their distinguishing property is their spin quantum numbers, as the spin of bosons only has integer values, while fermions have an odd half-integer spin values. The elementary bosons and fermions described in the SM are listed with their masses and electrical charges in Tab. 2.1.

While both, bosons and fermions, are excitations of a corresponding field, only the elementary bosons are mediators of a force. The only exception to this is the Higgs boson, which is not directly linked to a force but is the excitation of the Higgs field which is required to introduce mass terms to all particles. All other elementary bosons are gauge bosons mediating the forces described in the SM. The photon is the mediator of the electromagnetic force and couples proportionally to the electric charge. The weak force is mediated by the  $W^\pm$  and  $Z^0$  bosons coupling to the weak isospin ( $T_3$ ). The last type of gauge boson are the eight gluons which are the mediators of the strong force. The charge of the strong force is called colour and can have three different values, commonly referred to as “red”, “green” and “blue”, as well as their anti-colour counterparts.

The fermions can be separated into leptons and quarks, with one of their distinguishing properties being the quarks ability to interact via the strong force while leptons can not interact strongly. Both types of fermions can further be subdivided into three generations each, which are ordered by their mass.<sup>1</sup> Each of the generations corresponds to a different flavour of fermion.

---

<sup>1</sup>The masses of the neutrinos are not measured and only upper limits exist. As such the mass hierarchy can potentially be different for neutrinos.



The flavour of a fermion is a quantum number that is conserved in all interactions except for weak interactions involving a  $W^\pm$  boson. One major difference between the two different kinds of fermions is their electric charge. Where leptons only have integer values of charge in units of the elementary charge  $e$ , namely  $-1e$  for leptons ( $\ell^-$ ) and  $0e$  for lepton neutrinos ( $\nu_\ell$ ), quarks have either a charge of  $+\frac{2}{3}e$  for up-type quarks ( $u$ ) or  $-\frac{1}{3}e$  for down-type quarks ( $d$ ). With the exception of neutrinos, quarks and leptons can be both, left- and right-handed. The neutrinos have so far only been observed to be left-handed [64].

Table 2.1: Elementary bosons and fermions, subdivided into quarks and leptons, described in the Standard Model of particle physics and their masses and electric charges, values taken from [65].

| Bosons              |  |                             |        |  |  |
|---------------------|--|-----------------------------|--------|--|--|
| Name                |  | Mass                        | Charge |  |  |
| Photon ( $\gamma$ ) |  | $< 10^{-18} \text{ eV}/c^2$ | $0 e$  |  |  |
| Gluon ( $g$ )       |  | $0 \text{ eV}/c^2$          | $0 e$  |  |  |
| $W$                 |  | $80.379 \text{ GeV}/c^2$    | $1 e$  |  |  |
| $Z$                 |  | $91.1876 \text{ GeV}/c^2$   | $0 e$  |  |  |
| Higgs ( $h$ )       |  | $125.25 \text{ GeV}/c^2$    | $0 e$  |  |  |

| Quarks          |                          |                  | Leptons                       |                           |        |
|-----------------|--------------------------|------------------|-------------------------------|---------------------------|--------|
| Name            | Mass                     | Charge           | Name                          | Mass                      | Charge |
| Up ( $u$ )      | $2.16 \text{ MeV}/c^2$   | $\frac{2}{3} e$  | Electron ( $e$ )              | $510.99 \text{ keV}/c^2$  | $-1 e$ |
| Down ( $d$ )    | $4.67 \text{ MeV}/c^2$   | $-\frac{1}{3} e$ | Electron neutrino ( $\nu_e$ ) | $< 1.1 \text{ eV}/c^2$    | $0 e$  |
| Charm ( $c$ )   | $1.27 \text{ GeV}/c^2$   | $\frac{2}{3} e$  | Muon ( $\mu$ )                | $105.66 \text{ MeV}/c^2$  | $-1 e$ |
| Strange ( $s$ ) | $93 \text{ MeV}/c^2$     | $-\frac{1}{3} e$ | Muon neutrino ( $\nu_\mu$ )   | $< 0.19 \text{ MeV}/c^2$  | $0 e$  |
| Top ( $t$ )     | $172.76 \text{ GeV}/c^2$ | $\frac{2}{3} e$  | Tau ( $\tau$ )                | $1776.86 \text{ MeV}/c^2$ | $-1 e$ |
| Bottom ( $b$ )  | $4.18 \text{ GeV}/c^2$   | $-\frac{1}{3} e$ | Tau neutrino ( $\nu_\tau$ )   | $< 18.2 \text{ MeV}/c^2$  | $0 e$  |

### 2.1.3 Quantum chromodynamics

In the SM, the strong force is described by a renormalisable gauge theory called Quantum Chromodynamics (QCD) based on the  $SU(3)$  group. It describes the strong interaction in which only colour-charged particles partake. These particles are the quarks and gluons. Due to the colour being the charge associated to the  $SU(3)$ , it is often also labeled as  $SU(3)_C$ .

The strong force has the unique dynamical properties of asymptotic freedom and confinement. Asymptotic freedom describes the phenomenon that the effective coupling ( $\alpha_s$ ) is a function of the transferred four momentum ( $q^2$ ), namely it decreases with increasing  $q^2$  and even vanishes asymptotically. This dependency can be described by

$$\alpha_s = \frac{1}{b} \frac{\log q^2}{\Lambda_{QCD}^2},$$

where  $b$  is a constant and  $\Lambda_{QCD}$  is the energy scale at which the perturbatively-defined coupling of QCD would diverge. The numerical value of  $\Lambda_{QCD}$  is of the order of a few hundred MeV, however, in literature it has been increasingly more popular to directly state the used value for  $\alpha_s$  as the value of  $\Lambda_{QCD}$  only defines  $\alpha_s$  if it is known which approximations are used [65].

The other dynamical property, confinement, describes the empirical phenomenon that no colored state has been observed until now. Therefore, it is assumed that only states which are colourless can be free. A colourless state can be achieved if there are equal amounts of all three colours or equal amounts of a colour and its anti-colour, or a combination of both former options. Therefore, a single quark or a gluon cannot be free in the SM. As a consequence, any quark produced in an interaction will hadronise into a bound state [66]. Confinement is encoded in the interaction potential of the effective coupling as a function of the distance ( $r$ ) between two coloured objects. For small distances the potential behaves like the Coulomb potential and is proportional to  $\frac{\alpha_s}{r}$ , whereas it increases linearly with  $r$  for large distances. If two colour charged quarks separate in space, for example if they are created with large energies and back-to-back kinematics in pair-production, the potential between them increases to the point at which a quark anti-quark pair ( $q\bar{q}$ ) can be created from the vacuum. At some point, the energy of the initially produced quarks reduces so much, that the produced quarks and anti-quarks will form bound hadron states and no additional quark anti-quark pairs will be produced.

#### 2.1.4 Electroweak interaction

The electromagnetic and weak forces are commonly described together in the SM as a unified interaction based on an  $U(1) \times SU(2)$  symmetry. These are the electromagnetic and weak forces, and the unified interaction is referred to as the electroweak interaction. The charges conserved in the  $U(1) \times SU(2)$  symmetries are the weak hypercharge ( $Y_W$ ) and the third component of the weak isospin ( $T_3$ ), respectively. These charges are connected via the electric charge ( $Q$ ) by the relation

$$Q = T_{L/R}^3 + \frac{Y_{L/R}}{2},$$

where  $L$  and  $R$  indicates the chirality (handedness) of the fermions. A right-handed fermion has a positive helicity and the direction of the spin and the direction of its motion are in the same direction, while a left-handed fermion has negative helicity and the direction of the spin and its motion are opposite to each other. While the electromagnetic force acts upon all charged particles, the weak force only acts on left-handed particles. Given the left-handed nature of the weak force and the weak hypercharge being the charge of the  $U(1)$  group, the symmetry group is often referred to as  $U(1)_Y \times SU(2)_L$ .

The fields relevant to the electroweak interaction are the gauge fields  $B_\mu$  and  $W_\mu^A$ , where  $A$  enumerates the three  $W$  fields, as well as the left- or right-handed fermion fields ( $\Psi_{L,R}$ ). Using these fields, one can write the electroweak gauge covariant derivative ( $D_\mu^{EW} \Psi_{L,R}$ ) as

$$D_\mu^{EW} \Psi_{L,R} = \left[ \partial_\mu + ig \sum_{A=1}^3 t_{L,R}^A W_\mu^A + ig' \frac{Y_{L,R}}{2} B_\mu \right] \Psi_{L,R},$$

where  $\frac{Y_{L,R}}{2}$  is a scalar and the generator of the  $U(1)_Y$ , and  $t_{L,R}^A$  are the Pauli matrices and the generators of the  $SU(2)_L$ . Furthermore,  $g$  and  $g'$  are the couplings of the gauge fields. The commutation relations of the Pauli matrices,  $t_{L,R}^A$ , are given by

$$[t_{L,R}^A, t_{L,R}^B] = i\epsilon_{ABC} t_{L,R}^C,$$

where  $\epsilon_{ABC}$  is the Levi-Civita tensor. Here, the normalisation condition  $Tr[t^A, t^B] = \frac{\delta^{AB}}{2}$  is used. Using the gauge covariant derivative, one can write the electroweak Lagrangian ( $\mathcal{L}_{EW}$ ) as

$$\mathcal{L}_{EW} = -\frac{1}{4} \sum_{A=1}^3 F_{\mu\nu}^A F^{A\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \bar{\Psi}_L i \not{D}^{EW} \Psi_L + \bar{\Psi}_R i \not{D}^{EW} \Psi_R,$$

where  $F_{\mu\nu}^A$  is an abbreviation and given by

$$F_{\mu\nu}^A = \partial_\mu W_\nu^A - \partial_\nu W_\mu^A - g\epsilon_{ABC} W_\mu^B W_\nu^C,$$

$B_{\mu\nu}$  is given by

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

and  $\not{D}^{EW}$  uses the Feynman slash notation and is given as

$$\not{D}^{EW} = D_\mu^{EW} \gamma^\mu,$$

where the  $\gamma_\mu$  are the Dirac matrices.

Even though there are four gauge fields entering the Lagrangian and four observable gauge bosons, there is no one-to-one correspondence between them. The physically observable  $W^\pm$  bosons are linear combination of the  $W_\mu^1$  and  $W_\mu^2$ , and the photon ( $A^\mu$  or  $\gamma$ ) and the  $Z^0$  boson are linear combinations of gauge fields,  $B_\mu$  and  $W_\mu^3$ . The  $W^\pm$  bosons are given by the combination

$$W_\mu^\pm = \frac{W_\mu^1 \pm iW_\mu^2}{\sqrt{2}}$$

and the coupling to the charged current can be written as

$$\begin{aligned} g(t^1 W_\mu^1 + t^2 W_\mu^2) &= g \left[ \frac{(t^1 + it^2)}{\sqrt{2}} \frac{(W_\mu^1 - iW_\mu^2)}{\sqrt{2}} + \frac{(t^1 - it^2)}{\sqrt{2}} \frac{(W_\mu^1 + iW_\mu^2)}{\sqrt{2}} \right] \\ &= g \left[ \frac{t^+ W_\mu^-}{\sqrt{2}} + \frac{t^- W_\mu^+}{\sqrt{2}} \right], \end{aligned}$$

where  $t^\pm = t^1 \pm it^2$ . In the low-energy limit, meaning that the squared mass of the  $W^\pm$  boson ( $m_W^2$ ) is much larger than the transferred four-momentum ( $q^2$ ) of an interaction, the coupling to the weak force can also be given in terms of Fermi's coupling constant  $G_W$  at tree-level as

$$G_W = \frac{\sqrt{2}g^2}{8m_W^2}.$$

The photon and the  $Z^0$  boson are described by the relations

$$\begin{aligned} A_\mu &= \cos\theta_W B_\mu + \sin\theta_W W_\mu^3 \\ Z_\mu &= -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3, \end{aligned}$$

where  $\theta_W$  is the weak mixing angle of the gauge fields (Weinberg angle). Furthermore, the coupling of the photon can be used to relate the couplings  $g$  and  $g'$  with the condition

$$g \sin\theta_W = g' \cos\theta_W = e$$

as the photon couples equally to left- and right-handed particles and the coupling strength is proportional to the electric charge of the particle in units of the elementary charge  $e$ . From this condition, the weak mixing angle can be fixed by the relation

$$\tan \theta_W = \frac{g'}{g}.$$

The elementary charge  $e$  is also connected to the fine-structure constant  $\alpha_e$  by the relation  $e^2 = 4\pi\epsilon_0\hbar c\alpha_e$ . Unlike the photon, the  $Z^0$  boson does not couple equally to left- and right-handed particles, but couples only to left-handed particles.

### 2.1.5 Spontaneous symmetry breaking and Yukawa coupling

The gauge bosons of the electroweak sector have to be massless in order to fulfill the local gauge invariance of the  $U(1)\times SU(2)$  gauge symmetry group [62]. However, in measurements the  $W^\pm$  and  $Z^0$  bosons are found to have mass [67–74]. While these masses can be inserted by hand, it would destroy the rationale of using a gauge theory [62]. The other way for the  $W^\pm$  and  $Z^0$  bosons to obtain mass requires the spontaneous breaking of the electroweak  $U(1)\times SU(2)$  symmetry group (Higgs mechanism). A vector of scalar fields ( $\phi$ ) is needed to achieve the spontaneous symmetry breaking and the potential of these fields ( $V(\phi^\dagger\phi)$ ) is required to be symmetric under the  $U(1)\times SU(2)$  symmetry group and to contain only terms that are quadratic or quartic terms in the field  $\phi$ . If there were higher order terms, the theory would no longer be renormalisable [75]. Therefore, the potential can be written as

$$V(\phi^\dagger\phi) = -\frac{1}{2}\mu^2\phi^\dagger\phi + \frac{1}{4}\lambda(\phi^\dagger\phi)^2,$$

where  $\mu$  is a parameter with the same dimension as a mass and  $\lambda$  is a dimensionless parameter. To realize the spontaneous symmetry breaking of the electroweak sector, the Higgs potential is required to have a minimum at a non-vanishing value for  $\phi$ , explicitly  $\phi \neq \vec{0}$ . The position of the minimum ( $v$ ) of the potential is referred to as the vacuum expectation value, defined as  $\langle 0|\phi(x)|0\rangle = v \neq 0$ . This can be achieved by choosing  $\mu^2$  and  $\lambda$  to simultaneously have positive, real values. The masses of the gauge bosons are then generated by the coupling to the Higgs field and are absorbed in the gauge covariant derivative for the Higgs field ( $D_\mu^{Higgs}\phi$ ). The latter is defined to be

$$D_\mu^{Higgs}\phi = \left[ \partial_\mu + ig \sum_{A=1}^3 t^A W_\mu^A + ig' \frac{Y}{2} B_\mu \right] \phi.$$

In a Higgs doublet model with

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \text{ and } v = \begin{pmatrix} 0 \\ v \end{pmatrix}$$

the  $W^\pm$  and  $Z^0$  masses can be computed with the relations

$$(2.1) \quad m_W^2 = \frac{g^2 v^2}{2} \text{ and } m_Z^2 = \frac{(g^2 + g'^2)v^2}{2} = \frac{g^2 v^2}{2m_W^2}.$$

Given that the coupling strength  $g$  and  $g'$  are connected by the Weinberg angle, the latter also connects the masses of the  $W^\pm$  and  $Z^0$  bosons by the relation

$$\cos \theta_W = \frac{m_W}{m_Z}.$$

With this definition and Eq. 2.1, one can calculate the vacuum expectation value to be [50]

$$v = \frac{1}{2^{\frac{3}{4}} \sqrt{G_W}} = 174.1 \text{ GeV}.$$

As the photon interacts only with charged particles and the Higgs boson does not have an electric charge, the photon does not acquire mass by the Higgs mechanism and is massless even in the spontaneously broken  $U(1) \times SU(2)$  symmetry group.

The gauge covariant derivative and the potential of the Higgs field enter the Lagrangian ( $\mathcal{L}_{\text{Higgs}}$ ) which is given as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu^{Higgs} \phi)^\dagger (D_\mu^{Higgs} \phi) - V(\phi^\dagger \phi) - \bar{\Psi}_L \Gamma \Psi_R \phi - \bar{\Psi}_R \Gamma \Psi_L \phi,$$

where  $\Psi_{L,R}$  are the fermion mass eigenstates,  $\Gamma$  are matrices that makes the Yukawa couplings of the fermions to the Higgs field invariant under the Lorentz and gauge groups. The Yukawa couplings of the fermions are responsible for the fermion masses. The mass matrix for the fermions ( $M$ ) can be obtained from the Yukawa couplings by replacing the scalar field  $\phi$  with the vacuum expectation value in the mass term of the Lagrangian. The fermion mass matrix is then given by

$$M = \bar{\Psi}_L \mathcal{M} \Psi_R + \bar{\Psi}_R \mathcal{M}^\dagger \Psi_L,$$

where the matrix  $\mathcal{M}$  is defined as  $\mathcal{M} = \Gamma v$ .

### 2.1.6 The CKM mechanism

While the weak eigenstates and the mass eigenstates of fermions are not the same, the two states are related to another by a unitary transformation. Using separate unitary transformations on the weak eigenstates  $\Psi_L$  and  $\Psi_R$  according to

$$\Psi'_L = U \Psi_L \quad \text{and} \quad \Psi'_R = V \Psi_R,$$

one obtains the new, diagonal mass matrix

$$\mathcal{M}' = U^\dagger \mathcal{M} V.$$

The matrix  $V$  used to diagonalise the quark mass matrix is referred to as the Cabibbo-Kobayashi-Maskawa matrix (CKM matrix,  $V_{CKM}$ ) [76, 77]. It is defined as

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

In this matrix the diagonal elements correspond to the transitions in a given generation, while the off-diagonal elements are responsible for the transitions between different generations. The numerical values of the CKM matrix are given in Ref. [65] as

$$V_{CKM} = \begin{pmatrix} 0.97401 \pm 0.00011 & 0.22650 \pm 0.00048 & 0.00361^{+0.00011}_{-0.00009} \\ 0.22636 \pm 0.00048 & 0.97320 \pm 0.00011 & 0.04053^{+0.00083}_{-0.00061} \\ 0.00854^{+0.00023}_{-0.00016} & 0.03978^{+0.00082}_{-0.00060} & 0.999172^{+0.000024}_{-0.000035} \end{pmatrix}.$$

Considering the unitarity of the CKM matrix, the product of the CKM matrix and its conjugate transposed matrix is defined to be unity, explicitly  $V_{CKM}V_{CKM}^\dagger = V_{CKM}^\dagger V_{CKM} = 1$ . This condition gives rise to three sets of equation systems for the off-diagonals which form the unitarity triangles [78, 79]. Furthermore, the unitarity reduces the number of free parameters needed to describe a pair of  $3 \times 3$  matrices from 18 to 9. Out of these 9 parameters, 5 are can be absorbed into unobservable and, therefore, unphysical quark phases leaving 4 independent parameters. For the 4 remaining free parameters different parameterisation schemes can be used. In the standard parameterisation of  $V_{CKM}$  [80], the independent parameters are three Euler angles,  $\Theta_{12}$ ,  $\Theta_{13}$  and  $\Theta_{23}$  as well as a complex phase  $\delta$ . The phase  $\delta$  is responsible for any violation of the combined charge and parity symmetry ( $CP$ ) in the quark sector of the SM. Using the abbreviations  $c_{ij} = \cos \Theta_{ij}$  and  $s_{ij} = \sin \Theta_{ij}$ , one can write the CKM matrix as

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

## 2.2 Physics of the $B_s^0$ meson

In this section the physics of the  $B_s^0$  meson is introduced. Given the energy scales involved in the physics of the  $B_s^0$  meson, it is an important probe for QCD. Therefore, these decays can be used to test the completeness and accuracy of the SM. Decays of the  $B_s^0$  meson which can only occur at loop level, for example the decay  $B_s^0 \rightarrow \phi\mu^+\mu^-$ , are especially important for these tests as effects of New Physics can significantly contribute to these processes.

### 2.2.1 Production of $b\bar{b}$ pairs at LHCb

In order to produce any pair of quarks the minimal centre-of-mass energy ( $\sqrt{s}$ ) needed is twice the mass of the quark ( $m_q$ ), which is given by  $\sqrt{s} = m_{q\bar{q}}c^2 = 2m_qc^2$ . The produced quarks hadronise and can either form  $q\bar{q}$  resonances or other mesons and baryons. In this thesis, a  $b\bar{b}$  pair needs to be produced and one of the quarks is required to form a  $B_s^0$  meson which has a mass of  $m_{B_s^0} = 5366.3 \text{ MeV}/c^2$  [65]. While the minimum energy to produce the  $b\bar{b}$  pair is  $\sqrt{s} = 2 \cdot 4.16 \text{ GeV} = 8.36 \text{ GeV}$ , at least one of the two quarks needs to form a bound state with an  $s$  quark increasing the minimum energy on the order of a few (GeV). At the LHC the energy threshold to produce both, a  $b\bar{b}$  pair and a  $B_s^0$  meson, is exceeded by roughly a factor 1000 as the centre-of-mass energy is between 7 TeV and 13 TeV for proton-proton collisions depending on the data-taking conditions. Due to the large difference in centre-of-mass energy and the mass of the  $b\bar{b}$  pair as well as the potential imbalance of momenta of the quarks due to the parton distribution function of the proton [81–85], the  $b\bar{b}$  pair will likely not be produced at rest, but boosted along the beam, either in forward or backward direction. As a consequence, a  $B_s^0$  meson originating from the  $b\bar{b}$  pair will also be boosted along the beam. In Fig. 2.1 the effect of the boost on the  $b\bar{b}$  pair is shown using the colour-coded production cross section of a  $b\bar{b}$  pair in the pseudorapidity plane of the  $b\bar{b}$  pair ( $\eta_1$ - $\eta_2$ ) considering that large values for the pseudorapidity correspond to a large boost of light particles. The pseudorapidity of the  $b$  quark is labeled  $\eta_1$

and the pseudorapidity of the  $\bar{b}$  quark is labeled  $\eta_2$ . Regions of phasespace with low  $b\bar{b}$  cross sections are coloured blue with a colour gradient to red indicating increasing cross sections. There are two red regions in which the  $b\bar{b}$  pairs accumulate in forward and backward direction indicating the boost along the beam axis. The yellow and red boxes indicate the geometric acceptance of a typical general purpose detector and of the LHCb experiment, respectively.

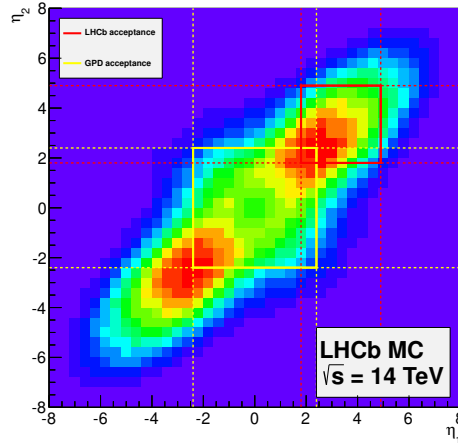


Figure 2.1: The production cross section of a  $b\bar{b}$  colour coded in the  $\eta_1$ - $\eta_2$  plane, where  $\eta_1$  is the pseudorapidity of the  $b$  quark and  $\eta_2$  the pseudorapidity of the  $\bar{b}$  quark. The geometric acceptance of a typical general purpose detector (GPD) and the geometric acceptance of LHCb are given as the yellow and red boxes. Taken from Ref. [86].

The production cross sections over the complete solid angle, are according to PYTHIA8 [87, 88],  $\sigma_{b\bar{b}}^{7 \text{ TeV}} = 251.8 \mu\text{b}$ ,  $\sigma_{b\bar{b}}^{8 \text{ TeV}} = 291.6 \mu\text{b}$  and  $\sigma_{b\bar{b}}^{14 \text{ TeV}} = 527.3 \mu\text{b}$ , respectively. In the  $\eta$  range  $2 < \eta < 5$ , which coincides with the red square in Fig. 2.1, the cross sections are measured [89] by the LHCb collaboration to be  $\sigma_{b\bar{b}}^{7 \text{ TeV, exp.}} = 72.0 \pm 0.3 \pm 6.8 \mu\text{b}$  and  $\sigma_{b\bar{b}}^{13 \text{ TeV, exp.}} = 144 \pm 1 \pm 21 \mu\text{b}$ , where the first uncertainty is of statistical nature and the second one is the systematic uncertainty for both values. These values are in agreement with those predicted by FONLL [90],  $\sigma_{b\bar{b}}^{7 \text{ TeV, theory}} = 62_{-22}^{+28} \mu\text{b}$  and  $\sigma_{b\bar{b}}^{13 \text{ TeV, theory}} = 111_{-44}^{+51} \mu\text{b}$ , where the uncertainty is of systematic nature.

## 2.2.2 $B_s^0$ meson mixing and time-evolution

Meson mixing is a quantum mechanical process in which a neutral meson oscillates into the corresponding anti-meson. To transition from a meson with the quark content  $q\bar{q}'$ , where  $q$  and  $\bar{q}'$  are different quark flavours, to the anti-meson, both quarks of the meson need to transition to the other quark flavour. This is possible as the CKM matrix is not diagonal and, therefore, quark transitions between generations are possible. Exemplary Feynman diagrams for the transition of a  $B_s^0$  meson to a  $\bar{B}_s^0$  meson are shown in Fig. 2.2.

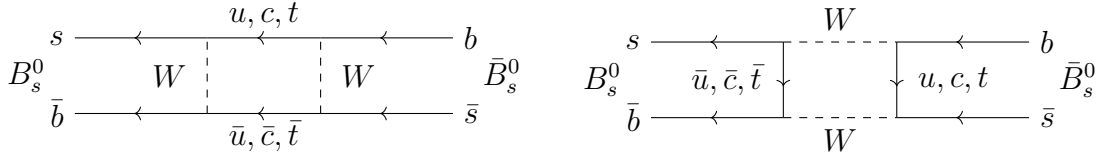


Figure 2.2: A  $B_s^0$  transitions to a  $\bar{B}_s^0$  via two different box diagrams.

When describing the mixed state of the  $B_s^0$  and  $\bar{B}_s^0$  mesons, a given initial state ( $|\Psi(0)\rangle$ ) has to be considered as a superposition of the meson and anti-meson state. Therefore, the initial state is given by

$$|\Psi(0)\rangle = a(0)|B_s^0\rangle + b(0)|\bar{B}_s^0\rangle,$$

where  $a(0)$  and  $b(0)$  are the fractions of  $B_s^0$  and  $\bar{B}_s^0$  in the initial state.<sup>2</sup> The parameters  $a(0)$  and  $b(0)$  are only coherent before the decay of the meson but lose practically all coherence once it decays.

To determine the time-dependent evolution of the initial state  $\Psi(0)$ , the time-dependent parameters  $a$  and  $b$  can be determined by solving the Schrödinger equation

$$i\frac{\partial}{\partial t}|\psi\rangle = \mathcal{H}|\psi\rangle,$$

where  $\mathcal{H}$  is the phenomenological Hamiltonian given by

$$\mathcal{H} = M - \frac{i}{2}\Gamma = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}.$$

The Hermitian matrices  $M$  and  $\Gamma$  appearing in the Hamiltonian correspond to transitions via intermediate off-shell and on-shell states, respectively. The diagonal elements of  $M$  and  $\Gamma$  describe the flavour-conserving transitions  $B_s^0 \rightarrow B_s^0$  and  $\bar{B}_s^0 \rightarrow \bar{B}_s^0$ , whereas the off-diagonal elements describe the flavour-changing transitions  $B_s^0 \leftrightarrow \bar{B}_s^0$ .

The Hamiltonian can be simplified assuming the charge, momentum and time symmetries ( $CPT$ ) to be conserved in the interaction at the same time. If the  $CPT$  symmetry is conserved, the laws of physics are identical for a universe and a copy of that universe, where all matter is replaced with antimatter, the positions of all particles are reflected through a fixed, arbitrary point in space and the momenta of all particles are inverted. In this case, the diagonal elements of each of the matrices  $M$  and  $\Gamma$  have to be the same, explicitly this means  $M_{11} = M_{22}$  and  $\Gamma_{11} = \Gamma_{22}$ . The hermiticity of  $M$  and  $\Gamma$  further implies  $M_{21} = M_{12}^*$  and  $\Gamma_{21} = \Gamma_{12}^*$ . Therefore, the Hamiltonian can be written as

$$\mathcal{H} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{11} - \frac{i}{2}\Gamma_{11} \end{pmatrix}.$$

For a non-diagonal Hamiltonian, the  $B_s^0$  and  $\bar{B}_s^0$  mesons are not mass eigenstates and, therefore, cannot have well defined masses and widths, while the eigenstates of  $\mathcal{H}$  have well defined masses and widths. For the  $B_s^0$  system these eigenstates are typically labeled as the light and heavy eigenstates  $B_{s,L}^0$  and  $B_{s,H}^0$ , respectively, where the mass of the heavy eigenstate ( $m_H$ ) is larger

<sup>2</sup>One of the two parameters can in principle be zero in strong productions leading to a pure initial state.



than the mass of the light eigenstate ( $m_L$ ). This results in the condition  $m_H > m_L$ . The solution to the eigenvalue problem of  $\mathcal{H}$  in the mass eigenbasis is given by

$$(2.2) \quad |B_{s,L/H}^0\rangle = p|B_s^0\rangle \pm q|\bar{B}_s^0\rangle,$$

where the coefficients  $p$  and  $q$  are normalized by  $|p|^2 + |q|^2 = 1$  assuming  $CPT$  and the fraction of the coefficients is fixed by the condition

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

The masses ( $m_{L,H}$ ) and decay widths ( $\Gamma_{L,H}$ ) of the states  $|B_{s,L/H}^0\rangle$  are given by the real and imaginary parts of the corresponding eigenvalues of  $\mathcal{H}$ . In the  $B_s^0$  system the mass and decay width differences between the light and heavy eigenstates ( $\Delta m_s$  and  $\Delta\Gamma_s$ ) are defined as

$$(2.3) \quad \Delta m_s = m_H - m_L \quad \text{and} \quad \Delta\Gamma_s = \Gamma_L - \Gamma_H$$

to guarantee both  $\Delta m_s$  and  $\Delta\Gamma_s$  to be positive. The components  $M_{11}$  of the matrix  $M$  and  $\Gamma_{11}$  of the matrix  $\Gamma$  are connected to  $M_{L,H}$  and  $\Gamma_{L,H}$  using matrix diagonalisation. These components are given as their averages

$$M_{11} = \frac{m_L + m_H}{2} \quad \text{and} \quad \Gamma_{11} = \frac{\Gamma_L + \Gamma_H}{2}.$$

Furthermore, one can relate  $M_{12}$  and  $\Gamma_{12}$  to the masses  $m_{L,H}$  and width  $\Gamma_{12}$  via

$$\begin{aligned} M_{L/H} - \frac{i}{2}\Gamma_{L,H} &= M_{11} - \frac{i}{2}\Gamma_{11} \mp \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)} \\ &= M_{11} - \frac{i}{2}\Gamma_{11} \mp \sqrt{|M_{12}|^2 - \frac{|\Gamma_{12}|^2}{4} - i|M_{12}||\Gamma_{12}|\cos(\Phi_\Gamma - \Phi_M)}, \end{aligned}$$

where the phases  $\Phi_\Gamma$  and  $\Phi_M$  are defined by  $\Phi_\Gamma = \arg \Gamma_{12}$  and  $\Phi_M = \arg M_{12}$ . Using the relations in Eq. 2.3, the differences in the mass and decay width can be written in terms of  $M_{12}$  and  $\Gamma_{12}$  as

$$\begin{aligned} \Delta m_s^2 - \frac{\Delta\Gamma_s^2}{4} &= 4|M_{12}|^2 - |\Gamma_{12}|^2 \\ \text{and} \quad \Delta m_s \Delta\Gamma_s &= -4|M_{12}||\Gamma_{12}|\cos(\Phi_\Gamma - \Phi_M). \end{aligned}$$

If  $\Delta\Gamma_s$  and  $\Gamma_{12}$  are simultaneously small, as is the case for example in the limit of  $CP$  conservation, these equations further simplify to

$$\Delta m_s = 2|M_{12}| \quad \text{and} \quad |\Delta\Gamma_s| = 2|\Gamma_{12}|.$$

The numerical values for the  $\Delta m_s$ ,  $\Gamma_s$  and  $\frac{\Delta\Gamma_s}{\Gamma_s}$  in the  $B_s^0$  system are [65]

$$\Delta m_s = (17.741 \pm 0.020) \times 10^{-12} \hbar s^{-1}, \quad \Gamma_s = (65.97 \pm 0.26) \times 10^{10} s^{-1} \quad \text{and} \quad \frac{\Delta\Gamma_s}{\Gamma_s} = 0.124 \pm 0.008.$$

The time-dependence of the  $B_s^0$  and  $\bar{B}_s^0$  mesons can be derived by solving the Schrödinger equation. Considering the Hamiltonian is diagonal in the mass basis and Eq. 2.2 directly links the wave functions of the  $B_s^0$  and  $\bar{B}_s^0$  mesons to those of the mass eigenstates  $B_{s,L}^0$  and  $B_{s,H}^0$ , it is more convenient to first derive the solution for the mass eigenstates  $B_{s,L}^0$  and  $B_{s,H}^0$  and then substitute the results to find the solutions for the  $B_s^0$  and  $\bar{B}_s^0$  mesons. The dependence of the  $B_s^0$  and  $\bar{B}_s^0$  mesons on the  $B_{s,L}^0$  and  $B_{s,H}^0$  eigenstates can be obtained by rearranging the two equations Eq. 2.2. Doing this, one finds the relations

$$(2.4) \quad |B_s^0(t)\rangle = \frac{1}{2p} (|B_{s,L}^0(t)\rangle + |B_{s,H}^0(t)\rangle)$$

$$(2.5) \quad \text{and } |\bar{B}_s^0(t)\rangle = \frac{1}{2q} (|B_{s,L}^0(t)\rangle - |B_{s,H}^0(t)\rangle) .$$

In the mass eigenstate basis the Schrödinger equation has the form

$$i \frac{\partial}{\partial t} \begin{pmatrix} |B_{s,L}^0(t)\rangle \\ |B_{s,H}^0(t)\rangle \end{pmatrix} = \begin{pmatrix} M_L - \frac{i}{2}\Gamma_L & 0 \\ 0 & M_H - \frac{i}{2}\Gamma_H \end{pmatrix} \begin{pmatrix} |B_{s,L}^0(t)\rangle \\ |B_{s,H}^0(t)\rangle \end{pmatrix}$$

and is a linear differential equation of the first order. It can be solved using the Ansatz

$$\begin{aligned} |B_{s,L}^0(t)\rangle &= e^{-iM_L t} e^{-\frac{\Gamma_L}{2}t} |B_{s,L}^0(0)\rangle \\ \text{and } |B_{s,H}^0(t)\rangle &= e^{-iM_H t} e^{-\frac{\Gamma_H}{2}t} |B_{s,H}^0(0)\rangle . \end{aligned}$$

With these solutions the time-dependent wave functions of the  $B_s^0$  and  $\bar{B}_s^0$  mesons can be obtained by substitution in Eq. 2.4 and Eq. 2.5. The time-dependent equations for the  $B_s^0$  and  $\bar{B}_s^0$  mesons are then given by

$$\begin{aligned} |B_s^0(t)\rangle &= \frac{1}{2} \left( e^{-iM_L t} e^{-\frac{\Gamma_L}{2}t} + e^{-iM_H t} e^{-\frac{\Gamma_H}{2}t} \right) |B_s^0(0)\rangle \\ &+ \frac{q}{2p} \left( e^{-iM_L t} e^{-\frac{\Gamma_L}{2}t} - e^{-iM_H t} e^{-\frac{\Gamma_H}{2}t} \right) |\bar{B}_s^0(0)\rangle \\ &= g_+(t) |B_s^0(0)\rangle + \frac{q}{p} g_-(t) |\bar{B}_s^0(0)\rangle \\ |\bar{B}_s^0(t)\rangle &= \frac{p}{2q} \left( e^{-iM_L t} e^{-\frac{\Gamma_L}{2}t} - e^{-iM_H t} e^{-\frac{\Gamma_H}{2}t} \right) |B_s^0(0)\rangle \\ &+ \frac{1}{2} \left( e^{-iM_L t} e^{-\frac{\Gamma_L}{2}t} + e^{-iM_H t} e^{-\frac{\Gamma_H}{2}t} \right) |\bar{B}_s^0(0)\rangle \\ &= \frac{p}{q} g_-(t) |B_s^0(0)\rangle + g_+(t) |\bar{B}_s^0(0)\rangle , \end{aligned}$$

where the  $g_{\pm}$  are given as

$$g_{\pm} = \frac{1}{2} \left( e^{-iM_L t} e^{-\frac{\Gamma_L}{2}t} \pm e^{-iM_H t} e^{-\frac{\Gamma_H}{2}t} \right) .$$

## 2.3 The rare decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$

The rare decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  is a flavour changing neutral current that only occurs at loop level in the SM. The leading order Feynman diagrams are shown in Fig. 2.3. The mathematical description of the decay in the SM is split into different parts. The quark-level  $b \rightarrow s \ell^+ \ell^-$  transition, as shown in Fig. 2.4, is commonly described using an effective field theory in the literature [55, 56]. The transition of a  $B_s^0$  meson into a  $\phi$  meson, also called the hadronic part of the decay, is described using form factors. The information of the  $b \rightarrow s \ell^+ \ell^-$  transition and the hadronic part of the decay is then combined into transversity amplitudes which in turn are used to provide formulae for the angular decay rates.

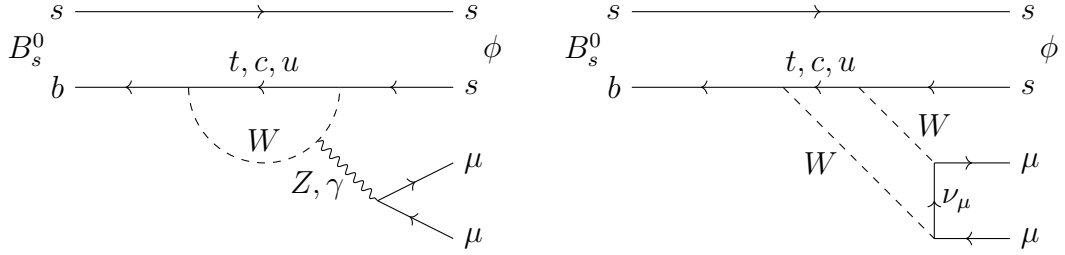


Figure 2.3: The decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  is mediated by (left) penguin diagrams and (right) box diagrams in the SM. The charges of the fermions are implied by the direction of the arrow.

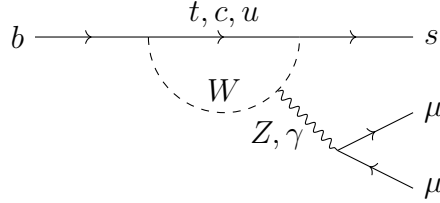


Figure 2.4: Feynman diagram of a  $b \rightarrow s \ell^+ \ell^-$  transition.

### 2.3.1 Effective field theory for $b \rightarrow s \ell^+ \ell^-$ transitions

Effective field theories (EFT) are a tool commonly used to analyse systems with multiple energy scales of interest, at least one of which is significantly larger than the others. In problems with multiple energy scales, an EFT uses the ratio of the energy scales as a expansion parameter. One famous example for an effective field theory is Fermi's theory of the neutron decay [91, 92], in which the interaction of the  $W^\pm$  boson with the neutron is approximated to be point-like. In this case, the inverse of the mass of the  $W^\pm$  boson serves as the expansion parameter. A general introduction to EFTs is given in Ref. [93].

In  $b \rightarrow s \ell^+ \ell^-$  transitions, the loop process involves multiple scales of interest. These are the masses of the  $t$  quark ( $m_t$ ) and of the  $W^\pm$  boson ( $m_W$ ) as these particles dominantly contribute to the loop, the square root of the transferred four momentum ( $\sqrt{q^2}$ ), the mass of the  $b$  quark ( $m_b$ ) and the energy scale  $\Lambda_{QCD}$ . The transferred four momentum is strictly smaller than the mass of the  $B_s^0$  meson and is as such also much smaller than the mass of the  $W^\pm$  boson. To

describe the transition of the  $b$  quark the operator product expansion (OPE) [94], is commonly used [95]. At the intermediate scale of the mass of the  $b$  quark ( $m_b$ ), the effective Hamiltonian is given in Ref. [53] as

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{\alpha_e}{4\pi} \sum_i \mathcal{C}_i(\mu_s = m_b) \mathcal{O}_i(\mu_s = m_b),$$

where  $G_F$  is Fermi's constant, the  $V_{jk}$  are CKM matrix elements,  $\alpha_e$  is the fine-structure constant,  $\mathcal{C}_i$  are Wilson coefficients [96],  $\mu_s$  is the renormalisation scale and  $\mathcal{O}_i$  are local operators with different Lorentz structure. The Wilson coefficients encode all short-range contributions and act as an effective coupling to the operators, which correspond to the long-range contributions. The most relevant operators for  $b \rightarrow s \ell^+ \ell^-$  transitions are  $\mathcal{O}_{7,9,10}$ , which have the form [53]

$$\begin{aligned} \mathcal{O}_7 &= \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_R b F_{\mu\nu}, \\ \mathcal{O}_9 &= \bar{s} \gamma_\mu P_L b \bar{\ell} \gamma^\mu \ell \\ \text{and } \mathcal{O}_{10} &= \bar{s} \gamma_\mu P_L b \bar{\ell} \gamma^\mu \gamma_5 \ell, \end{aligned}$$

where  $\gamma^\mu$  and  $\gamma_5$  are the Dirac matrices,  $\sigma^{\mu\nu} = \frac{1}{4}(\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu)$ ,  $P_{L,R} = \frac{1 \mp \gamma_5}{2}$  denotes the left- and right-handed chiral projections and  $F_{\mu\nu}$  denotes the electromagnetic field strength tensor. The operator  $\mathcal{O}_7$  corresponds to a flavour changing neutral current in which a photon is emitted, whereas the operators  $\mathcal{O}_{9,10}$  correspond to the electroweak decays with a vector- or an axialvector-current, respectively. An illustration for the interactions is given in Fig. 2.5. In New Physics (NP) scenarios the chirality flipped operators  $\mathcal{O}'_i$  with the corresponding Wilson coefficients  $\mathcal{C}'_i$  can compete with the SM operators  $\mathcal{O}_i$ .

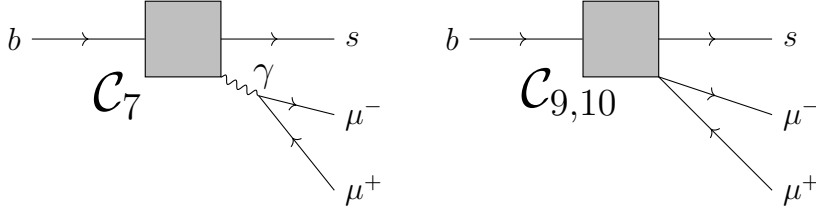


Figure 2.5: Visual representation of the local operators of  $b \rightarrow s \ell^+ \ell^-$  transitions, on the left photonic operator  $\mathcal{O}_7$  is illustrated, on the right the vector- and axialvector-currents of the operators  $\mathcal{O}_{9,10}$ . The Wilson coefficients  $\mathcal{C}_i$  are numerical values representing the described process.

The Wilson coefficients  $\mathcal{C}_{7,9,10}$  can be calculated in the SM by matching the amplitudes of the full electroweak theory to the effective Hamiltonian  $\mathcal{H}_{\text{eff}}$ . The coefficients can be calculated at  $\mu_s = m_b$  using the renormalisation group equation assuming SM dynamics [97] at next-to-leading-logarithmic order to be [98]

$$\mathcal{C}_7^{\text{eff}} = -0.308 \quad \mathcal{C}_9 = +4.154 \quad \mathcal{C}_{10} = -4.261.$$

The Wilson coefficients are typically used to determine the SM expectation for observables such as branching fractions or angular observables.

In the presence of heavy NP contributions, the values for the Wilson coefficients would be shifted from the SM prediction ( $\Delta\mathcal{C}_i = \mathcal{C}_i^{\text{NP}} - \mathcal{C}_i^{\text{SM}}$ ). These shifts in the Wilson coefficients would also translate to a modification of the effective Hamiltonian of the form

$$\Delta H^{\text{NP}} = \sum_i \frac{c_i^{\text{NP}}}{\Lambda_{\text{NP}}^2} \mathcal{O}_i = \sum_i C_i^{\text{NP}} \mathcal{O}_i$$

where the index  $i$  indicates all affected Wilson coefficients,  $c_i^{\text{NP}}$  is the coupling of the NP,  $\Lambda_{\text{NP}}$  the energy scale and  $\mathcal{O}_i$  the affected operator. The value of the Wilson coefficients  $\mathcal{C}_{7,9,10}$  can be measured in data by studying for example branching fractions or angular distributions of  $b \rightarrow s\ell^+\ell^-$  decays. If a new, preferred value for the Wilson coefficients is found, the energy scale of a NP scenario can be estimated by assuming the coupling structure of the NP.

### 2.3.2 Form factors for the $B_s^0 \rightarrow \phi$ transition

While an EFT can be used to describe the quark level  $b \rightarrow s\ell^+\ell^-$  transition, the  $B_s^0 \rightarrow \phi\mu^+\mu^-$  decay includes a  $B_s^0 \rightarrow \phi$  transition that cannot be described using an EFT. Instead the hadronic transition is described using form factors. There are seven different form factors, one for the vector-current ( $V$ ) and three each for the axial-vector-currents ( $A_{0,1,2}$ ) and tensor-currents ( $T_{1,2,3}$ ). All seven hadronic form factors depend on the transferred four-momentum ( $q^2$ ). To calculate these form factors, there are two different, complementary approaches. The first one is the use of light-cone sum rules (LCSR) [99] and the second one relies on lattice QCD [100] computations. While LCSR are valid if the energy of the  $\phi$  meson is large, corresponding to low values of  $q^2$ , lattice QCD is valid if the  $\phi$  meson has little energy, corresponding to large values of  $q^2$ . Typically, LCSR are used in the  $q^2$  region below  $6 \text{ GeV}^2/c^4$  and lattice QCD is used in the  $q^2$  region above  $15 \text{ GeV}^2/c^4$ . The two predictions can also be combined using a fit to cover the complete  $q^2$  range. The intermediate  $q^2$  region, namely the region of  $q^2 \in [6, 15 \text{ GeV}^2/c^4]$ , is sensitive to long-distance effects from the charmonium resonances, which cannot be theoretically calculated from first principles in the SM [56] as the narrow charmonium resonances lead to violations of the assumed quark hadron duality [101, 102]. The influence of the intermediate  $q^2$  region on the low  $q^2$  region is another topic of discussion [103–107].

### 2.3.3 Transversity amplitudes

The decay rate of the process  $B_s^0 \rightarrow \phi\mu^+\mu^-$  can be described by a set of seven different, complex transversity amplitudes,  $A_0^{\text{L,R}}$ ,  $A_{\parallel}^{\text{L,R}}$  and  $A_{\perp}^{\text{L,R}}$ , where L and R indicates the left- or right-handed chirality of the dimuon system, as well as  $A_t$ . The indices 0,  $\parallel$  and  $\perp$  indicate the different polarisation modes of the final state. These are illustrated in Fig. 2.6. In addition, the dimuon system can also be in a spin-0 configuration giving rise to the time-like amplitude  $A_t$ . This amplitude couples only to axial-vector currents which is given as  $q^\mu(\bar{\mu}\gamma^\mu\gamma_5\mu = 2im_\mu(\bar{\mu}\gamma^\mu\mu)$  and is, therefore, suppressed in the SM by the small lepton mass. However, in the presence of NP contributions, the amplitude  $A_t$  can become sizeable as can additional scalar and tensor amplitudes [108].

The seven considered amplitudes depend on the Wilson coefficients  $\mathcal{C}_{7,9,10}$ ,  $q^2$  and the hadronic form factors. They are given by the relations [56]

$$\begin{aligned}
A_0^{\text{L,R}} &= -\frac{N}{2m_\phi\sqrt{q^2}} \left[ [(C_9 - C'_9) \mp (C_{10} - C'_{10})] \right. \\
&\quad \times [(m_{B_s^0}^2 - m_\phi^2 - q^2)(m_{B_s^0} + m_\phi)A_1(q^2) - \lambda \frac{A_2(q^2)}{m_{B_s^0} + m_\phi}] \\
&\quad \left. + 2m_b(C_7 - C'_7)[(m_{B_s^0}^2 + 3m_\phi^2 - q^2)T_2(q^2) - \lambda \frac{T_3(q^2)}{m_{B_s^0}^2 - m_\phi^2}] \right], \\
A_{\parallel}^{\text{L,R}} &= -N\sqrt{2}(m_{B_s^0}^2 - m_\phi^2) \left[ [(C_9 - C'_9) \mp (C_{10} - C'_{10})] \frac{A_1(q^2)}{m_{B_s^0} - m_\phi} + \frac{2m_b}{q^2}(C_7 - C'_7)T_2(q^2) \right], \\
A_{\perp}^{\text{L,R}} &= N\sqrt{2}\lambda \left[ [(C_9 + C'_9) \mp (C_{10} + C'_{10})] \frac{V(q^2)}{m_{B_s^0} + m_\phi} + \frac{2m_b}{q^2}(C_7 + C'_7)T_1(q^2) \right] \\
\text{and } A_t &= \frac{N}{\sqrt{q^2}}\sqrt{\lambda} \left[ 2(C_{10} - C'_{10}) + \frac{q^2}{m_\mu}(C_P - C'_P) \right] A_0(q^2),
\end{aligned}$$

where the index  $C_P$  corresponds to Wilson coefficient of the pseudoscalar current, and  $N$  and  $\lambda$  are abbreviations defined as [56]

$$\begin{aligned}
\lambda &= m_{B_s^0}^4 + m_\phi^4 + q^4 - 2(m_{B_s^0}^2 m_\phi^2 + m_\phi^2 q^2 + m_{B_s^0}^2 q^2), \\
\text{and } N &= V_{tb}V_{ts}^* \left[ \frac{G_F^2 \alpha^2}{3 \times 2^{10} \pi^5 m_{B_s^0}^3} q^2 \sqrt{\lambda} \beta_\mu \right]^{\frac{1}{2}},
\end{aligned}$$

with  $\beta_\mu = \sqrt{1 - \frac{4m_\mu^2}{q^2}}$ .

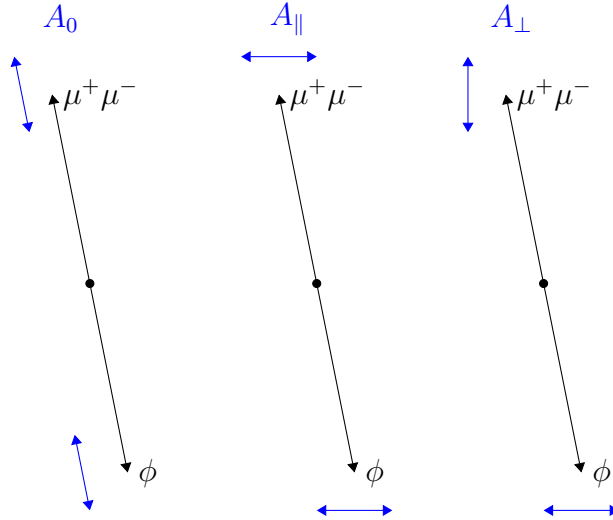


Figure 2.6: Visual representation of the polarisation modes of the final state of the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  for the transversity amplitudes  $A_0$ ,  $A_{\parallel}$  and  $A_{\perp}$ .

### 2.3.4 Differential decay rate

The differential angular decay rate for the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  is given in Refs. [55, 56] as

$$\begin{aligned}
 \frac{d^4\Gamma(B_s^0 \rightarrow \phi \mu^+ \mu^-)}{dq^2 d\cos\theta_l d\cos\theta_K d\phi} &= \sum_i J_i(q^2) f_i(\cos\theta_l, \cos\theta_K, \phi) \\
 &= \frac{9}{32\pi} [J_1^s \sin^2\theta_K + J_1^c \cos^2\theta_K \\
 &\quad + J_2^s \sin^2\theta_K \cos 2\theta_l + J_2^c \cos^2\theta_K \cos 2\theta_l \\
 &\quad + J_3 \sin^2\theta_K \sin^2\theta_l \cos 2\phi + J_4 \sin 2\theta_K \sin 2\theta_l \cos \phi \\
 &\quad + J_5 \sin 2\theta_K \sin \theta_l \cos \phi + J_6^s \sin^2\theta_K \cos \theta_l \\
 &\quad + J_7 \sin 2\theta_K \sin \theta_l \sin \phi + J_8 \sin 2\theta_K \sin 2\theta_l \sin \phi \\
 &\quad + J_9 \sin^2\theta_K \sin^2\theta_l \sin 2\phi],
 \end{aligned}
 \tag{2.6}$$

where  $J_i$  are angular coefficients and  $\theta_l$ ,  $\theta_K$  and  $\phi$  are the helicity angles of the decay.<sup>3</sup> The decay angle  $\theta_l$  is defined as the angle between the negatively charged muon,  $\mu^-$ , and the  $B_s^0$  in the  $\mu^+ \mu^-$  rest frame

$$\cos\theta_l = \frac{\vec{p}_{\mu^-}^{\mu^+ \mu^-} \cdot \vec{p}_{B_s^0}^{\mu^+ \mu^-}}{|\vec{p}_{\mu^-}^{\mu^+ \mu^-}| |\vec{p}_{B_s^0}^{\mu^+ \mu^-}|},$$

$\theta_K$  is the angle between the negatively charged kaon,  $K^-$ , and the  $B_s^0$  in the  $K^+ K^-$  rest frame

$$\cos\theta_K = \frac{\vec{p}_{K^-}^{K^+ K^-} \cdot \vec{p}_{B_s^0}^{K^+ K^-}}{|\vec{p}_{K^-}^{K^+ K^-}| |\vec{p}_{B_s^0}^{K^+ K^-}|},$$

and  $\phi$  is the angle between the planes spanned by the  $\mu^+ \mu^-$  and  $K^+ K^-$  systems in the  $B_s^0$  rest frame. It is defined via the two relations

$$\cos\phi = \vec{n}_{K^+ K^-}^{B_s^0} \cdot \vec{n}_{\mu^+ \mu^-}^{B_s^0} \quad \text{and} \quad \sin\phi = (\vec{n}_{K^+ K^-}^{B_s^0} \times \vec{n}_{\mu^+ \mu^-}^{B_s^0}) \cdot \frac{\vec{p}_{K^+ K^-}^{B_s^0}}{|\vec{p}_{K^+ K^-}^{B_s^0}|}.$$

The set of the three decay angles ( $\vec{\Omega}$ ) are illustrated in Fig. 2.7.

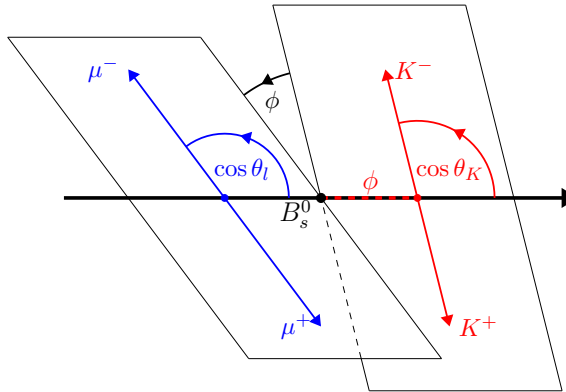


Figure 2.7: A sketch to explain the three angles,  $\theta_K$ ,  $\theta_l$  and  $\phi$ .

<sup>3</sup>In literature the  $J_i$  are sometimes given as  $I_i$ .

The  $J_i$  are  $q^2$  dependent observables that are given by combinations of seven complex polarisation amplitudes ( $A_0^{L,R}$ ,  $A_{\parallel}^{L,R}$ ,  $A_{\perp}^{L,R}$  and  $A_t$ ) as [55, 56]

$$\begin{aligned}
J_1^s &= \frac{(2 + \beta_\mu^2)}{4} [|A_{\perp}^L|^2 + |A_{\parallel}^L|^2 + (L \rightarrow R)] + \frac{4m_\mu^2}{q^2} \Re(A_{\perp}^L A_{\perp}^{R*} + A_{\parallel}^L A_{\parallel}^{R*}) \\
J_1^c &= |A_0^L|^2 + |A_0^R|^2 + \frac{4m_\mu^2}{q^2} [|A_t|^2 + 2\Re(A_0^L A_0^{R*})] \\
J_2^s &= \frac{\beta_\mu^2}{4} [|A_{\perp}^L|^2 + |A_{\parallel}^L|^2 + (L \rightarrow R)] \\
J_2^c &= -\beta_\mu^2 [|A_0^L|^2 + (L \rightarrow R)] \\
J_3 &= \frac{\beta_\mu^2}{2} [|A_{\perp}^L|^2 - |A_{\parallel}^L|^2 + (L \rightarrow R)] \\
J_4 &= \frac{\beta_\mu^2}{\sqrt{2}} [\Re(A_0^L A_{\parallel}^{L*}) + (L \rightarrow R)] \\
J_5 &= \sqrt{2}\beta_\mu [\Re(A_0^L A_{\perp}^{L*}) - (L \rightarrow R)] \\
J_6^s &= 2\beta_\mu [\Re(A_{\parallel}^L A_{\perp}^{L*}) - (L \rightarrow R)] \\
J_7 &= \sqrt{2}\beta_\mu [\Im(A_0^L A_{\parallel}^{L*}) - (L \rightarrow R)] \\
J_8 &= \frac{\beta_\mu^2}{\sqrt{2}} [\Im(A_0^L A_{\perp}^{L*}) + (L \rightarrow R)] \\
(2.7) \quad J_9 &= \beta_\mu^2 [\Im(A_{\parallel}^{L*} A_{\perp}^L) + (L \rightarrow R)],
\end{aligned}$$

where  $L$  and  $R$  refer to the chirality of the dimuon system, the notation  $L \rightarrow R$  indicates an additional set of the left-handed terms with  $R$  replacing  $L$  and  $\beta_\mu^2$  is defined by  $\beta_\mu^2 = (1 - 4m_\mu^2/q^2)$ . For the decay of a  $\bar{B}_s^0$  meson the replacements  $J_{1,2,3,4,7} \rightarrow +\bar{J}_{1,2,3,4,7}$  and  $J_{5,6,8,9} \rightarrow -\bar{J}_{5,6,8,9}$  need to be made in Eq. 2.6 [55, 56]. The observables  $J_1^s$  and  $J_2^s$  as well as  $J_1^c$  and  $J_2^c$  are generally not independent and obey the relations  $J_1^s = 3J_2^s$  and  $J_1^c = -J_2^c$  for a value of  $\beta_\mu^2 = 1$ . This is a good approximation for  $q^2$  values that are large compared to the square of the muon mass. Furthermore, the observables  $J_{1,2}^s$  and  $J_{1,2}^c$  are connected to each other by the normalisation of the decay rate, as the condition

$$\frac{1}{d\Gamma/dq^2} \iiint \frac{d^4(\Gamma)}{dq^2 d\cos\theta_l d\cos\theta_K d\phi} d\cos\theta_l d\cos\theta_K d\phi = \frac{1}{d\Gamma/dq^2} \frac{d\Gamma}{dq^2} = 1$$

holds. The analogous equation for the  $\bar{B}_s^0$  meson also holds. Integration of Eq. 2.6 over the three angles yields

$$\frac{1}{d\Gamma/dq^2} \frac{d\Gamma}{dq^2} = \frac{2J_1^s + J_1^c}{4} - \frac{2J_2^s + J_2^c}{4} = 1$$

which can be simplified for  $\beta_\mu^2 \approx 1$  to

$$J_1^s = \frac{3}{4} (1 - J_1^c) .$$



The flavour of the  $B_s^0$  meson at the time of decay is unknown considering the final state  $\phi[\rightarrow K^+K^-]\mu^+\mu^-$  is flavour-symmetric and flavour tagging is not possible due to the limited statistics. The measurement is, therefore, performed untagged and time-integrated and the differential decay rates for the  $B_s^0$  decay and the  $\bar{B}_s^0$  decay are combined. The combined untagged angular decay rate of the  $B_s^0$  and  $\bar{B}_s^0$  mesons results in  $CP$ -averaged observables for the terms  $J_{1,2,3,4,7}$  and  $CP$ -asymmetric observables for the terms  $J_{5,6,8,9}$ . In any  $q^2$  region the combined decay rate is given by

$$(2.8) \quad \frac{d^4(\Gamma + \bar{\Gamma})}{dq^2 d\cos\theta_l d\cos\theta_K d\phi} = \frac{9}{32\pi} \left[ \frac{3}{4}(1 - F_L) \sin^2 \theta_K (1 + \frac{1}{3} \cos 2\theta_l) + F_L \cos^2 \theta_K (1 - \cos 2\theta_l) \right. \\ \left. + S_3 \sin^2 \theta_K \sin^2 \theta_l \cos 2\phi + S_4 \sin 2\theta_K \sin 2\theta_l \cos \phi \right. \\ \left. + A_5 \sin 2\theta_K \sin \theta_l \cos \phi + \frac{4}{3} A_{\text{FB}}^{CP} \sin^2 \theta_K \cos \theta_l \right. \\ \left. + S_7 \sin 2\theta_K \sin \theta_l \sin \phi + A_8 \sin 2\theta_K \sin 2\theta_l \sin \phi \right. \\ \left. + A_9 \sin^2 \theta_K \sin^2 \theta_l \sin 2\phi \right]$$

where the

$$S_i(q^2) = (J_i + \bar{J}_i) / \frac{d(\Gamma + \bar{\Gamma})}{dq^2}$$

denote  $CP$ -averaged quantities and the

$$A_i(q^2) = (J_i - \bar{J}_i) / \frac{d(\Gamma + \bar{\Gamma})}{dq^2}$$

are  $CP$ -asymmetries. In Eq. 2.8 the longitudinal polarisation fraction ( $F_L$ ), a  $CP$ -average, as well as the  $CP$ -asymmetry of the forward-backward asymmetry between the  $B_s^0$  and the  $\bar{B}_s^0$  meson ( $A_{\text{FB}}^{CP}$ ) are used. They are defined as  $F_L = S_1^c$  and  $A_{\text{FB}}^{CP} = \frac{3}{4} A_6^s$ . Furthermore, the lepton masses are assumed to be small compared to  $q^2$  in Eq. 2.8.

Following Ref. [57], the angular decay rate also has a dependence on time. The time-dependence originates from the  $B_s^0$ -meson mixing and the  $K^+K^-\mu^+\mu^-$  final state being a  $CP$ -conjugate state of itself, and, therefore, identical for the decay of a  $B_s^0$  meson and that of a  $\bar{B}_s^0$  meson. The time-dependence can be absorbed into the  $CP$ -averaged and  $CP$ -asymmetric observables,  $S_i$  and  $A_i$ , by using two new angular observables ( $h_i$  and  $s_i$ ). The  $h_i$  and  $s_i$  are respectively connected to the real ( $\mathcal{Re}$ ) and imaginary ( $\mathcal{Im}$ ) parts of combinations of the transversity amplitudes and offer an additional avenue to probe for NP contributions. Their definitions are given in App. A. The time-dependence of the  $CP$ -averages and  $CP$ -asymmetries is explicitly given as [57]

$$(2.9) \quad J_i(t) + \bar{J}_i(t) = e^{-\Gamma_s t} \left[ (J_i + \bar{J}_i) \cosh \left( \frac{\Delta\Gamma_s}{2\Gamma_s} t \right) - h_i \sinh \left( \frac{\Delta\Gamma_s}{2\Gamma_s} t \right) \right]$$

$$(2.10) \quad \text{and } J_i(t) - \bar{J}_i(t) = e^{-\Gamma_s t} \left[ (J_i - \bar{J}_i) \cos \left( \frac{\Delta m_s}{\Gamma_s} t \right) - s_i \sin \left( \frac{\Delta m_s}{\Gamma_s} t \right) \right].$$

The differences are defined in Eq. 2.3. The differences of the lifetimes and masses are often abbreviated as  $x = \frac{\Delta m_s}{\Gamma_s}$  and  $y = \frac{\Delta\Gamma_s}{2\Gamma_s}$ . For incoherent production of a  $b\bar{b}$  quark pair, as is typically the case at a hadron collider, the time-integrated observables one can measure are given by

$$S_i = \langle J_i + \bar{J}_i \rangle_{\text{Hadronic}} = \frac{1}{\Gamma} \left[ \frac{J_i + \bar{J}_i}{1-y^2} - \frac{y \cdot h_i}{1-y^2} \right]$$

$$\text{and } A_i = \langle J_i - \bar{J}_i \rangle_{\text{Hadronic}} = \frac{1}{\Gamma} \left[ \frac{J_i - \bar{J}_i}{1+x^2} - \frac{x \cdot s_i}{1+x^2} \right],$$

where  $\langle J_i \pm \bar{J}_i \rangle$  indicates the time-integrated combination of the observables  $J_i$  and  $\bar{J}_i$ . In state-of-the art SM predictions, for example as provided in the **flavio** software package [109], the time-integrated formulae are used to calculate the SM predictions.

# Chapter 3

## The LHCb experiment

In this Chapter, the LHCb experiment, located at the Large Hadron Collider, will be introduced. The Large Hadron Collider (LHC), which provides the proton-proton ( $pp$ ) collisions recorded with the LHCb experiment, will be covered in Sec. 3.1. The subdetectors and core concepts of the LHCb experiment will be discussed in Sec. 3.2.

### 3.1 The Large Hadron Collider

The Large Hadron Collider [110] is the most recent in a series of particle accelerators operated by the Conseil Européen pour la Recherche Nucléaire (CERN) and is located near Geneva. CERN is able to operate the LHC with protons, heavy ions or both at the same time. For this thesis only the  $pp$  collisions are relevant and, therefore, only this type of collision will be discussed.

The protons are obtained by stripping hydrogen gas, which originates from a hydrogen bottle, of the electron and accelerating the remaining nucleus to 9 keV in a duoplasmatron [111]. They are then injected into a radio frequency quadrupole (RFQ) [112]. The protons are focussed and accelerated to 750 keV using radio frequency (RF) cavities. In RF cavities a charged particle is accelerated using an electric field (E-field) oscillating at a frequency in the radio frequency range. Due to the oscillation, the size of each RF cavity and the timing of the protons entering the RF cavity have to be chosen carefully such that the proton always arrives at an accelerating E-field and never at a decelerating one. The RF cavities are employed both in the RFQ and in all subsequent accelerators. In addition to focusing and accelerating the beam, the RFQ also divides the continuous beam of protons adiabatically into 6 parts (bunches). These bunches are then injected into the LINAC 2 [113] and further accelerated to 50 MeV. While both the RFQ and the LINAC 2 are linear accelerators, all further links in the chain, are circular accelerators [114]. The first one of these is the proton synchrotron booster (PS Booster) which accelerates the protons to a final energy of 1.4 GeV. In the proton synchrotron (PS) the 6 bunches are split in two steps into 72 bunches. These bunches are injected into the super proton synchrotron (SPS) at an energy of 25 GeV. They leave the SPS at an energy of 450 GeV and finally enter the LHC. The LHC is designed for a maximal energy of 14 TeV and uses RF cavities with a frequency of 400 MHz. Once the beam is at the operating energy of the LHC, protons entering the RF cavity at the right timing and chosen energy will experience no accelerating voltage and are kept at the same energy. In contrast, protons with differing energies will either arrive earlier or later and thus be accelerated or decelerated. These proton

bunches will converge after a few cycles at the operating energy of the LHC. In Fig. 3.1 all the accelerators starting from LINAC 2 are sketched.

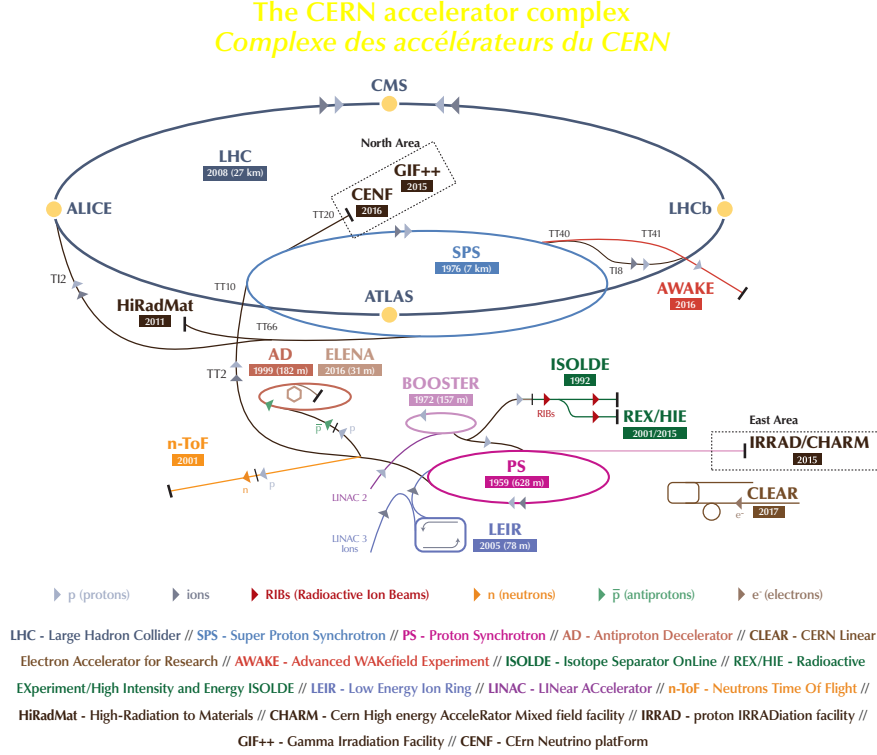


Figure 3.1: The accelerator complex hosted by CERN. The year indicates when the accelerator or experiment was taken into operation and their length or circumference (when applicable) are given. Taken from Ref. [115].

The LHC is designed for two beams with 2808 bunches, corresponding to 39 SPS fills, containing a total of  $1.15 \cdot 10^{11}$  protons [116]. The two beams traverse the LHC simultaneously in opposite directions which requires a set of two beam pipes. The beams are kept on a circular path using dipole magnets with opposite magnetic fields for the two beams and focused with quadrupole magnets. Around the storage ring there are four interaction points at which the beams are brought to collision with a bunch crossing frequency of 40 MHz. At each of these points CERN hosts one large experiment. Two interaction points are dedicated to general purpose detectors, ATLAS [36] and CMS [37], and the other two are dedicated to specialised experiments, ALICE [117] and LHCb [38]. While the two general purpose detectors have a broad physics program, ranging from the study of the Standard Model of particle physics to the search for New Physics, ALICE is dedicated to heavy-ion physics, specifically the study of strongly interacting matter at extreme energy densities, and LHCb is dedicated to heavy flavour physics and will be introduced in Sec. 3.2.

So far, there have been two data taking periods, respectively called Run 1 and Run 2. The data in Run 1 was taken between 2010 and 2012, where data taken in 2010 contributes a negligible amount and is usually not taken into account, and the data in Run 2 was taken between the years 2015 and 2018. The amount of data taken is typically given by the integrated

luminosity [118] ( $\mathcal{L}_{\text{int}}$ ), which is defined as

$$(3.1) \quad \mathcal{L}_{\text{int}} = \int \mathcal{L}(t) dt = \int \frac{f \cdot N_1(t) N_2(t)}{4\pi\sigma_x\sigma_y} dt,$$

where  $\mathcal{L}$  is the instantaneous luminosity,  $f$  is the bunch crossing frequency,  $N_1$  and  $N_2$  are the number of protons per bunch of the two beams and  $\sigma_x$  and  $\sigma_y$  are the transverse beam sizes in the  $x$  or  $y$  direction, respectively. The LHC is designed for an instantaneous luminosity of  $\mathcal{L} = 1 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . In 2017 the LHC provided an instantaneous luminosity of  $\mathcal{L} = 2.06 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  [119], roughly doubling the design value.

## 3.2 The LHCb detector

The Large Hadron Collider beauty (LHCb) experiment [38, 120, 121] is designed to fully exploit the potential at the LHC to measure heavy flavour decays and to probe CP-violation in the heavy flavour sector. A detailed description of the LHCb experiment can be found in Ref. [38].

Among the four large experiments at the LHC, LHCb is unique in many aspects. While ALICE, ATLAS and CMS follow a hermetic detector design, covering almost  $4\pi$  steradians of solid angle, LHCb is a single arm forward spectrometer, as shown in Fig. 3.2. The LHCb experiment has an angular coverage of the region between roughly 10 mrad to 300(250) mrad in the (non-)bending plane. This design choice follows the predominant production of  $b\bar{b}$  pairs in a forward or backward cone, as shown in Fig. 2.1 in Sec. 2.2.1. Approximately 25% of all produced  $b\bar{b}$  pairs are covered in the LHCb acceptance.

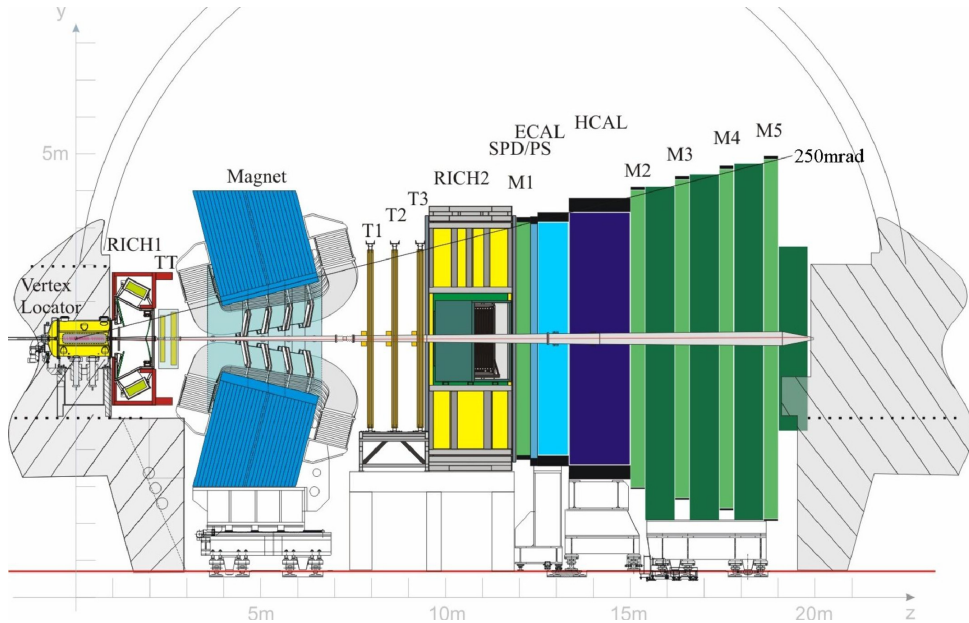


Figure 3.2: Schematic side view of the LHCb experiment. The subdetectors and the warm magnet are labeled. Taken from Ref. [38].

In contrast to ATLAS and CMS, both of which are designed to operate at the design instantaneous luminosity of the LHC, LHCb is designed for a lower instantaneous luminosity of  $\mathcal{L} = 2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ . This design choice will be discussed further in Sec. 3.2.1. The tracking

system at LHCb will be described in Sec. 3.2.2. The Ring Imaging Cherenkov detectors, the electronic and hadronic calorimeters, and the muon stations will be covered in Sec. 3.2.3, Sec. 3.2.4 and Sec. 3.2.5, respectively. Lastly, the trigger system of LHCb will be discussed in Sec. 3.2.6.

### 3.2.1 Luminosity at LHCb

In contrast to the general purpose detectors at the LHC, the LHCb experiment is not built to record data at LHC's design luminosity of  $\mathcal{L} = 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  but at an instantaneous luminosity of  $\mathcal{L} = 2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ . The reduced instantaneous luminosity has been chosen by the LHCb collaboration as the detector occupancy remains low and the radiation damage is reduced compared to larger values for the instantaneous luminosity [38]. An increased detector occupancy would be detrimental to the tracking efficiency and subdetectors which experience more radiation damage have worse resolution [120]. Both of these effects are especially disadvantageous when probing  $B$  physics which requires good resolution and a reliably high tracking efficiency when reconstructing the  $B$  vertex. In addition, a lower occupancy is preferable considering the slight dependence of the particle identification efficiency on the occupancy [122]. A further consequence of the instantaneous luminosity used at LHCb is that bunch crossings (events) are dominated by a single  $pp$  interaction as shown in Fig. 3.3. This follows from the average number of  $pp$  interactions in visible interactions ( $\mu$ ) per bunch crossing (pile-up,  $P$ ) being reduced at lower instantaneous luminosities. A visible interaction is defined as having at least two charged particles with sufficient measurements (hits) in the VELO and the T1-T3 stations to be reconstructible. The instantaneous luminosity, the number of average visible  $pp$  interactions per bunch crossing and the pile-up are connected, as the pile-up is defined as

$$P = \frac{\mu}{1 - e^{-\mu}},$$

and  $\mu$  is given by [123]

$$\mu = \frac{\mathcal{L} \sigma_{\text{tot}}}{f N_{\text{bunches}}},$$

where  $\sigma_{\text{tot}}$  is the total cross section at the LHC,  $f$  is the bunch crossing frequency and  $N_{\text{bunches}}$  is the number of bunches per beam in the LHC. At the design luminosity of LHCb, the pile-up is  $P \approx 2.1$ . In Fig. 3.3 the probability of the number of interactions ( $n$ ), ranging from 0 to 5, at a given instantaneous luminosity is shown using the continuous counterpart of the Poisson distribution [124]. As the probability for a larger number of interaction is located at larger values for  $\mathcal{L}$ , the pile-up increases with the instantaneous luminosity.

There are several options to reduce the instantaneous luminosity [125] at LHCb while ensuring the design luminosity at the other interaction points. One possibility is to change the focus of the beam at the LHCb interaction point, namely by increasing either of the transverse beam sizes,  $\sigma_x$  or  $\sigma_y$ . The focus of the beam often is also given as the  $\beta$  function, which is connected to the beam size by the formula

$$\epsilon \beta = \sigma_x \sigma_y,$$

where  $\epsilon$  is the emittance, which is a conserved quantity following Liouville's theorem [126]. The value for the  $\beta$  function at the interaction point is referred to as  $\beta^*$ . At LHCb the value for  $\beta^*$  is typically between 3 m and 10 m, whereas ATLAS and CMS typically use  $\beta^* = 0.4$  m.

Another option is the reduction of the overlap of the crossing beams, which reduces the effective number of particles ( $N_{1,2}$ ) entering the calculation of the instantaneous luminosity in Eq. 3.1. This can be done either by changing the crossing angle, often done with crab cavities and, therefore, called crab-crossing [127], or by introducing a transverse offset.

At LHCb all three methods are used in conjunction with each other to ensure the instantaneous luminosity to be at the same level for a fill during data taking. This is commonly referred to as luminosity leveling [128]. The transverse offset and the cross angle are continuously changed to increase the overlap of the bunches and counteract the depleting number of protons in a bunch. Additionally,  $\beta^*$  starts at a large value at LHCb to decrease the peak luminosity but during a fill the  $\beta^*$  value is decreased once the collisions deplete the number of particles in the bunches in order to increase the peak luminosity again. Changing the  $\beta^*$  value is also the only possible avenue to ensure the same level of luminosity once there is no transverse offset between the beams and they are colliding head-on.

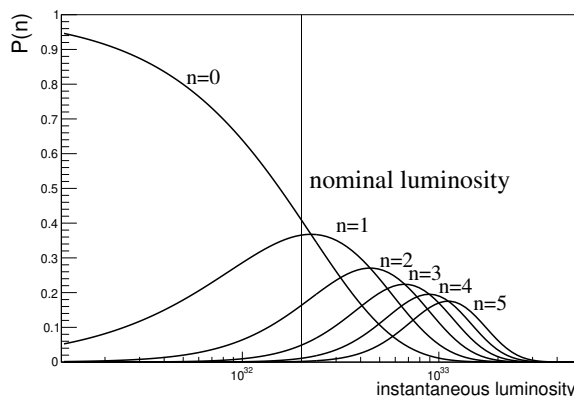


Figure 3.3: The Poisson probability of a given number of interactions ( $n$ ) is given as a function of the instantaneous luminosity. The instantaneous luminosity corresponding to the most probable value for a specific number of interactions increases with the number of interactions.

### 3.2.2 Tracking at LHCb

The tracking system at LHCb has two main objectives. For one, it has to precisely measure the momenta of particles traversing the detector. This is achieved by leveraging multiple tracking systems over a length of approximately 10 m to determine the bending in the magnetic field provided by a warm dipole magnet. These tracking systems are the VERtex LOcator [129] (VELO), the Tracker Turicensis [38, 130] (TT), sometimes also called trigger tracker, located before the magnet and three tracking stations called T1, T2 and T3 downstream of the magnet.

The other main objective is the precise measurement of track coordinates in the area surrounding interaction point. This area is especially interesting in the context of  $B$ -physics as  $B$  hadrons have significant lifetimes and are highly boosted along the direction of the beam. As such  $B$  hadrons travel on average a few mm before the decay. For example a  $B^0$  meson has an average flight distance of 7 mm [131]. Therefore, both the vertex at which  $B$  hadrons are produced and at which they decay are of particular interest. This means that the point where the proton-proton collision takes place, called the primary vertex (PV), as well as the secondary vertex (SV) at which the  $B$  hadron decays, need to be resolved with high precision. The resolution of the PV measured in LHCb depends on the number of tracks originating from

it. For example, for a PV with 25 tracks a resolution of  $13\text{ }\mu\text{m}$  in  $x$ - and  $y$ -directions, and  $71\text{ }\mu\text{m}$  in  $z$ -direction can be achieved [132], well below the average decay length of a  $B$  meson. Given the precise measurement of the PV and the SV, the displacement of the SV can be used to differentiate  $B$  hadron decays from decays of short-lived particles.

### Vertex Locator

The VERtEX LOcator [129] is a silicon strip detector and the detector closest to the interaction point. Its positioning is chosen to precisely measure the PV and SVs. The VELO consists of two movable halves which are at a distance of 7 mm from the beam during data taking, once the proton beam is considered stable, with the first active strip at a distance of 8.2 mm. If the beam is considered unstable, for example during the injection, the halves are move apart from each other to protect the detector from being hit with the LHC beam. As the half aperture of the LHCb beam is approximately 27 mm during injection [129, 133], the distance between the two halves is approximately 6 cm during ramp-up.

Each VELO half has 21 silicon modules instrumented with sensors in radial ( $R$ ) and azimuthal ( $\phi$ ) directions along the beam axis. The VELO is built from silicon detectors, as they are radiation hard, provide good resolution and are fast. In Fig. 3.4 the 21 modules of the VELO and the VELO in both the fully closed and fully open positions are shown. There are two additional stations on each side of the beam axis which can be used to veto events in the hardware trigger with multiple interactions by measuring the total backward charged track multiplicity [38]. These additional stations are only instrumented with radial sensors.

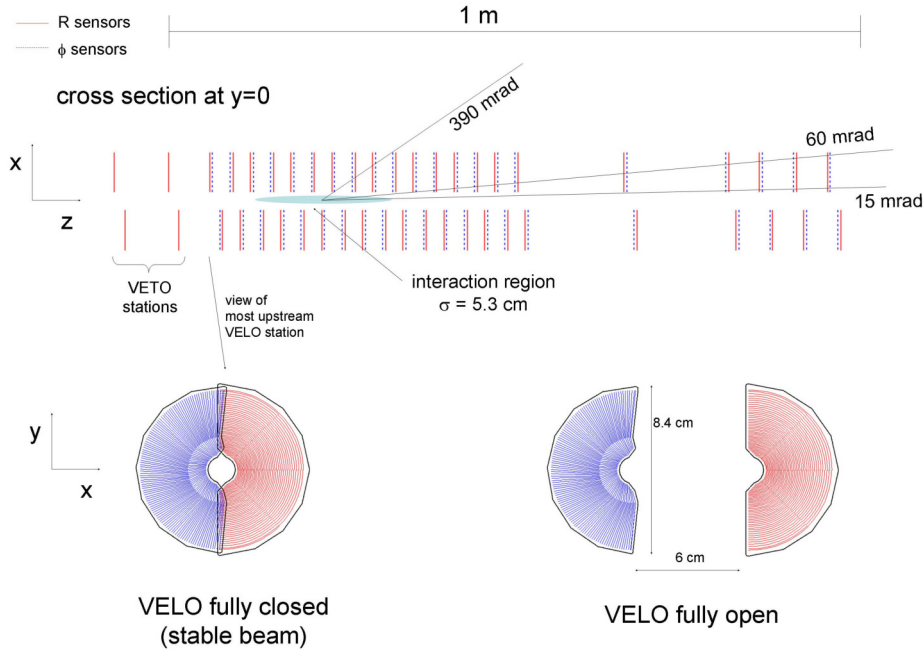


Figure 3.4: (Top) The positioning of the 21 silicon modules of the VELO along the beam axis as well as the (bottom left) the fully closed and (bottom right) fully open position of each half of the VELO. Taken from Ref. [38].

Each  $R$  sensor is a concentric semi-circle with its centre located at the nominal LHC beam



spot. They are built from four segments, each of which covers an angle of  $45^\circ$  and consists of 512 strips. The pitch of each strip increases linearly with the radial distance from the centre. At the innermost radius the pitch is  $40\text{ }\mu\text{m}$  and increases to  $101.6\text{ }\mu\text{m}$  at the outermost radius [38]. The  $R - \phi$  geometry of the VELO sensors is illustrated in Fig. 3.5.

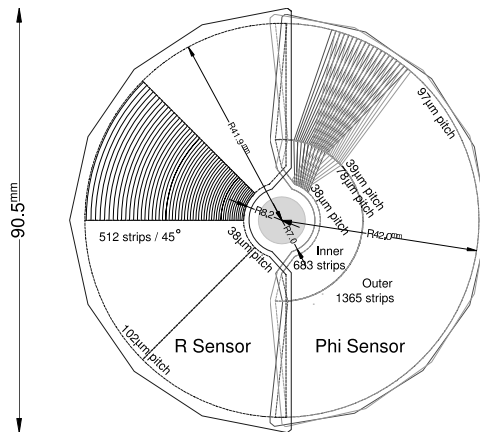


Figure 3.5: The geometry of the  $R$  and  $\phi$  sensors. For clarity, only a portion of the strips are shown. Taken from Ref. [38].

The  $\phi$  sensors are subdivided into an inner and an outer part with 683 and 1365 radially mounted strips, respectively. This partition reduces the occupancy of the strips and avoids the pitch becoming too large at the outer edge of the sensor. The inner part covers the region with radii smaller than  $17.25\text{ mm}$  [38]. The pitch of the  $\phi$  strips increases linearly from the inner edge to the outer edge, namely from  $37.7\text{ }\mu\text{m}$  to  $79.5\text{ }\mu\text{m}$  and from  $39.8\text{ }\mu\text{m}$  to  $96.9\text{ }\mu\text{m}$  for the inner and outer parts, respectively [38].

The resolution of the VELO depends on the strip pitch and of the projected angle. The best hit resolution of the  $R$  sensor is measured to be approximately  $4\text{ }\mu\text{m}$  at the optimal projected angle of  $8^\circ$  and the minimum pitch of  $40\text{ }\mu\text{m}$ . The resolution of the  $\phi$  sensor is comparable with that of the  $R$  sensor [134].

### Tracker Turicensis

The Tracker Turicensis (TT) [38, 130] is also a silicon strip detector located upstream of the LHCb dipole magnet. The TT is used to measure low momentum particles, which are bent out of the acceptance by the dipole magnet.

The TT consists of four detection layers of which the second and third layers are rotated around the  $z$ -axis with respect to the first and last layer at a stereo angle of  $-5^\circ$  and  $+5^\circ$ , respectively. The use of the stereo angle between the different layers allows not only the coordinate in  $x$  direction to be measured but also in  $y$  direction. Furthermore, the first pair of layers is separated from the second pair of layers by approximately  $27\text{ cm}$  along the beam axis [38]. Each of the TT layers is  $150\text{ cm}$  wide and  $130\text{ cm}$  high corresponding to an active area of  $8.4\text{ m}^2$  [38, 130]. The layout of the detection layer is shown in Fig. 3.6. The regions to the left and right of the beam pipe are covered with seven rows of full modules for the first two detection layers and eight rows for the second two detection layers and the region directly

above and below the beam pipe is covered by a single half module in each direction. A full module consists of 14 sensors and is assembled by joining two half modules end-to-end, where each half module covers half the height of the detector acceptance.

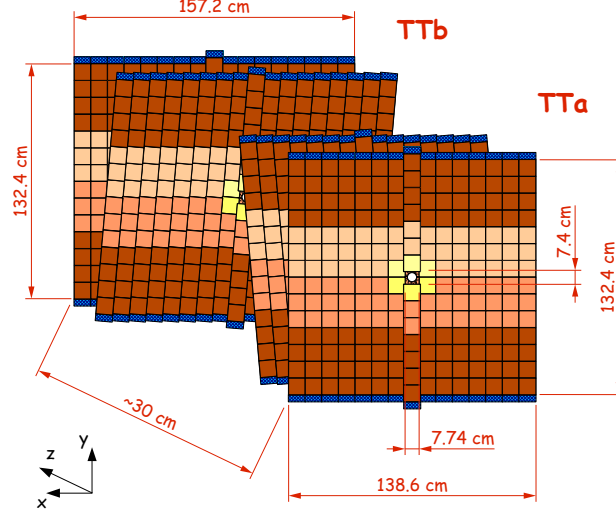


Figure 3.6: Schematic view of the layers of the TT station, taken from Ref. [135].

Each half module comprises seven sensors, which are readout with either two or three readout hybrids [136]. In each readout hybrid a four layer Kapton flex circuit hosting four ASIC frontend chips is used to readout the signals of the sensors. The hybrid also carries an 80-pin board-to-board connector which provides the control signals, low voltage and bias voltage to the half modules [38]. Each silicon sensor is 500  $\mu\text{m}$  thick, 9.64 cm wide and 9.44 cm long and carries 512 readout strips with a pitch of 183  $\mu\text{m}$  [38]. The single-hit spatial resolution of any sensor is approximately 50  $\mu\text{m}$  [132].

## Magnet

To be able to measure the momenta of charged particles in LHCb a warm dipole magnet [38, 137] is used to bend the tracks. The measurement of the momentum  $p$  relies on the Lorentz force connecting the bending radius and the magnetic field strength  $B$  via the formula  $p[\text{GeV}/c] = 0.3B[\text{T}]r[\text{m}]$  for the case of a particle with the charge  $\pm 1e$ . The magnetic field strength of the dipole used in LHCb has been measured along the  $z$ -axis and its maximal value is determined to be around 1.1 T, as can be seen in Fig. 3.7. This corresponds to a bending power of  $\int B d\ell = 4 \text{ Tm}$  for a track with a length of 10 m [38, 137]. To control potential systematic effects due to detector asymmetries, the magnetic field is flipped regularly between fills.

The shape of the magnetic field strength is constrained by two conflicting requirements of the LHCb experiment on the magnet: For a given precision on the bending radius  $r$ , the precision of momentum measurement is directly proportional to the magnetic field strength and, therefore, requires the largest magnetic field strength possible. In contrast, the Ring Imaging Cherenkov detectors require the magnetic field strength to be below 3 mT for its photon detectors to operate at full efficiency. With shielding against the magnetic field, the fringe field strength is allowed to be approximately 60 mT [38]. From these two requirements the magnetic field strength was designed to be as small as possible close to the Ring Imaging Cherenkov detectors but

as large as possible near the tracking system. As the VELO is located before the first Ring Imaging Cherenkov detector, the magnetic field is practically negligible in the VELO and a momentum measurement only with hits located there is nearly impossible.

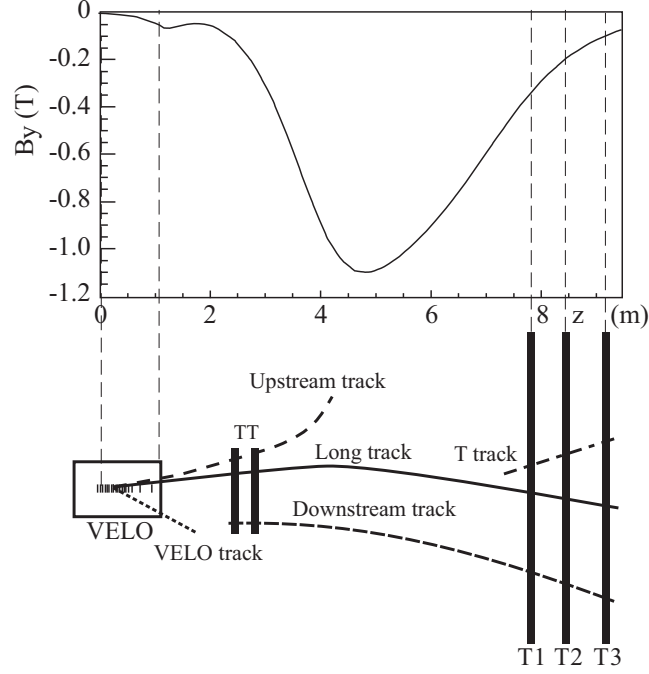


Figure 3.7: The main component of the magnetic field strength as a function of the  $z$  position and the different track types used in LHCb. Taken from Ref. [121].

### Tracking Stations T1, T2 and T3

The tracking stations T1, T2 and T3 are located downstream of the magnet. Each of the stations is divided in an inner [38,130] and outer part [38,138]. The two parts have different detector designs.

The inner tracker (IT) covers the region close to the beam pipe and is designed in analogy to the TT. It is also a silicon tracker with each station consisting of four layers with the same arrangement as found in the TT and an active area of  $4.0 \text{ m}^2$  [130]. It covers a cross shape which is  $125.6 \text{ cm}$  wide and  $41.4 \text{ cm}$  high [130], as shown in Fig. 3.8,. A single layer consists of seven detector modules each to the left and right of the beam pipe and seven modules above and below the beam pipe. The modules to the left and right are two sensor modules while the top and bottom modules are one sensor modules. Each sensor is  $7.6 \text{ cm}$  wide and  $11 \text{ cm}$  long. They consist of 384 readout strips with a pitch of  $198 \text{ }\mu\text{m}$ . The sensors in the two sensor modules have a thickness of  $410 \text{ }\mu\text{m}$  and the sensors in the one sensor modules thickness is  $320 \text{ }\mu\text{m}$  [38,130]. The modules are readout with a readout hybrid similar to the one used in the TT. In contrast to the four frontend ASICs used in the TT, the IT readout hybrid uses three front-end ASICs [38]. Their single-hit spatial resolution is about  $50 \text{ }\mu\text{m}$  [132], similar to the resolution of the TT.

In contrast to the IT, the outer tracker (OT) follows a different detection principle. Here, a drift-time measurement with gas-tight straw tube modules is performed instead of using silicon sensors. The different approach is chosen as instrumenting the large active area of one of the

three OT stations of  $5971 \times 4850 \text{ mm}^2$  or roughly  $28.96 \text{ m}^2$  [38, 138] with silicon detectors is very expensive while straw tubes are a cheaper option. Like silicon detectors, the straw tube detectors are radiation hard and fast but have worse resolution than silicon detectors. Analogous to the TT and IT, each OT station has four layers with the second and third layers being rotated at an angle of  $-5^\circ$  and  $+5^\circ$  around the  $z$ -axis with respect to the first and last layer. Each layer comprises 14 long and 8 short straw tube modules. The short modules are located above and below the beam pipe, whereas the long modules are located to the left and right of the beam pipe. The long modules consist of 256 single drift tubes, arranged in two staggered layers, while the short modules have 128 single drift tubes. Each drift tube has a pitch of 5.25 mm [38]. With these drift tubes, a maximum drift time of about 35 ns and a drift-coordinate resolution of  $205 \text{ } \mu\text{m}$  [132] can be achieved.

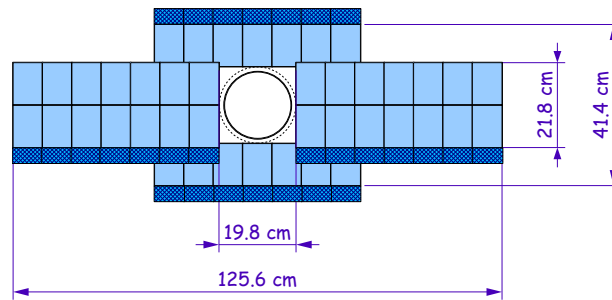


Figure 3.8: Layout of the first layer in T2, the dimensions are given in cm and refer to the sensitive surface covered by the Inner Tracker. Taken from Ref. [130].

### Track reconstruction algorithms

The measurements performed with the VELO, TT and the tracking stations T1, T2 and T3 are used to reconstruct the tracks of charged particles traversing the detector. It is advantageous to define different track types considering that some particles only leave hits in some subdetectors because they are either bent out of the acceptance, decay in flight or originate from long-lived particles. There are five different types of tracks based on the subdetectors contributing to the reconstruction of the tracks. These are illustrated in Fig. 3.7. The tracks reconstructed only from hits left in the modules of the VELO are referred to as VELO tracks and analogously tracks built only from hits in the tracking stations T1, T2 and T3 are referred to as T-tracks. In addition to tracks only built from one subdetector, there are the upstream and downstream tracks which are reconstructed from two different subdetectors. Upstream tracks consist of hits in the VELO and the TT, and typically originate from low momentum tracks bent out of the detector acceptance by the dipole magnet. In contrast, downstream tracks are reconstructed from hits in the TT and the tracking stations T1, T2 and T3. These tracks can originate from the decay products of long-lived particles such as the  $K_s^0$  meson. The last track type are long tracks. These are built from all three subdetectors and have the best momentum resolution due to the length of the track. They are, therefore, most commonly used in physics analyses.

All tracks are based on initial track candidates (seeds) [139–150] in the VELO regions or the T stations. These regions are chosen as these are the regions with low magnetic fields, meaning that the track candidates are likely to be straight. The VELO and T-tracks reconstructed from a seed are then used to form the other three types of tracks.

Two different algorithms can be used to reconstruct long tracks, **forward tracking** or **track matching**. When performing **forward tracking**, the starting point is the VELO track which is extrapolated as a straight line into the T stations [151]. There, a T station hit is connected with the track and additional hits in the TT are optionally added from in a window around the extrapolated track to form the long track. In contrast, in the **track matching** approach the VELO and T-tracks are extrapolated to the bending plane inside the magnet. If the two tracks match with each other, additional hits in the TT close to the extrapolated track are added to form the long track. A track needs to have a minimum number of hits in each subdetector to be considered reconstructible. Furthermore, tracks in simulated events are considered to be reconstructed successfully if more than 70% of the hits originate from a real particle. The efficiency of the tracking algorithm is above 96% in the momentum range of  $5 \text{ GeV}/c < p < 200 \text{ GeV}/c$  [38, 132, 151]. If less than 70% hits used to build a track originate from a real particle, they are considered to be ghost tracks.

The momentum of the long track is determined by performing a fit using a Kalman filter [152]. The use of a Kalman filter accounts for multiple scatterings and, therefore, corrects for the  $dE/dx$  energy loss improving the momentum resolution [38]. From this fit, the track quality can be inferred from the  $\chi^2_{\text{track}}$  and the number of degrees of freedom in the fit. The resolution of the track depends on its momentum and is found to be approximately 5 per mille for particles with a momentum below 20 GeV/c and approximately 8 per mille for particles with a momentum around 100 GeV/c [38, 132, 151].

### 3.2.3 Ring Imaging Cherenkov Detectors

In LHCb two Ring Imaging Cherenkov detectors [38, 153], called RICH1 and RICH2, are used to distinguish different types of charged particles, such as pions and kaons. The first RICH detector (RICH1) is located upstream of the magnet and before the TT. In contrast, the second RICH detector (RICH2) is located downstream of the magnet after the tracking stations T1-3. A schematic view of each of the RICH detectors is shown in Fig. 3.9. The two RICH detectors are filled with gasses to induce the emission of photons if a charged particle traverses the detector (Cherenkov radiation) [154]. These photons are emitted with a given angle (Cherenkov angle), defined relative to the path of the particle. The Cherenkov angle ( $\theta_C$ ) is given by the formula

$$\cos \theta_C = \frac{1}{n\beta},$$

where  $\beta = \frac{v}{c}$ ,  $v$  is the velocity of the charged particle,  $c$  is the speed of light in vacuum and  $n$  is the refractive index of the gas. The emitted photons are focused on a detector plane instrumented with hybrid photo detectors (HPD) using a combination of spherical and flat mirrors. The HPDs are placed in a region outside of the detector acceptance where the fringe magnetic field strength is approximately 50 mT. Considering that the HPDs operate at full efficiency in  $B$ -fields below 3 mT, magnetic shield boxes are used to attenuate the external fields by a factor 20 or larger to ensure maximum efficiency [38]. To further maximise the information gained from the two RICH detectors, they are filled with different gases. RICH1 is filled with  $C_4F_{10}$  ( $n = 1.0014$  [153]) and additionally had a silica aerogel ( $n = 1.03$  [153]) radiator in Run 1 which was removed in Run 2, whereas RICH2 is filled with  $CF_4$  ( $n = 1.0005$  [153]). Due to the different gas radiators, the momentum coverage of the two RICH detectors differs. While the covered momentum range depends on the mass of the particle,

RICH1 generally covers the low momentum range between 1 and 60 GeV/c, whereas RICH2 generally covers the high momentum range between 15 and 100 GeV/c. The aerogel was used in Run 1 due to its high refractive index and low density, which helps with the identification of particle with low momenta of a few GeV/c [38]. However, due to new requirements of the event reconstruction in Run 2, the aerogel was removed, which reduces the number of photon candidates in the RICH reconstruction by roughly half and, therefore, speeding up the reconstruction significantly [155]. Furthermore, the angular momentum range covered by the two detectors also differs. The RICH1 detector covers the full LHCb acceptance between 25 to 300 (250) mrad in the horizontal (vertical) plane, while RICH2 covers a more limited range between roughly 15 and 120 (100) mrad in the horizontal (vertical) plane [153].

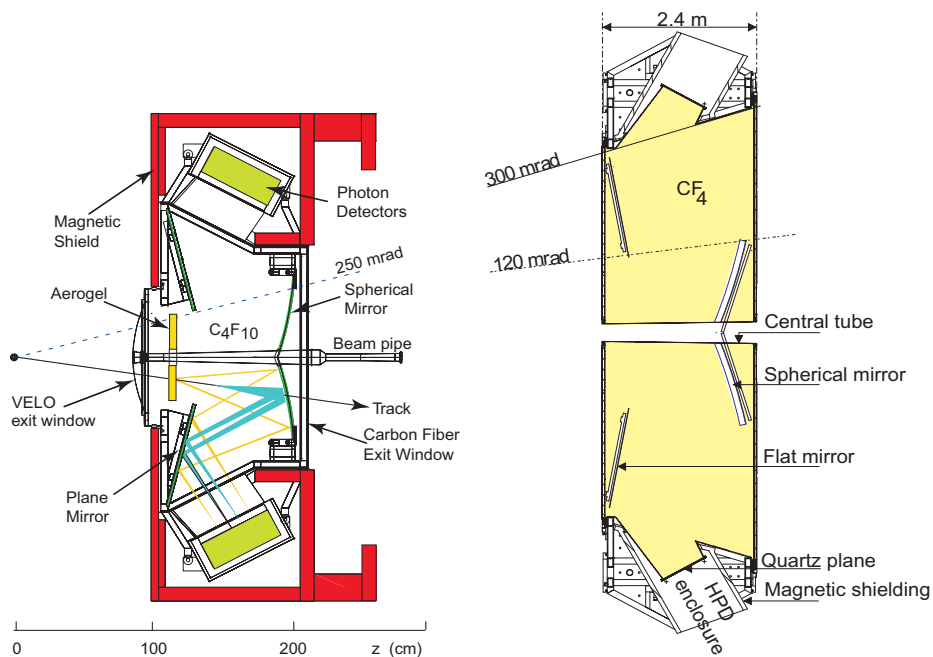


Figure 3.9: (Left) A schematic view of the RICH1 from the side and (right) a schematic view of RICH2 from above. Both are taken from Ref. [38].

### 3.2.4 Calorimeters

The calorimeters in LHCb [38, 156] are located downstream of the RICH2 detector and are designed with three main goals: To perform an energy measurement, to contribute to the particle identification and to contribute to the trigger.

While the momenta of electrons and charged hadrons can be measured with the tracking system, the energy can directly measured with the calorimeters. In contrast, an uncharged particles does not leave hits in the detector and its momentum can not be measured by the tracking system. Therefore, for these particles an energy measurement needs to be performed with the calorimeter from which the momentum can be inferred. The energy is determined using alternating layers of an scintillator plane and an absorber plane. The scintillators are used to measure an electromagnetic shower originating from the absorber plane. Photons emitted in the scintillator are transmitted to photo-multiplier tubes using wavelength-shifting fibers, where they are detected. Due to the different interactions length required to stop hadrons

compared to electrons and photons, there is a dedicated calorimeter in LHCb for both types of particles. Electrons and photons typically deposit most or all of their energy in a specially designed electromagnetic calorimeter (ECAL), whereas hadrons typically are not stopped before the hadronic calorimeter (HCAL). The information on the location and the distribution of the deposited energy can, therefore, be leveraged to separate electrons and photons from hadrons. To distinguish electrons from photons in the hardware trigger, another detector, the scintillating pad detector (SPD), is used in which only charged particles deposit energy. A thin lead absorber and another scintillating pad detector, called the preshower detector (PS), is then used to distinguish electrons from charged light hadrons, primarily from pions, discussed in detail below. The shower shapes used in the particle identification are illustrated in Fig. 3.10. Furthermore, all these bits of information are combined in the hardware trigger stage of LHCb. A decision based on for example the deposited transverse energy is made within 4  $\mu\text{s}$  of the interaction [156, 157].

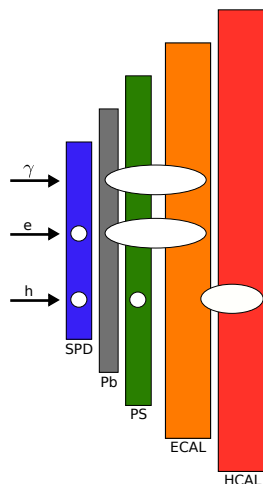


Figure 3.10: Typical shower shapes left in the calorimeters of photons, electrons and charged hadrons shown as white ellipses. The different sub detectors, the lead absorber and the calorimeters are not to scale.

### Scintillating pad detector and preshower detector

The scintillating pad and preshower detectors are both types of scintillating pad detectors and separated by a 15 mm thick lead absorber. The spatial separation allows the separation of photons from electrons and light hadrons as photons do not interact with the SPD but can leave hits in the PS if they convert into pairs of charged particles in the absorber. Given the different shower shapes in the PS detectors, as shown in Fig. 3.10, electrons can be distinguished from light hadrons using the information of the PS.

The two detectors are built from 12032 rectangular scintillating pads each and have a total sensitive area of 7.6 m in width and 6.2 m in height corresponding to 47.12 m<sup>2</sup> [38]. The PS and SPD detector planes are subdivided into a left and a right half, each of which is further subdivided into three regions. They are called the inner, the middle and the outer sections and shown in Fig. 3.11. In these sections, there are 3072, 3584 and 5376 square-shaped cells, respectively, with side lengths of 4.04 cm, 6.06 cm and 12.12 cm cells [38, 156].

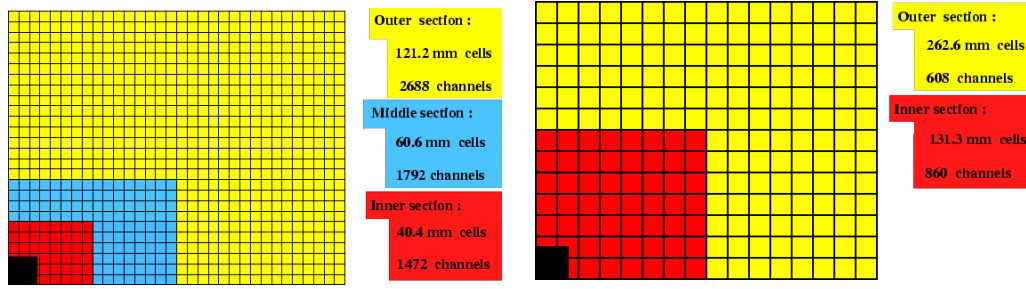


Figure 3.11: Lateral segmentation of the (left) SPD/PS and ECAL as well as the (right) HCAL. Taken from [156].

### Electromagnetic Calorimeter

The electromagnetic calorimeter consists of alternating layers of 4 mm scintillator material, 120  $\mu\text{m}$  thick, white, reflective TYVEK paper and 2 mm lead are used [38]. This choice allows for a highly reliable and reasonably radiation hard detector that also has a fast time response and reasonably good energy resolution. The length of the ECAL has been chosen designed to correspond to 25 radiation length ( $X_0$ ) and 1.1 hadronic interaction lengths ( $\lambda_I$ ). The TYVEK paper is used to improve the light reflection into the scintillating material. The energy resolution has been determined [158] to be

$$\frac{\sigma_E}{E} = \frac{(9.0 \pm 0.5)\%}{\sqrt{E}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.003}{E \sin \theta}$$

where  $E$  is the energy of a particle in GeV, the  $\oplus$  symbol implies addition in quadrature and  $\theta$  is the angle between the beam axis and a line from the LHCb interaction point and the centre of the ECAL cell. The ECAL has the same segmentation as the SPD and PS detectors, shown in Fig. 3.11.

### Hadronic Calorimeter

The last of the calorimeters is the hadronic calorimeter. It is a sampling calorimeter built from iron as the absorber and scintillating tiles as the active material, which are oriented parallel to the beam. The HCAL covers 5.6 hadronic interaction length across its length. In contrast to the other three calorimeters, each HCAL half is only segmented into an inner and an outer section, as shown in Fig. 3.11. The HCAL cells are square shaped and have a side length of 13.13 cm in the inner and 26.26 cm in the outer section [156].

Using SPS test beam data, the energy resolution has been determined [158] to be

$$\frac{\sigma_E}{E} = \frac{(67 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%.$$

## 3.2.5 Muon chambers

In contrast to hadrons, electrons and photons, muons traverse the calorimeters without depositing large amounts of energy. Therefore, a dedicated detector, the muon system [38, 159–161], has been built and installed to provide information for the trigger and particle identification and give an additional measurement of the momenta of muons. The muon system measures



the momenta of muons based on multi-wire proportional chambers (MWPC) [162–164] and triple gas electron multiplier detectors [165, 166] (triple-GEM detectors). The information on the momentum can be obtained in less than 25 ns and is used in the hardware trigger stage of LHCb.

The muon system consists of five stations, referred to as M1 to M5. The first muon station (M1) is positioned between the RICH2 detector and the calorimeters, while the other four muon stations (M2 to M5) are positioned downstream of the calorimeters. This choice has been taken in order to increase the length of the muon track. The increased length of the muon track corresponds to an improvement in the momentum resolution of 30% [167]. Between each of the M2 to M5 stations 80 cm thick iron absorbers are placed to filter any particle species other than muons [38]. These absorbers also place a threshold on the minimum momentum of the muons needed to traverse all five muon stations, namely  $p_{\min}(\mu) \approx 6 \text{ GeV}/c$  [38]. Considering that electrons and hadrons are stopped on average before the muon station M2, it has also an essential role when identifying muons.

All five muon stations are divided into four regions, called R1 to R4. The  $x$  and  $y$  dimensions of the segments R1:R2:R3:R4 each scale as 1:2:4:8 and the areas scale as 1:4:16:64 as can be seen in Fig. 3.12. This segmentation was chosen in order to approximately have the same flux in each region. However, due to its positioning, the region R1 of the muon station M1 is expected to have the largest particle flux. It has to be capable to handle rates of up to 500 kHz/cm<sup>2</sup> and, therefore, has to fulfill especially stringent requirements on radiation hardness. For this reason, triple-GEM detectors are used in this region as they are designed to not show any aging effects after 10 years of LHCb operation. The MWPCs, used in all other regions and muon stations, would only be able to achieve this when decreasing the gain and, therefore, sacrificing signal-to-noise ratio and in turn sensitivity.

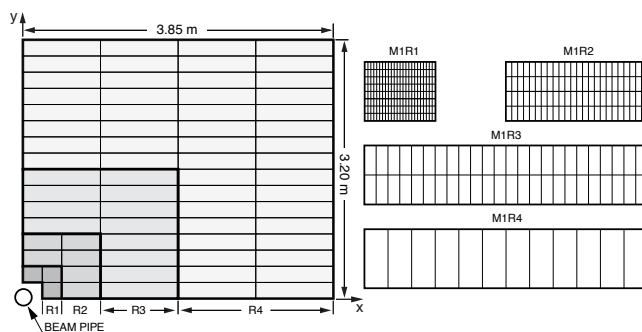


Figure 3.12: On the left, the front view of a quadrant of a muon station is shown. Each rectangle represents one chamber. On the right, the division of the four chambers into logical pads belonging to the four regions of station M1 is shown. In each region of stations M2-M3 (M4-M5) the number of pad columns per chamber is double (half) the number in the corresponding region of station M1, while the number of pad rows per chamber is the same. Taken from [38].

### 3.2.6 Trigger

The LHC provides beams with a bunch crossing rate of 40 MHz and the total cross section at the LHC at the centre of mass energies of  $\sqrt{s} = 7, 8$  and 13 TeV has been measured [168–170]

to be  $\sigma_{\text{tot}}^{7 \text{ TeV}} = (95.35 \pm 1.36) \text{ mb}$ ,  $\sigma_{\text{tot}}^{8 \text{ TeV}} = (96.07 \pm 0.93) \text{ mb}$  and  $\sigma_{\text{tot}}^{13 \text{ TeV}} = (110.6 \pm 3.4) \text{ mb}$ , respectively. Despite the bunch crossing rate of 40 MHz, the rate of visible interactions in LHCb is reduced to approximately 10 MHz at the instantaneous luminosity of  $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  as not all bunch crossings lead to a visible collision [157]. Furthermore, out of the 10 MHz of visible events, only approximately 100 kHz contains at least one  $b\bar{b}$  pair and 15% of those have all their final state particles in the LHCb acceptance [157]. Considering this, it is crucial to efficiently select heavy flavour decays to save to tape whilst rejecting visible interactions from non-heavy flavour decays and random combinations of tracks. For this reason a trigger system has been developed for LHCb [157], which also needs to account for computational restraints. To accommodate for the computational restraints, the trigger system is split into two stages, a hardware stage and a software stage. The two stages are also referred to as the Level 0 (L0) trigger and the High Level Trigger (HLT), respectively. The structure of the trigger used in LHCb is illustrated in Fig. 3.13. A major difference between the two stages is that the decision making of the L0 trigger operates synchronously with the bunch crossing frequency, while the HLT is run asynchronously on a processor farm, called the event filter farm [171]. This choice has been made in order to avoid the need to readout the full detector at the rate of visible interactions of 10 MHz. At this rate only the calorimeters and the muon stations as well as parts of the VELO are readout for the L0 trigger.

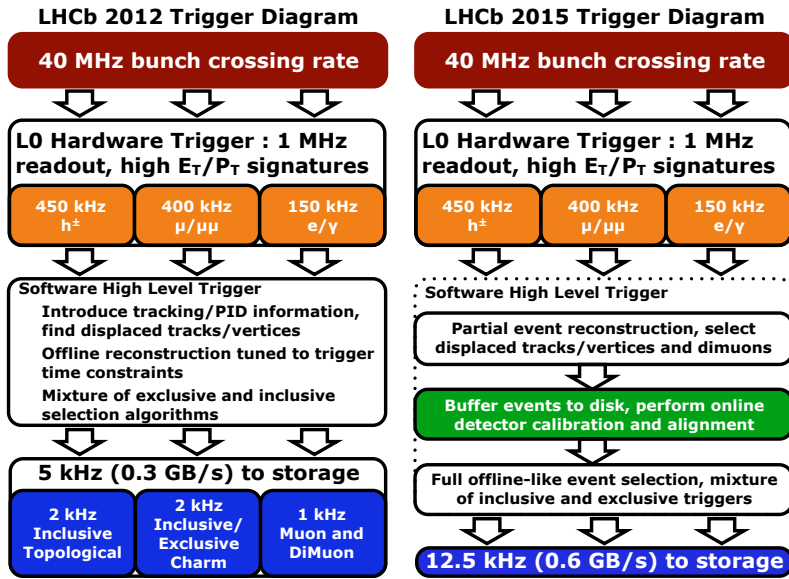


Figure 3.13: Overview of the trigger schemes used in Run 1 and Run 2. Taken from Ref. [172].

The main purpose of the L0 trigger is the reduction of the bunch crossing rate of 40 MHz to the rate of approximately 1 MHz at which the whole detector can be read out [38]. The L0 trigger is able to use information of the pile-up system located in the VELO, the calorimeters and the muon system to make a decision. Due to the large mass and boost of a  $B$  meson, its daughter particles likely have a large transverse momentum ( $p_T$ ). Therefore, the calorimeter and the muon systems are used to estimate the transverse momentum on custom electronics. In the calorimeter trigger, clusters are formed from  $2 \times 2$  cells and the transverse energy ( $E_T$ ) is calculated. Furthermore, information from the SPD, PS, ECAL and HCAL is used to assign an initial particle hypothesis to the cluster. At this stage, the calorimeter is used to differentiate between electrons, photons and hadrons. The electron, photon and hadron clusters with the

largest transverse energy is triggered, if specific, running period dependent thresholds are exceeded. For hadrons the typical threshold is at  $E_T \gtrsim 3.5$  GeV and for electrons and photons it is at  $E_T \gtrsim 2.5$  GeV [173,174]. In contrast, muons are triggered using the muon system. There are single and dimuon triggers, typically requiring a  $p_T \gtrsim 1.5$  GeV/ $c$  and  $\sqrt{p_T^1 p_T^2} \gtrsim 1.5$  GeV/ $c$ , respectively [173,174]. The transverse momentum is estimated by using the slope in the first two muon stations [167]. In Run 1 the rate of events passing the L0 was 0.87 MHz [173] compared to 1 MHz in Run 2 [174],

Events passing the L0 trigger enter the software trigger stage. As the HLT is a software based trigger, the selection threshold for the triggers can be changed during data taking using the Trigger Configuration Key (TCK). The TCK can be adjusted to utilize, for example, the remaining buffer size for the HLT as best as possible. The HLT is subdivided into two parts, called HLT1 and HLT2. In HLT1, the tracks and primary vertices (PV) are reconstructed using information from the VELO and tracking stations. This partial event reconstruction is used to verify the L0 decision and further select events containing particles with large momenta and a large impact parameter with respect to any PV. Such signatures are consistent with long-lived  $B$  or charm hadrons. The HLT1 is able to reduce the rate to approximately 43 kHz in Run 1 [174] and approximately 110 kHz in Run 2 [174]. In Run 2, the remaining events are then written to a buffer, at which point detector calibrations and alignments are performed. In the HLT2 the information of all subdetectors is used to fully reconstruct the event. This reconstruction slightly differs from the final offline reconstruction in that no Kalman filter [152] is used in the online reconstruction as this would be too CPU intensive. Not using a Kalman filter does not take the  $dE/dx$  energy loss correctly into account and degrades the momentum resolution [38]. The HLT2 selects events with  $B$  decays either inclusively or exclusively based on information on the invariant mass of particle combinations and the secondary vertex displacement, amongst other criteria. Exclusive selections are mostly used for  $B$  hadrons decaying to hadrons. In contrast, inclusive selections are frequently used to identify a broad spectrum of multi-body decay topologies of  $B$  hadrons or displaced dimuon decays. In Run 1 the HLT2 output rate was approximately 5 kHz [175,176] and 12.5 kHz in Run 2 [174].

# Chapter 4

## Analysis tools

In this Chapter, important tools and concepts for the analysis will be introduced. These are the multivariate analysis in Sec. 4.1, the maximum likelihood technique for parameter estimation in Sec. 4.2, largely based on Ref. [177], and the *sPlot*-technique [178] to statistically subtract a background contribution from a sample in Sec. 4.3.

### 4.1 Multivariate analysis

A data set containing events including both signal decays and background sources can be separated into its components using variables such as the invariant mass of composite particles such as the  $\phi$  and  $B_s^0$  mesons. While placing a selection on a single variable can be efficient, the separation power can often be improved by combining variables into a single discriminator using multivariate analysis. One example where this is necessary are events originating from random combinations of tracks (combinatorial background). For this type of background, there is no single variable strongly discriminating signal from background. Therefore, boosted decision trees (BDT) [179–181] are used in this thesis to combine several input variables in a discriminator and thereby increase the signal to background separation for the combinatorial background.

To understand BDTs, two core concepts are important, decision trees [182] and boosting [183, 184]. Decision trees can be classified as supervised machine learning techniques and follow the general concept of a selection based on placing requirements on a combination of variables. A schematic overview of a decision tree is given in Fig. 4.1. Each selection of a decision tree is performed at a node. They aim to separate signal from background components and create two branch in the process. The final branches (leaves) are classified either as signal or background leaves depending on the category the leaf contains the most of. For a set of events with weights  $w_{s,b}$ , the purity ( $P$ ) of a branch is given as  $P = \frac{\sum_s w_s}{\sum_s w_s + \sum_b w_b}$  [181]. To derive the exact threshold at each node a training step has to be performed. In the training the separation power of each node is maximised using a parameter called the *Gini*-index. The *Gini*-index is defined as  $Gini = (\sum_e w_e) P(1 - P)$ , where the sum is performed over all events  $e$  relevant to that node. The *Gini*-index is by construction 0 for a leaf containing only signal or background events. The separation power can then be maximised by minimizing the criterion

$$\Delta = Gini_{\text{node}} - Gini_{\text{left branch}} - Gini_{\text{right branch}}.$$

While decision trees are a powerful tool to classify signal and background events, they are

also sensitive to small fluctuations in the training sample. To avoid this sensitivity to small fluctuations, boosting was introduced. When performing boosting, a decision tree is trained with unweighted events. Events mis-classified by the decision tree are then given higher weights (boosted) and a second decision tree is trained. A known signal event is considered mis-classified if it lands on a background leaf and vice versa. The training of new decision trees is repeated multiple times, typically  $\mathcal{O}(1000)$  times. After the training of the decision trees, each event is determined to be either a signal- or background-like event. For this purpose a score is assigned by taking the difference of the number of trees the event lands on a signal leaf and on a background leaf. As a larger total score corresponds to an event being more signal-like, the total score is used to separate signal and background events. To avoid any biases when determining the threshold on the total score data set is split into three independent samples, a training sample, a testing sample and an evaluation sample. This can be done using the  $k$ -fold cross-validation technique [185], where a data set is split into  $k$  samples (folds) of which up to  $k - 2$  are used for training, one is used to test for overtraining and to determine the final threshold on the discriminator and the last is used for evaluation. To avoid reducing the effective data set size, one can construct  $k$  BDTs, each of which uses the same  $k$  folds but uses a unique composition of the folds for the training, testing and evaluation. This way, each event is used in the training of up to  $k - 2$  BDTs, testing of one BDT and is kept or rejected according to the last independent BDT.

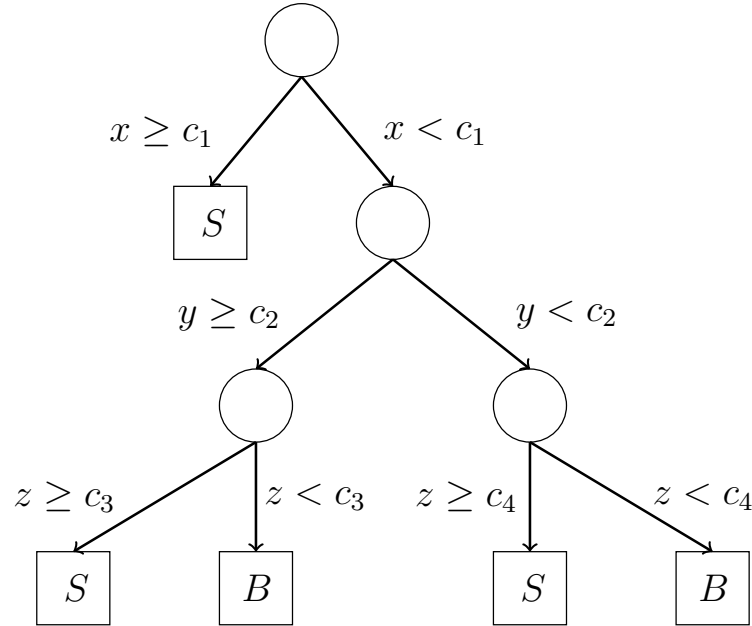


Figure 4.1: Schematic overview of a decision tree, in which the signal ( $S$ ) is separated from background ( $B$ ) using three variables ( $x$ ,  $y$  and  $z$ ) with four different selections ( $c_1$ ,  $c_2$ ,  $c_3$  and  $c_4$ ).

## 4.2 Maximum likelihood technique for parameter estimation

A probability density function (PDF),  $\mathcal{P}$ , relating  $N$  independent measurements ( $\vec{x}$ ) and  $j$  physics parameters ( $\vec{\lambda}_{\text{physics}}$ ), describes the probability density distribution of events for a given set of values of  $\vec{\lambda}_{\text{physics}}$  or the probability density distribution for a set of parameters given the measurements  $\vec{x}$ . These two concepts are connected by Bayes' theorem [186] as

$$\mathcal{P}(\vec{\lambda}_{\text{physics}}|\vec{x}) = \frac{p(\vec{\lambda}_{\text{physics}})\mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics}})}{p(\vec{x})},$$

where  $p(\vec{x})$  is the prior probability to observe the measurements  $\vec{x}$ ,  $p(\vec{\lambda}_{\text{physics}})$  is the prior probability to observe the physics parameters  $\vec{\lambda}$ ,  $\mathcal{P}(\vec{\lambda}_{\text{physics}}|\vec{x})$  is the conditional probability of finding the values  $\vec{\lambda}_{\text{physics}}$  given the measurements  $\vec{x}$  and  $\mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics}})$  is the conditional probability of finding the measurements  $\vec{x}$  given the values  $\vec{\lambda}_{\text{physics}}$ .

The expected distribution of events for a given set of physics observables is determined from the conditional probability  $\mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics}})$ . This is used to generate pseudoexperiments used to estimate systematic uncertainties or to validate the fit behaviour. In contrast, the second conditional probability,  $\mathcal{P}(\vec{\lambda}_{\text{physics}}|\vec{x})$ , is used to extract the physics parameters, namely the angular observables introduced in Sec. 2.3.4, from the distributions of the decay angles ( $\vec{\Omega}$ ) in  $B_s^0 \rightarrow \phi\mu^+\mu^-$  decays in  $q^2$  regions. The likelihood function, defined as

$$\mathcal{L}(\vec{x}|\vec{\lambda}_{\text{physics}}) = \prod_{i=1}^N \mathcal{P}(x_i|\vec{\lambda}_{\text{physics}}),$$

is maximised for a given data set to which makes the observed data sample most likely. Therefore, the maximum likelihood condition is given by

$$\frac{\partial}{\partial \lambda_i} \mathcal{L}(\vec{x}|\vec{\lambda}_{\text{physics}})|_{\hat{\vec{\lambda}}_{\text{physics}}} = 0,$$

where  $\hat{\vec{\lambda}}_{\text{physics}}$  is the estimator of the set of parameters for which the likelihood is maximal. When dealing with a large numbers of events, the product of the probabilities can quickly become a very small number. This can lead to computational instabilities which can be avoided by determining the estimator  $\hat{\vec{\lambda}}_{\text{physics}}$  using the negative logarithm of the likelihood (NLL,  $-\ln \mathcal{L}$ ). The logarithm can be used to determine the best estimator as it is a monotonically increasing function and, therefore, the ordering of the probabilities stays the same, *i.e.*

$$\mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics},1}) < \mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics},2}) \Leftrightarrow \ln \left( \mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics},1}) \right) < \ln \left( \mathcal{P}(\vec{x}|\vec{\lambda}_{\text{physics},2}) \right).$$

Furthermore, the use of the logarithm is advantageous as the numerical values usually do not encounter problems with the floating point precision as the numerical values for the likelihood might. The logarithm also converts the product of probabilities into a sum of logarithms, which is a computationally faster operation. The NLL is given as

$$\ln \left( \mathcal{L}(\vec{x}|\vec{\lambda}_{\text{physics}}) \right) = \ln \left( \prod_{i=1}^N \mathcal{P}(x_i|\vec{\lambda}_{\text{physics}}) \right) = \sum_{i=1}^N \ln \left( \mathcal{P}(x_i|\vec{\lambda}_{\text{physics}}) \right).$$

Given that the logarithm is a monotonically increasing function, the position of the extrema of the likelihood and the log-likelihood are the same. The maximum likelihood condition can, therefore, be written as

$$-\frac{\partial}{\partial \lambda_i} \ln \mathcal{L}(\vec{x}|\vec{\lambda}_{\text{physics}})|_{\hat{\lambda}_{\text{physics}}} = 0.$$

In this thesis the minimisation of the negative log-likelihood is performed with the **Minuit** [187] implementation in the **ROOT** framework [188]. When determining the best estimators  $\hat{\lambda}_{\text{physics}}$  on a dataset in which no event is weighted, the uncertainties on the parameters are estimated using **Minuit** as implemented in the **HESSE** function. For a data set in which each event receives a weight  $w_i$ , the uncertainties on the best estimators are determined using bootstrapping [189]. In this case, the NLL function can be adjusted to account for the weights. The NLL is then given as

$$\ln \left( \mathcal{L}(\vec{x}|\vec{\lambda}_{\text{physics}}) \right) = \sum_{i=1}^N w_i \times \ln \left( \mathcal{P}(x_i|\vec{\lambda}_{\text{physics}}) \right).$$

A fit using this type of NLL is commonly referred to as a weighted fit. Following Ref. [190], bootstrapping is expected to give the asymptotically correct parameter uncertainties. Therefore, the uncertainties on the angular observables are estimated by bootstrapping the data set 1000 times and repeating the weighted fit for each bootstrap. The uncertainties are then defined asymmetrically around the original fit result, covering 68.3% to each side. An ordered list of all fit results is constructed to find the uncertainties to either side. The positions of the fit results covering 68.3% of the area to the left or right ( $\text{index}_{\text{low,high}}$ ) of the original fit result ( $\text{index}_{\text{original}}$ ) are given by

$$\begin{aligned} \text{index}_{\text{low}} &= 0.683 \cdot \text{index}_{\text{original}} \\ \text{and index}_{\text{high}} &= 0.683 \cdot (1000 - \text{index}_{\text{original}}) + \text{index}_{\text{original}}, \end{aligned}$$

where the indices  $\text{index}_{\text{low,high}}$  are rounded to the nearest integer. The difference between the fit results at the positions  $\text{index}_{\text{low,high}}$  and the original fit result are taken as the uncertainties. The coverage of this uncertainty estimation of the angular observables for both cases, a weighted and unweighted fit to the data sets, is discussed in detail in Sec. 7.3.3.

## 4.3 Statistically subtracting backgrounds with the *sPlot*-technique

An event of a data set can be sorted into one of two categories, signal and background events. Depending on their origin background events can be split into further subcategories. For the signal and background events, there are generally two types of variables, control variables ( $x$ ) and discriminating variables ( $y$ ). Control variables are variables for which the distributions are either unknown or assumed to be unknown and discriminating variables are those with known distributions. One example for a discriminating variable is the invariant mass of the  $B_s^0$  meson candidate, with which the number of events of each category in the data set can be determined. In the context of the *sPlot*-technique [178] this is done with an extended maximum likelihood [191] analysis. Considering each subcategory of signal and background events as an

individual species  $i$ , the log-likelihood can be written as

$$\ln \mathcal{L} = \sum_{e=1}^N \ln \left( \sum_{i=1}^{N_s} N_i f_i(y(e)) \right) - \sum_{i=1}^{N_s} N_i,$$

where  $N$  is the number of events in the data sample,  $N_s$  is the number of species of events in the data sample,  $N_i$  is the average number of expected events for the species  $i$ ,  $y(e)$  is the set of values of the discriminating variables for each event  $e$ ,  $f_i(y(e))$  is the value taken by the PDF of species  $i$  ( $f_i$ ) at  $y(e)$ . The set of control variables  $x$  does not appear in this equation by definition.

The distribution of the control variables of species  $i$  can generally be inferred using per-event weights determined on the discriminating variables. For the case that the control and discriminating variables are uncorrelated, the per-event weights are called *sWeights* and are defined as

$${}_s\mathcal{P}_i(y(e)) = \frac{\sum_{j=1}^{N_s} V_{ij} f_j(y(e))}{\sum_{k=1}^{N_s} N_k f_k(y(e))},$$

where  $V_{ij}$  is the covariance between the species  $n$  and  $j$ . When determining *sWeights* the only free parameter for the PDF is the yield. The distribution  $M_i(x)$  of species  $i$  in  $x$  can be deduced by histogramming  $x$  while applying the *sWeights* to each event. It should be noted that *sWeights* can become negative.



# Chapter 5

## Selection of data and simulation samples

In this Chapter, the selection of the data and simulated samples will be discussed. First, the data samples will be introduced in Sec. 5.1, then the event generation in LHCb and the subsequent simulated samples will be discussed in Sec. 5.2. The partitioning in the separate  $q^2$  regions will be discussed in Sec. 5.3 and the selection of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay will be covered in Sec. 5.4. The corrections applied to the simulation to sufficiently align it with data will be described in Sec. 5.5.

### 5.1 Data samples

The data analysed in this thesis were taken in Run 1 (2011–2012) and Run 2 (2016–2018) of the LHC and correspond to a total integrated luminosity of  $8.4 \text{ fb}^{-1}$ . The integrated luminosity recorded in each year is given in Tab. 5.1 and visually presented in Fig. 5.1. The centre-of-mass energy  $\sqrt{s}$  differs for the data taken in 2011, 2012 and Run 2 with a centre-of-mass energy of 7 TeV in 2011, 8 TeV in 2012 and 13 TeV in Run 2. Considering the increase in  $b\bar{b}$  production cross section for the increased centre-of-mass energies as discussed in Sec. 2.2.1, each  $\text{fb}^{-1}$  of integrated luminosity in Run 2 contains approximately twice as many  $b\bar{b}$ -pairs as in Run 1. The addition of the Run 2 data, therefore, constitutes an approximate four-fold increase in the total number of  $b\bar{b}$ -pairs with respect to the Run 1 sample. Furthermore, the data taken in Run 2 is split into two parts for the analysis presented in this thesis. The data taken in 2016 is referenced as Run 2p1 and the data sets taken in 2017 and 2018 is referenced as Run 2p2. This split is motivated by differing trigger thresholds between the two periods. Data taken in 2015 is not used in this analysis as the L0 trigger configuration for the 2015 data significantly differs from the other running periods. This in turn would require the 2015 data to be treated as an additional data set. However, given the smaller integrated luminosity, the background components could not be reliably be able to be determined in a maximum likelihood fit.

Table 5.1: The integrated luminosity of each year is given.

| year                            | integrated luminosity in $\text{fb}^{-1}$ |
|---------------------------------|-------------------------------------------|
| 2011                            | 1.0                                       |
| 2012                            | 2.0                                       |
| $\Sigma$ Run 1                  | 3.0                                       |
| 2016                            | 1.66                                      |
| 2017                            | 1.63                                      |
| 2018                            | 2.08                                      |
| $\Sigma$ Run 2                  | 5.37                                      |
| $\Sigma$ Run 1 + $\Sigma$ Run 2 | 8.4                                       |

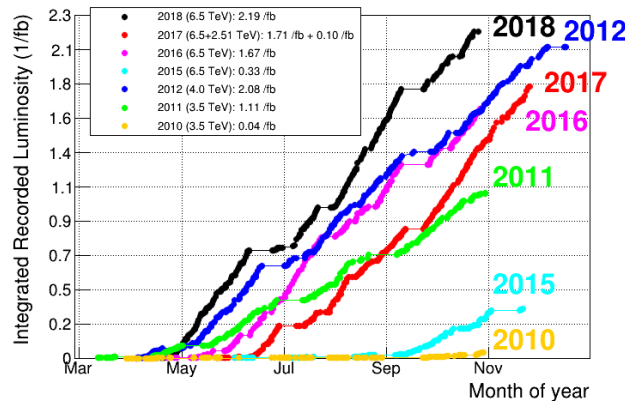


Figure 5.1: The recorded integrated luminosities by the LHCb experiment are shown. The integrated luminosities recorded range from  $0.04 \text{ fb}^{-1}$  in 2010 to  $2.19 \text{ fb}^{-1}$  in 2018. Taken from Ref. [192].

## 5.2 Samples of simulated events

Simulated events are extensively used in physics analyses in LHCb. Simulated events are crucial when testing analysis methods and techniques, when developing specific selections and when studying the effects of the detector on the signal candidates. The Monte Carlo technique (MC) [193] is used to generate these events. In the LHCb collaboration, a dedicated simulation framework has been developed based on the **Gaudi** framework [194, 195]. The simulation framework consists of five projects, each handling a different part of the event generation. They are **Gauss** [196, 197], **Boole** [198], **Moore** [199], **Brunel** [200] and **DaVinci** [201]. The **Gaudi** framework is also used to record and process data. There the packages **Moore**, **Brunel** and **DaVinci** are also used, but not the **Gauss** and **Boole** packages.

The **Gauss** software package [196, 197] is used to interface different programs for the various steps in the event generation. The proton-proton collisions are simulated by **PYTHIA** [202, 203], a general purpose MC event generator, which is specifically tuned for the LHCb environment [204]. Aside from the  $pp$  collision, **PYTHIA** is also used to simulate the  $b\bar{b}$ -pair production. The subsequent decay of the  $B$  hadrons is handled by the **EvtGen** package [205]. This generator has been developed to describe the decay of  $B$  hadrons using decay amplitudes. The **EvtGen** package is also able to describe the interference effects in decay and mixing of  $B$  mesons for

specific decays. The emission of photons in the initial or final state is modeled using the **PHOTOS** package [206]. The interaction of the particles produced in a given proton-proton interaction with the detector is simulated using the **Geant4** software [207–209]. Using the output of the **Gauss** software as an input, **Boole** [198] then simulates the detector response to hits in the various subdetectors. It is also used to emulate the response of the readout electronics and the L0 trigger hardware. The output of **Boole** is designed to match the format of data read-out by the detector. The resulting simulated events from hereon go through the same steps as recorded data. The simulated and recorded events are passed to the **Moore** package [199] in order to determine the HLT response. Afterwards **Brunel** [200] is used to reconstruct the particles. In simulated events, **Brunel** is also responsible for the matching of reconstructed tracks to the originally produced events. This feature is only available for simulation where the generator-level information is available. The last step in the generation process is handled by **DaVinci** [201] which is used to perform a first selection of the events. The selection will be introduced in Sec. 5.4. Events simulated by the LHCb collaboration are saved on the worldwide LHC computing grid [210, 211].

In Tab. 5.2 the simulated samples used in this analysis, the event types, which is a unique identifier for models used for the simulation, generation conditions, the used model and the number of events generated without filtering ( $N_{\text{DST}}$ ) are given. For the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays two different models are used to model the decay of the  $B_s^0$  meson. One sample decays the  $B_s^0$  meson purely according to the available phase space of the decay. This means that the distribution of the angles between the final state particles (decay angles) is generated flat. This model is referred to as the phase space model (PHSP). These samples are used to determine the angular acceptance which encodes the effect of the detector acceptance, reconstruction, trigger and selection and is introduced in detail in Sec. 7.1. A phase space model is preferable for determining the acceptance as the shape of distribution across the decay angles after all selection directly encodes the efficiency due to the flat generation. In contrast, an additional set of samples use a proper physics model in the generation, the **BTOSLLBALL v6** model [212, 213]. It is a  $b \rightarrow s \ell^+ \ell^-$  decay model [212], which takes form factors from Ref. [213]. These samples have the advantage of describing the angles and the  $q^2$  distribution in the SM, making them more appropriate for the validation of the angular acceptance as described in Sec. 7.3.1.

Table 5.2: The simulated samples used in this analysis.

| Decay                                       | Eventtype | $N_{\text{DST}}$ | Generation Conditions | Model         |
|---------------------------------------------|-----------|------------------|-----------------------|---------------|
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$        | 13114005  | 5002941          | 2012                  | PHSP          |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$        | 13114005  | 6101210          | 2016                  | PHSP          |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$        | 13114005  | 8009627          | 2017                  | PHSP          |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$        | 13114006  | 2046275          | 2012                  | BTOSLLBALL v6 |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$        | 13114006  | 3010950          | 2016                  | BTOSLLBALL v6 |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$        | 13114006  | 3010721          | 2017                  | BTOSLLBALL v6 |
| $B_s^0 \rightarrow J/\psi \phi$             | 2014253   | 13144001         | 2012                  | -             |
| $B_s^0 \rightarrow J/\psi \phi$             | 2419154   | 13144001         | 2016                  | -             |
| $B_s^0 \rightarrow J/\psi \phi$             | 9282494   | 13144001         | 2017                  | -             |
| $B^0 \rightarrow K^{*0} \mu^+ \mu^-$        | 11114002  | 2520250          | 2012                  | BTOSLLBALL v6 |
| $B^0 \rightarrow K^{*0} \mu^+ \mu^-$        | 11114001  | 1365262          | 2016                  | BTOSLLBALL v6 |
| $B^0 \rightarrow K^{*0} \mu^+ \mu^-$        | 11114002  | 4014543          | 2017                  | BTOSLLBALL v6 |
| $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$ | 15114011  | 2019539          | 2012                  | PHSP          |
| $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$ | 15114011  | 7974997          | 2016                  | PHSP          |
| $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$ | 15114011  | 2225233          | 2017                  | PHSP          |
| $B^0 \rightarrow J/\psi K^{*0}$             | 11144001  | 4422426          | 2012                  | physics       |
| $B^0 \rightarrow J/\psi K^{*0}$             | 11144001  | 2449607          | 2016                  | physics       |
| $B^0 \rightarrow J/\psi K^{*0}$             | 11144001  | 4127863          | 2017                  | physics       |

### 5.3 Partitioning in $q^2$ regions

The angular analysis is performed in regions of the invariant dimuon mass squared ( $q^2$ ). The partitioning into  $q^2$  regions is changed with respect to the previous analysis in Ref. [15] to have finer regions, which maximises the sensitivity to potential short-distance NP contributions whilst ensuring stable fit behaviour. The  $q^2$  regions are given in Tab. 5.3. The regions of width  $\sim 2\text{--}3 \text{ GeV}^2/c^4$  are referred to as the *narrow*  $q^2$  regions. In addition to the *narrow*  $q^2$  regions, the results of the angular analysis is provided in two *wide*  $q^2$  regions, namely  $q^2 \in [1.1, 6.0] \text{ GeV}^2/c^4$  and  $q^2 \in [15.0, 18.9] \text{ GeV}^2/c^4$ . For the high  $q^2$  region, the upper-boundary at  $18.9 \text{ GeV}^2/c^4$  corresponds to the allowed phase-space limit in  $q^2$  given by  $m^2(B_s^0) - m^2(\phi)$ .

The excluded  $q^2$  regions, corresponding to  $q^2 \in [0.98, 1.1] \text{ GeV}^2/c^4$ ,  $q^2 \in [8.0, 11.0] \text{ GeV}^2/c^4$  and  $q^2 \in [12.5, 15.0] \text{ GeV}^2/c^4$ , remove the  $\phi$ ,  $J/\psi$  and  $\psi(2S)$  resonances, respectively. The remaining  $q^2$  regions are collectively referred to as the rare mode.

The  $B_s^0 \rightarrow J/\psi \phi$  decays, found in the region of  $q^2 \in [8.0, 11.0] \text{ GeV}^2/c^4$ , are referred to as the control mode in this analysis. These decays are used for example to determine corrections to the simulation, as introduced in Sec. 5.5 and to validate the angular fit in Sec. 7.3.

Table 5.3: The  $q^2$  binning scheme used in the analysis. The regions 1–5 correspond to the narrow binning scheme, and regions 6–7 correspond to the wide binning scheme

| region | $q^2$ range [ $\text{GeV}^2/c^4$ ] | veto            |
|--------|------------------------------------|-----------------|
| 1      | 0.1 – 0.98                         | $\phi$ veto     |
| -      | 0.98 – 1.1                         |                 |
| 2      | 1.1 – 4.0                          |                 |
| 3      | 4.0 – 6.0                          |                 |
| 4      | 6.0 – 8.0                          |                 |
| -      | 8.0 – 11.0                         | $J/\psi$ veto   |
| 5      | 11.0 – 12.5                        |                 |
| -      | 12.5 – 15.0                        | $\Psi(2S)$ veto |
| 6      | 1.1 – 6.0                          |                 |
| 7      | 15.0 – 18.9                        |                 |

### 5.4 Event selection

In this section the event selection will be discussed. It is divided into six parts. Important quantities associated to particle candidates in the event are introduced in Sec. 5.4.1, the trigger selection is discussed in Sec. 5.4.2, and the truth matching of simulated events is described in Sec. 5.4.3. The LHCb-wide preselection (stripping) is outlined in Sec. 5.4.4, the additional analysis specific preselection based on properties of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  signal decay is introduced in Sec. 5.4.5 and the BDT against combinatorial background is covered in Sec. 5.4.6. Lastly, the removal of candidates for events in which multiple  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates pass the full selection is described in Sec. 5.4.7.

### 5.4.1 Definition of important quantities of particle candidates

The signal selection deployed in this thesis relies on placing a combination of requirements on important quantities of final state particles and combined particles. The variables used for these purposes can either be quantities directly measurable, such as the energy or the momentum of a particle or can be calculated such as particle identification variables. These variables will be introduced in the following.

#### Energy, momentum and invariant mass

The momentum of a charged particle leaving a track in the LHCb experiment is measured with the tracking system as the warm dipole magnet induces a curvature of the track. The total momentum ( $p$ ) as well as the components in the  $x$ -,  $y$ - and  $z$ -directions ( $p_x$ ,  $p_y$  and  $p_z$ ) are measured. Furthermore, the transverse momentum ( $p_T$ ) of a track, defined as the momentum transverse to the beam direction, can be determined using these components.

The energy ( $E$ ) of a particle leaving a track in the detector is computed from the measured momentum by giving a hypothesis on the particle type ( $m_0$ ), *i.e.* using particle identification information, to a track as these quantities are related via  $E^2 = p^2 c^2 + m_0^2 c^4$ , where  $c$  is the speed of light in vacuum. While the energy of charged particles could in principle also be measured in the calorimeters, the computation of the energy with the momentum measurement is advantageous as the momentum resolution is typically better than the energy resolution from the calorimeter system at LHCb.

With the energy and the momentum the particle candidates being known, the energy and momentum of a combined particle such as the  $\phi$  or  $B_s^0$  mesons can be determined by adding the four-momentum ( $p^\mu$ ) of the involved particles. For example the four-momenta of a  $\phi$  meson decaying into a pair of kaons is given as  $p^\mu(\phi) = p^\mu(K^+) + p^\mu(K^-)$ . The invariant mass of such a composite particle can then be inferred from the energy-momentum relation  $E^2 = p^2 c^2 + m_0^2 c^4$ . This can also be done with the four-momentum as their components are the energy and the momentum vector, for example for the  $B_s^0$  meson the invariant mass is given as  $m(B_s^0) = \sqrt{(p^\mu(K^+) + p^\mu(K^-) + p^\mu(\mu^+) + p^\mu(\mu^-))^2}$ .

When considering backgrounds originating from misidentifying one or more final state particles, the invariant mass of a composite particle where one of the final state particles is given another mass hypothesis can be a powerful discriminant. For example in a  $B_s^0 \rightarrow J/\psi \phi$  decay a pair of kaon and muon of the same charge can be misidentified as each other, such that the kaon is reconstructed under the muon mass hypothesis and vice-versa. The invariant mass of the  $J/\psi$  and of the  $\phi$  mesons would be different in the case of a misidentification compared to the correct identification. Therefore, a selection on the mass of the  $J/\psi$  meson under the background hypothesis ( $m((K^\pm \rightarrow \mu^\pm), \mu^\mp)$ ) can be used to reject backgrounds, where the muon misidentified and reconstructed as a kaon is given the muon mass hypothesis and combined with the oppositely charged muon in the final state.

#### Particle identification

Particle identification information is available from the muon system, the RICH detectors and the calorimeters. The information of the subdetectors can be used to help identify the particle species of a track. For this purpose, boolean flags saving information if a track has hits in the muon system (`isMuon`) [214] or in at least one of the RICH detectors (`HASRICH`) are created during the reconstruction. Considering that muons are most likely to leave a track in the muon

system, the `isMuon` variable is a strong discriminator between muons and other particle types. Another approach to obtain information on the particle species is to combine the information of the subdetectors into a single particle identification variable [122,132]. In this approach the likelihood of a track having a particular mass hypothesis ( $X$ ) is determined separately for the RICH, calorimeter and muon subdetectors and linearly combined into a single likelihood and compared to the pion mass hypothesis ( $\Delta \log \mathcal{L}(X - \pi)$ ) [122,132]. Following this definition, this variable is a measure for how likely a track is a particle of species  $X$  compared to being a pion. For a track reconstructed as species  $Y$  the variable is referred to as `PIDX( $Y$ )`, where  $X$  is the tested particle type hypothesis. Another way to combine the information of the subdetectors is the use of multivariate analyses [122,132,215,216]. The multivariate analyses use information of the muon system, the RICH detectors and the calorimeters and adds the combined likelihoods as well as information on the track momentum and track quality. For each mass hypothesis a different multivariate analysis is trained resulting in several variables called `ProbNNX( $Y$ )`, where  $X$  is the mass hypothesis in question for a track reconstructed as species  $Y$ .

It is also important to ensure that the reconstructed tracks correspond to a particle traversing the detector. Reconstructed tracks that do not correspond to a particle traversing the detector can originate from mismatched hits from different particles or from detector noise. These tracks are referred to as fake tracks or ghosts. To reject these ghosts a neural network has been developed [151,217,218] that uses the  $\chi^2$  of the Kalman filter, the  $\chi^2$  values for the fits for the different track segments, the number of hits in the different tracking detectors and the  $p_T$  of the track. The output of the neural network is then transformed to correspond to a probability of a track to be a ghost (`ghostProb`).

## Topological variables

The geometric properties of reconstructed tracks can be used to differentiate signal from background. One such variable is the direction angle (`DIRA`) defined as the cosine of the angle between the direction of flight of a particle and the vector connecting the PV and the decay vertex of the particle question. Another important angle is the pseudorapidity ( $\eta$ ), a Lorentz-invariant quantity, defined as  $\eta = -\log(\tan \frac{\theta}{2})$ , where  $\theta$  is the polar angle of a particle.

A second class of important topological variables used in LHCb are variables that behave like a  $\chi^2$  distribution. Therefore, this second class of topological variables are referred to as  $\chi^2$  variables and are unitless. The three most important variables of this type for this thesis are  $\chi_{\text{IP}}^2$ ,  $\chi_{\text{FD}}^2$  and  $\chi_{\text{vtx}}^2$ . The first of these variables,  $\chi_{\text{IP}}^2$ , concerns the impact parameter (IP). The IP is defined as the distance of closest approach to the PV of a track extrapolated towards the PV. The  $\chi_{\text{IP}}^2$  can be used to distinguish tracks originating the PV from tracks originating from secondary vertices. The second variable takes the flight distance (FD) into account, which defined as the distance between the vertex of origin and the vertex of decay of a particle. The  $\chi_{\text{FD}}^2$  can be used to differentiate long-lived particles from short-lived ones. These two variables are constructed by determining the difference in the vertex-fit  $\chi^2$  of a given PV reconstructed with or without the track in question. The third variable,  $\chi_{\text{vtx}}^2$ , is a measure for the fit quality of the vertex and is as such usually used normalised to the number of degrees of freedom (ndof). It is a powerful tool to separate tracks not originating from vertices from those which do.

## 5.4.2 Trigger

At LHCb a trigger system, introduced in more detail in Sec. 3.2.6, is used to reduce the rate at which events are recorded from the bunch crossing rate at 40 MHz to 5 kHz (12.5 kHz) in Run 1 (Run 2) while selecting signal candidates with high efficiency. For an event to be considered in this analysis, it has to pass at least one trigger in the L0, HLT1 and HLT2 stages. The positive trigger decision is required to be taken on a part of the signal decay (TOS). An alternative approach is to require the trigger to fire independent of signal (TIS). This alternative is used when calibrating the L0 trigger performance in data, as discussed in Sec. 5.5.3. In Tab. 5.4 the trigger lines required for this analysis are shown.

Table 5.4: Trigger lines used for selection. The `DiMuonDetached` trigger is not used for 16 as its dimuon momentum threshold is set to the same value as for the `DiMuonDetachedHeavy` due to a bug. This bug is fixed for most of 2017 and 2018, so there it is included. As `TrackAllL0` is not available for Run 2, the `(Two)TrackMVA` lines have been added.

| stage | Run 1                                                                                                                                             | Run 2                                                                                                                                        |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| L0    | Muon<br>DiMuon                                                                                                                                    | Muon<br>DiMuon                                                                                                                               |
| HLT1  | TrackAllL0<br>TrackMuon<br>-                                                                                                                      | TrackMuon<br>TrackMVA<br>TwoTrackMVA                                                                                                         |
| HLT2  | Topo2BodyBBDT<br>Topo3BodyBBDT<br>Topo4BodyBBDT<br>TopoMu2BodyBBDT<br>TopoMu3BodyBBDT<br>TopoMu4BodyBBDT<br>DiMuonDetached<br>DiMuonDetachedHeavy | Topo2Body<br>Topo3Body<br>Topo4Body<br>TopoMu2Body<br>TopoMu3Body<br>TopoMu4Body<br>DiMuonDetached <sup>2017/18</sup><br>DiMuonDetachedHeavy |

The L0 triggers require a single muon to have a transverse momentum greater than 1.5 GeV/ $c$  in the `Muon` L0 trigger or a dimuon pair to have a product of transverse momenta greater than 1.5 GeV/ $c$  in the `DiMuon` L0 trigger. In the HLT1 stage, the `TrackAllL0` and `TrackMuon` triggers are used in Run 1 to select events. Events triggered by the `TrackAllL0` line selects tracks with hits in the VELO and requires the track to be displaced tracks from any PV in the event using the  $\chi_{\text{IP}}^2$  variable. Furthermore, the track is required to have a transverse momentum above 1.3 GeV/ $c$ . The `TrackMuon` line operates similarly albeit with less strict requirements but additionally requires hits in the muon station and a positive `Muon` L0 trigger decision. As the HLT1 `TrackAllL0` trigger is not available in Run 2, it is replaced by the HLT1 `TrackMVA` and `TwoTrackMVA` triggers in Run 2. The trigger lines use a multi-dimensional classifier based on kinematic variables such as the transverse momentum and  $\chi_{\text{IP}}^2$  to select events. In the HLT2 stage, the `Topo2,3,4Body` lines are designed to select partially reconstructed events containing a  $B$  hadron. These lines achieve this by selecting events consistent with a two, three or four body topology using an MVA [219–221]. The `TopoMu2,3,4Body` lines function similarly to the `Topo2,3,4Body` lines albeit with less strict requirements but also require at least 1 track to be identified as a muon. The `DiMuonDetached` and `DiMuonDetachedHeavy` lines trigger events containing a pair of muons originating from a displaced secondary vertex. Furthermore, the `DiMuonDetachedHeavy` line requires the dimuon candidate to exceed a mass threshold of 2950 MeV/ $c^2$ . The `DiMuonDetached` trigger line in the HLT2 is not used for

Run 2p1, as the dimuon momentum threshold was erroneously set to the similar value as in the `DiMuonDetachedHeavy`. As this bug is fixed for most of Run 2p2 data, it is included again for this data set.

### 5.4.3 Truth-matching of simulated events

To ensure the simulated muons and kaons selected after the reconstruction were generated as muons and kaons, each of the reconstructed particles is matched to a generated particle. This is referred to as truth matching. When performing truth matching, the particle ID, as introduced in Chapter 44 of Ref. [65], of the final state particles are required to match those of muons and kaons. Explicitly, this means that the `TRUEIDs` of positively charged kaons are required to be +321, for negatively charged kaons -321, for negatively charged muons +13 and for positively charged muons -13. In addition to having the correct particles in the final state, it is required that all final state particles originate from the same mother, and that this mother is a  $B_s^0$  meson. For truth matched background candidates, the final state particle `TRUEIDs` are matched to the appropriate hypotheses such that there was a misidentification.

### 5.4.4 LHCb wide preselection

Events passing the trigger stages are written to tape, which is not directly accessible by analysts. Given the storage limitations, only a fraction of the events written to tape can be written to disk. This reduction is achieved with LHCb-wide selection lines (stripping lines), which are run over the events stored on tape in periodic central productions. The stripping lines select reconstructible signal decays with a high efficiency and write the resulting output to disk, where they can be accessed by analysts. The stripping line used for this analysis is the `B2XMuMu` line. The `B2XMuMu` stripping line selects events in which all tracks of the  $K^+K^-\mu^+\mu^-$  final state have a significant  $\chi_{\text{IP}}^2$  to any PV. Furthermore, the final state muons have to exceed thresholds on the particle identification variables and the final state kaons have to have hits in the `RICH` detector. The four final state particles are used to fit a vertex for the  $B_s^0$  meson. This vertex needs to be of good quality, and significantly displaced from any PV in the event. The complete list of selection criteria and their numerical values are listed in Tab. 5.5. Criteria that differ between Run 1 and Run 2 are also identified in this table.



Table 5.5: Stripping selection criteria. The given values are the settings for every stripping version that is used in this analysis, unless indicated. The \* symbol indicated PID selections that are removed from the stripping of simulation and applied offline to allow correction before the application.

| Variable                                    | Selection                                                                                                                             |
|---------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| $m(B_s^0)$                                  | $> 4900^{\text{Run 1}} > 4700^{\text{Run 2}} \text{ MeV}/c^2$                                                                         |
| $m(B_s^0)$                                  | $< 7000 \text{ MeV}/c^2$                                                                                                              |
| $\chi_{\text{vtx}}^2(B_s^0)/\text{ndof}$    | $< 8.0$                                                                                                                               |
| $\chi_{\text{IP}}^2(B_s^0)$                 | $< 16.0$                                                                                                                              |
| $\text{DIRA}(B_s^0)$                        | $> 0.9999$                                                                                                                            |
| $\chi_{\text{FD}}^2(B_s^0)$                 | $> 121.0$                                                                                                                             |
| $\max(\text{daughter } \chi_{\text{IP}}^2)$ | $> 9.0$                                                                                                                               |
| $m(\phi)$                                   | $\in [0, 6200] \text{ MeV}/c^2$                                                                                                       |
| $\chi_{\text{FD}}^2(\phi)$                  | $> 9.0^{\text{Run 1}} > 16.0^{\text{Run 2}}$                                                                                          |
| $\chi_{\text{IP}}^2(\phi)$                  | $> 0.0$                                                                                                                               |
| $\max(K^+, K^- \chi_{\text{IP}}^2)$         | $> 9.0$                                                                                                                               |
| $\chi_{\text{vtx}}^2(\phi)/\text{ndof}$     | $< 12.0$                                                                                                                              |
| $\text{DIRA}(\phi)$                         | $> -0.9$                                                                                                                              |
| $m(J/\psi)$                                 | $< 7100 \text{ MeV}/c^2$                                                                                                              |
| $\chi_{\text{FD}}^2(J/\psi)$                | $> 9.0$                                                                                                                               |
| $\chi_{\text{vtx}}^2(J/\psi)/\text{ndof}$   | $< 12.0$                                                                                                                              |
| $\text{DIRA}(J/\psi)$                       | $> -0.9$                                                                                                                              |
| $\max(\text{daughter } \chi_{\text{IP}}^2)$ | $> 9.0$                                                                                                                               |
| $\text{PIDmu}(\mu^\pm)$                     | $> -3.0^*$                                                                                                                            |
| $\chi_{\text{IP}}^2(\mu^\pm)$               | $> 9.0$                                                                                                                               |
| $\text{isMuon}(\mu^\pm)$                    | $\text{true}^*$                                                                                                                       |
| $\chi_{\text{IP}}^2(K^\pm)$                 | $> 6.0$                                                                                                                               |
| $\text{HASRICH}(K^\pm)$                     | $\text{true}$                                                                                                                         |
| tracks ghostProb                            | $< 0.35^{\text{Run 1}} < 0.5^{\text{Run 2}}$                                                                                          |
| hits in SPD                                 | $< 600$                                                                                                                               |
| trigger filters <sup>Run 1</sup>            | $(\text{HLT1TrackAllL0} \mid \text{Hlt1TrackMuon})$<br>$\& (\text{HLT2Dimuon}^* \mid \text{Hlt2Topo}^* \mid \text{Hlt2SingleMuon}^*)$ |

### 5.4.5 Analysis specific preselection

Once the criteria of the B2XMuMu stripping line are passed, an additional analysis specific preselection is applied to further reduce background contributions. The list of criteria, together with the numerical values are given in Tab. 5.6. While muons are already required to be positively identified using particle identification variables in the B2XMuMu stripping line, the equivalent for kaons is done in the analysis specific preselection. Furthermore, the kaons of a  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay are required to originate from a  $\phi$  meson. The narrowness of the  $\phi \rightarrow K^+ K^-$  resonance is exploited to further reduce background contributions. Given the total decay width of a  $\phi$  meson is approximately  $4 \text{ MeV}/c^2$  [65], the chosen mass range of 12 MeV in either direction of the known phi mass selects more than 95% of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays.

A secondary set of selection criteria is used to specifically target background contributions originating from  $B$  decays where one or more of the daughters have been assigned an incorrect mass hypothesis (peaking backgrounds). A complete list of the considered backgrounds is given in Sec. 6.1. One example of such a decay is the process  $B_s^0 \rightarrow J/\psi \phi$  where one kaon is misidentified as a muon and vice versa, such that the final state stays the same, but the value of  $q^2$  differs and is no longer rejected. To reduce backgrounds of this form involving either a  $J/\psi$  or a  $\psi(2S)$  meson, more stringent PID criteria are applied to the kaon candidate, if the  $J/\psi$  or  $\psi(2S)$  candidates has a mass reconstructed under the background hypothesis ( $m((K^\pm \rightarrow \mu^\pm), \mu^\mp)$ ) within  $45 \text{ MeV}/c^2$  of the known mass of the  $J/\psi$  or  $\psi(2S)$  mass [65]. An analogous selection is performed to reject decays  $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$  and  $\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-$ , in which case the kaon is a mis-identified proton. The exact selection criteria are given in Tab. 5.7.

Table 5.6: A summary of the applied offline selection criteria.

| Variable                 | Selection                                     |
|--------------------------|-----------------------------------------------|
| $m(K^- K^+ \mu^+ \mu^-)$ | $\in [5270, 5700] \text{ MeV}/c^2$            |
| $m(K^+ K^-)$             | $\in [m_{\phi(1020)} \pm 12 \text{ MeV}/c^2]$ |
| $\text{ProbNNk}(K^\pm)$  | $> 0.025$                                     |

Table 5.7: Table of background processes and the applied mis-id vetoes against them.

| decay                                                 | veto                                                                                         | signal efficiency |
|-------------------------------------------------------|----------------------------------------------------------------------------------------------|-------------------|
| $B_s^0 \rightarrow J/\psi \phi$ double mis-id         | $ m((K^\pm \rightarrow \mu^\pm), \mu^\mp) - m_{J/\psi}  < 45 \text{ MeV}/c^2$                | $> 99.9\%$        |
| $B^0 \rightarrow J/\psi K^{*0}$ double mis-id         | and $(\text{isMuon}(K^\pm) \text{ or } \text{PIDmu}(K^\pm) > 5.0)$                           |                   |
| $B_s^0 \rightarrow \psi(2S) \phi$ double mis-id       | $ m((K^\pm \rightarrow \mu^\pm), \mu^\mp) - m_{\psi(2S)}  < 45 \text{ MeV}/c^2$              | $> 99.9\%$        |
|                                                       | and $(\text{isMuon}(K^\pm) \text{ or } \text{PIDmu}(K^\pm) > 5.0)$                           |                   |
| $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$           | $ m((K^\pm \rightarrow p^\pm), K^\mp, \mu^+, \mu^-) - m_{\Lambda_b^0}  < 50 \text{ MeV}/c^2$ | $> 99\%$          |
| $\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-$ | and $\text{ProbNNp} \cdot (1 - \text{ProbNNk})(K^\pm) > 0.35$                                |                   |

### 5.4.6 BDT

The last set of selection criteria used in this analysis is used to specifically reduce background from random combinations of tracks (combinatorial background). While peaking backgrounds can be removed using either a single variable or simple combinations of variables, this is not possible for combinatorial background, while retaining high signal efficiency. For this reason, a BDT classifier as implemented in the TMVA software package [222, 223] is trained on data to reject these contributions. Control mode  $B_s^0 \rightarrow J/\psi \phi$  events in the  $q^2$  region  $[8, 11] \text{ GeV}^2/c^4$  and a  $\pm 50 \text{ MeV}/c^2$  mass range around the known  $B_s^0$  mass [65] serve as a signal proxy in the training. Even though, these requirements on the control mode already ensure a high purity,

the *sPlot* technique is used to statistically subtract the remaining combinatorial background contribution. To extract the *sWeights* for the Run 1, Run 2p1 and Run 2p2 data sets, two fits per data set are performed. Both fits use the same fit model, namely two Gaussians each with a power-law tail to either the upper or lower mass, in the following referred to as double Crystal Ball functions [224] (DCB), to describe the  $B_s^0 \rightarrow J/\psi \phi$  signal and an exponential function to describe the combinatorial background. The fit model is described in more detail in Sec. 7.2. While the first fit is performed in the region of  $m(K^+K^-\mu^+\mu^-) \in [5170, 5570] \text{ MeV}/c^2$  to fix all floating parameters except for the yields, the second fit is performed only floating the yields to determine explicitly the *sWeights* in the region  $m(K^+K^-\mu^+\mu^-) \in [5316.7, 5416.7] \text{ MeV}/c^2$ . These fits are shown in Fig. 5.2. As background proxy, the upper mass sideband, defined as  $m_{B_s^0} > 5567 \text{ MeV}/c^2$  and enriched in combinatorial background by relaxing the  $12 \text{ MeV}/c^2$   $m(K^+K^-)$  range to a  $50 \text{ MeV}/c^2$  range, is used. The signal and background proxies are indicated respectively as blue and red boxes in the  $m_{K^+K^-\mu^+\mu^-} - q^2$  plane in Fig. 5.3 for each of the three data sets.

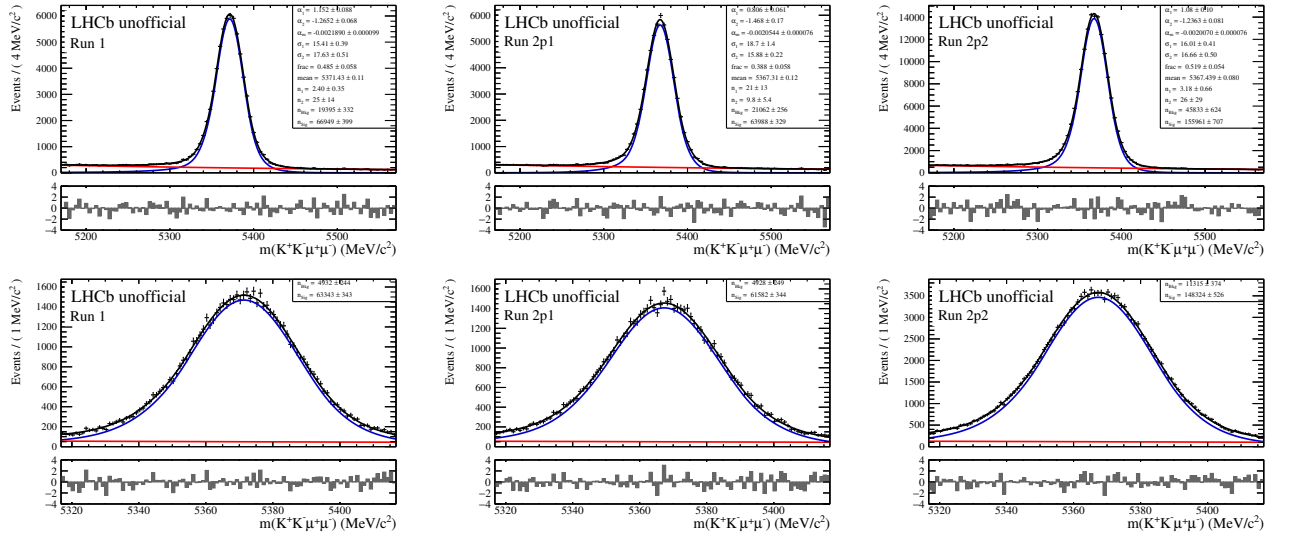


Figure 5.2: Fits (top) to fix all floating parameters but the event yields of the control mode  $B_s^0 \rightarrow J/\psi \phi$  and (bottom) to extract the *sWeights* in (left) Run 1, (middle) Run 2p1 and (right) Run 2p2 data. The blue(red) component corresponds to the  $B_s^0 \rightarrow J/\psi \phi$  signal (combinatorial background).

The BDT classifier is based on the particle identification variables of the final state particles, the  $\chi_{\text{IP}}^2$  of the final state particles and the  $B_s^0$  meson, as well as the  $B_s^0$  vertex fit quality and its displacement. Any bias is avoided using the k-fold cross-validation technique, as described in Sec. 4.1, using ten folds. Eight are used to train each of the BDTs, one is used for testing and the last for the evaluation. The complete list of the input variables is given in Tab. 5.8. The input distributions are shown using the Run 1 BDT as an example in Fig. 5.4 for one of the ten folds.

For each of the used data sets a different BDT is trained. The BDT responses, as well as the receiver operating characteristic (ROC) curves, are shown for Run 1, Run 2p1 and Run 2p2 in Fig. 5.5. The areas under the ROC curve for the BDTs are 0.996, 0.995 and 0.996 for Run 1, Run 2p1 and Run 2p2, respectively. The distribution of the BDT classifier output from events used in the testing follows the distribution for events used in training. Therefore, no signs of

overtraining are observed.

The selection criterion on the BDT classifier is determined using the figure of merit (FOM)  $S/\sqrt{S+B}$ , where  $S$  is the number of signal events and  $B$  is the number of background events, as a guideline. As this analysis is performed blind, the number of signal events  $S$  is also kept blind until the selection and validation are finalised. Therefore, the FOM is determined using an expected number of signal events. The number of expected signal events is obtained by scaling the number of  $B_s^0 \rightarrow J/\psi \phi$  candidates with the branching fraction ratio of the two modes of  $f = 0.013$  [65] while assuming the BDT to have the same efficiency for the rare mode and the control mode. In order to calculate the number of expected background events  $B$  in the signal region, an exponential function is fitted to the mass distribution of the rare mode, where the signal region  $5316 \text{ MeV}/c^2 < m(B_s^0) < 5416 \text{ MeV}/c^2$  is excluded from the fit. The number of expected combinatorial background events  $B$  are obtained by interpolating this exponential into this excluded signal region. The result of this scan over the FOM can be found in Fig. 5.6 for Run 1, Run 2p1 and Run 2p2 data, separately. The FOM has a plateau in the region of  $[0.02, 0.1]$  for each of the BDT classifiers and a single threshold on the BDT classifier is placed at 0.05. While a slightly more stringent requirement could have been used, the increase in the number of rare mode candidates stabilises the angular fit.

### 5.4.7 Multiple candidates

As the pile-up at the design luminosity of the LHCb is 2.1, it is likely to have more than one PV in a given event. Therefore, it is possible to have more than one signal candidate in an event. However, it is highly unlikely to have two signal decays in a given event considering the branching fractions. In the case that an event contains multiple  $B_s^0$  candidates passing the full selection, one candidate is, therefore, randomly chosen and the other is rejected. This is done in both, the data and simulated samples. The fraction of events that contain multiple signal candidates is given in Tab. 5.9 for both the rare mode region as well as the resonant mode. All the fractions in the signal simulation are well below one percent.

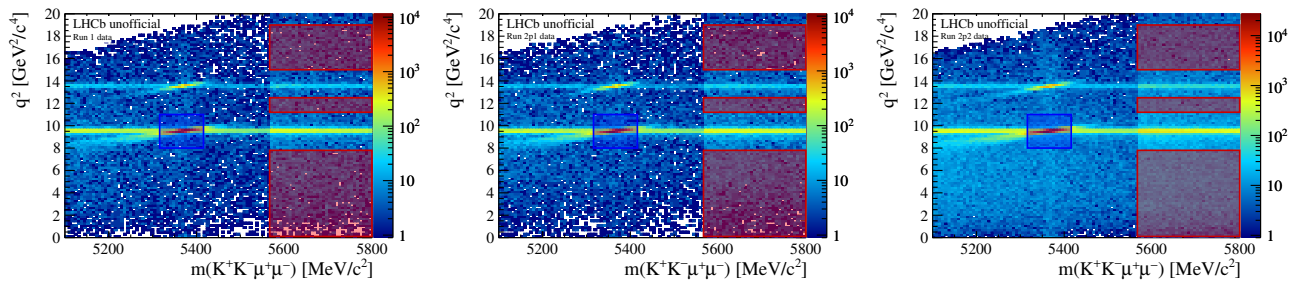


Figure 5.3: The (left) Run 1, (middle) Run 2p1 (right) and Run 2p2 data sets after the preselection used to determine the signal and background proxies from the blue and red regions, respectively, for the BDT classifier training.

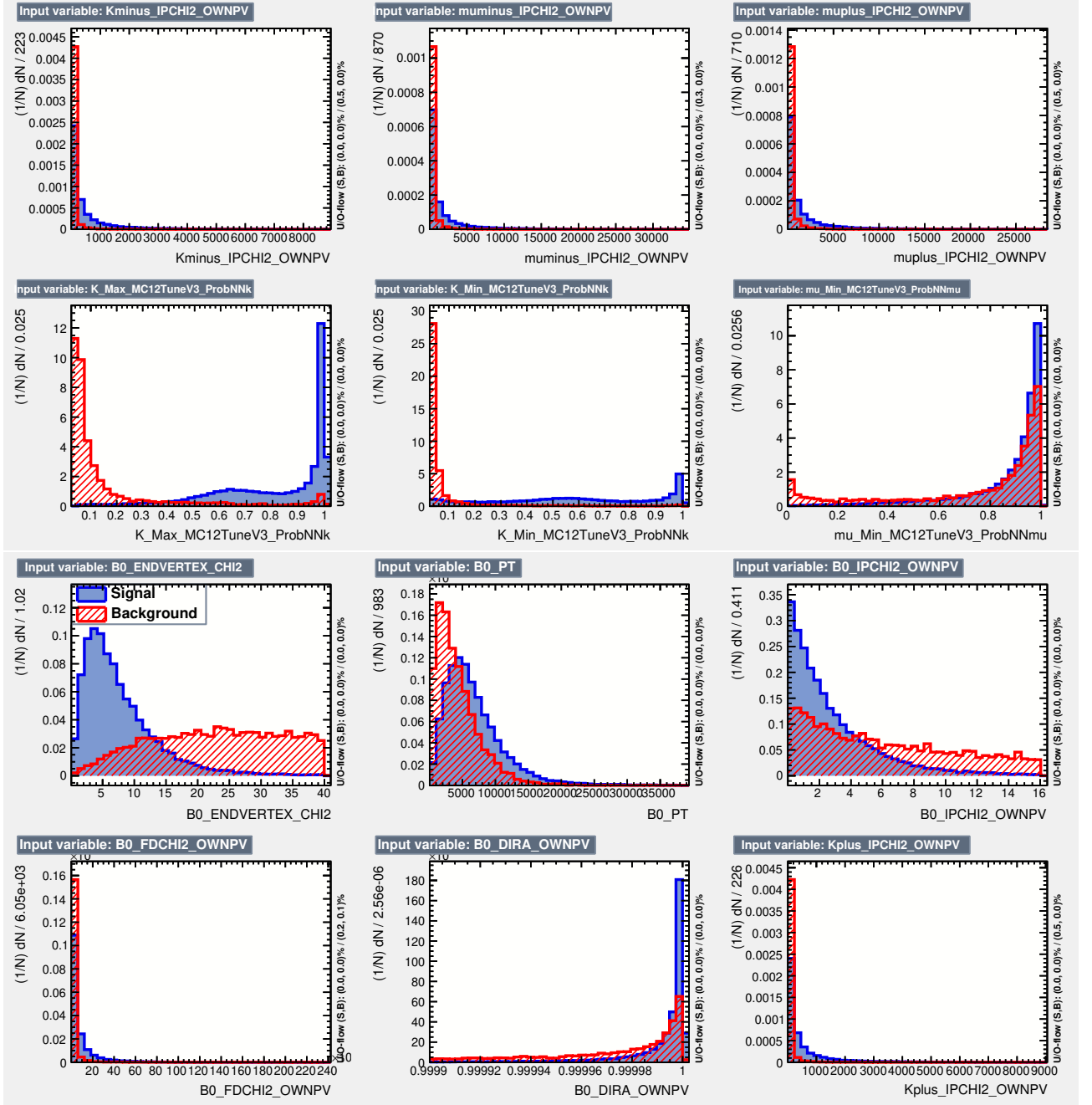


Figure 5.4: Examples of the signal (blue) and background (hashed red) distributions for the input variables used in the BDT training. Example taken from one of the folds in the Run 1 data-set.

Table 5.8: List of BDT input parameters in order of separation power.

| BDT input parameters                                   |
|--------------------------------------------------------|
| $\min(\text{ProbNNk}(K^+), \text{ProbNNk}(K^-))$       |
| $\chi_{\text{vtx}}^2(B_s^0)$                           |
| $\max(\text{ProbNNk}(K^+), \text{ProbNNk}(K^-))$       |
| $p_T(B_s^0)$                                           |
| $\text{DIRA}(B_s^0)$                                   |
| $\chi_{\text{IP}}^2(B_s^0)$                            |
| $\min(\text{ProbNNmu}(\mu^+), \text{ProbNNmu}(\mu^-))$ |
| $\chi_{\text{IP}}^2(K^+)$                              |
| $\chi_{\text{FD}}^2(B_s^0)$                            |
| $\chi_{\text{IP}}^2(\mu^-)$                            |
| $\chi_{\text{IP}}^2(\mu^+)$                            |
| $\chi_{\text{IP}}^2(K^-)$                              |

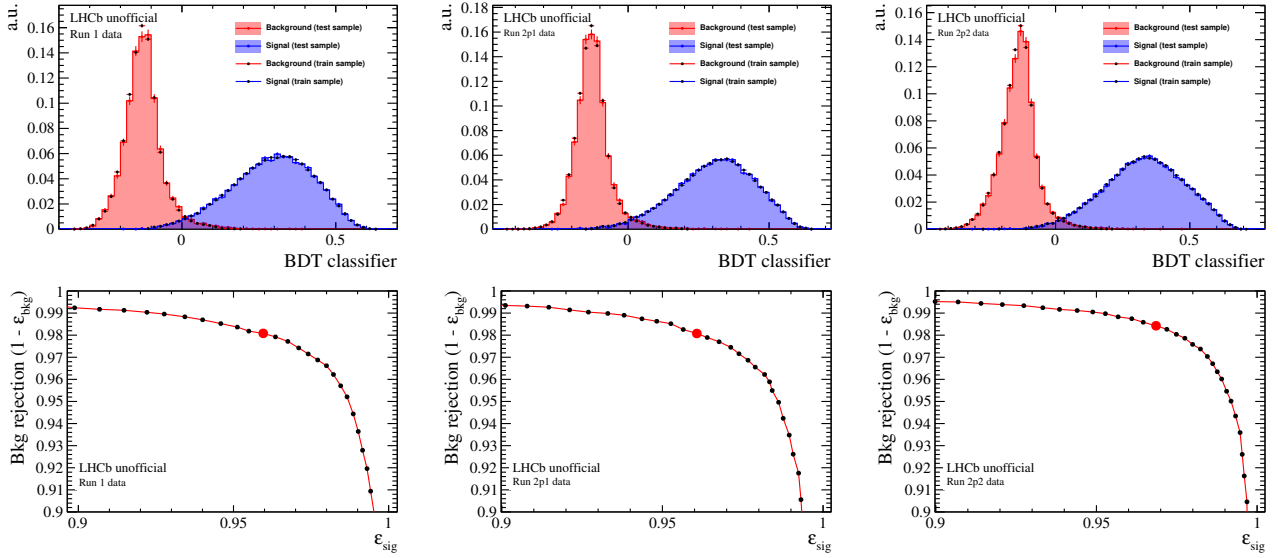


Figure 5.5: Overtraining check (top) and ROC (bottom) for the different BDTs, integrated over all folds. The BDT efficiency corresponding to the chosen selection is indicated.

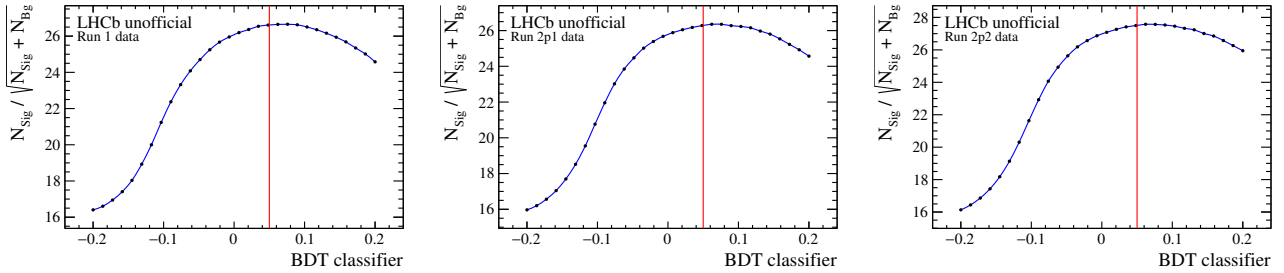


Figure 5.6: FOM in dependence of the BDT classifier for the four different BDTs. The red lines indicates the BDT response criterion which is used for analysis (*i.e.* 0.05).

Table 5.9: Percentage of events with multiple candidates in the fully selected data sets in a  $50 \text{ MeV}/c^2$  window around the  $B_s^0$  mass. In the case of simulation, this is truth matched. For 2012 rare mode data, no event with multiple candidate is found. For the other years, between 1 – 3 events with multiple candidates are found. The fluctuations are, therefore, of statistical nature given the small size of the rare mode data set.

|      | $B_s^0 \rightarrow \phi \mu^+ \mu^-$ sim. | $B_s^0 \rightarrow J/\psi \phi$ sim. | rare $q^2$ data | resonant $q^2$ data |
|------|-------------------------------------------|--------------------------------------|-----------------|---------------------|
| 2011 | 0.017%                                    | 0.004%                               | 0.787%          | 0.090%              |
| 2012 | 0.015%                                    | 0.008%                               | none            | 0.099%              |
| 2016 | 0.015%                                    | 0.005%                               | 0.674%          | 0.197%              |
| 2017 | 0.002%                                    | 0.035%                               | 0.355%          | 0.181%              |
| 2018 | 0.002%                                    | 0.056%                               | 0.149%          | 0.202%              |

## 5.5 Corrections to simulation

Simulation in LHCb is known to model some distributions of variables imperfectly. For this reason, the data-simulation agreement of variables used in this thesis is tested using  $B_s^0 \rightarrow J/\psi \phi$  control mode decays. The variables found to not be modelled sufficiently well are the track multiplicity (**nTracks**), the transverse momentum of the  $B_s^0$  meson ( $p_T(B_s^0)$ ), the L0 trigger efficiency and the response of particle identification (PID) variables. The first three variables are corrected for by reweighing the distributions, and the PID variables are corrected using the **MEERKAT** tool [225–227]. When determining the corrections as few selection criteria as necessary are used. Furthermore, the corrections are not determined on the rare mode, but on the control mode due to the higher yields. The weight-based corrections are obtained in the control mode and are also used to correct the rare mode simulation. Furthermore, the weight-based corrections are determined after taking the previous corrections into account. The track multiplicity weights are obtained first, the  $p_T(B_s^0)$  weights second and the L0 trigger efficiency correction is determined last.

### 5.5.1 Reweighting of the track multiplicity

The track multiplicity of simulated events is corrected by determining per-event weights from the comparison of data and fully reconstructed simulation of the control mode. These weights are then applied to both, control mode and rare mode simulation. The application to the rare mode utilises the fact that the decay mode of the  $B_s^0$  meson is largely independent of the track multiplicity. In Fig. 5.7 the **nTracks** distributions of  $B_s^0 \rightarrow J/\psi \phi$  and  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  simulation are shown, indicating this assumption to be valid.

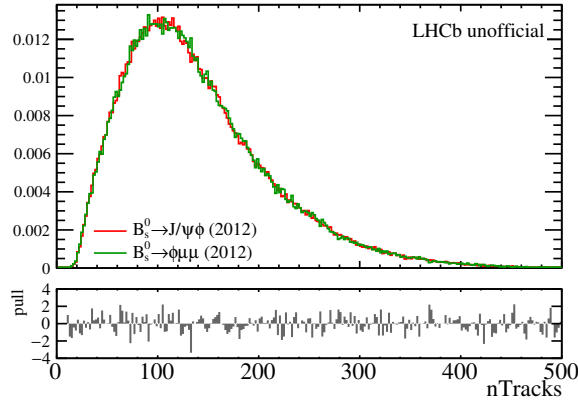


Figure 5.7: Comparison of **nTracks** between truth matched simulated samples of the control and rare decay modes directly after the stripping selection.  $B_s^0 \rightarrow J/\psi \phi$  simulation is compared to  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  for 2012 simulation conditions.

To derive the weights, the same selection is applied for the control mode data and simulation, namely any event needs to pass the trigger and stripping selections, as well as be in a  $\pm 50 \text{ MeV}/c^2$  region around the known mass of the  $B_s^0$  meson and a  $\pm 12 \text{ MeV}/c^2$  region around the known mass of the  $\phi$  meson. The combinatorial background is statistically removed from data by applying *sWeights* derived on a control mode sample with the same selection. The weights to



correct the track multiplicity are determined as the ratio of the thus selected events on data and simulation using the `nTracks` distribution, split into 50 bins with roughly equal number of entries, in the range of 0 to 600. Bins of equal population, as opposed to equal width, are taken to avoid large fluctuations in weights in regions of low density. The weights are shown in Fig. 5.8 for Run 1 Run 2p1 and Run 2p2 in different colors. The agreement of  $B_s^0 \rightarrow J/\psi \phi$  data and simulation is shown in Fig. 5.9 for the different data taking periods. In this figure, the data is shown as black markers, the green line indicates the uncorrected simulation whereas the red line represents fully corrected simulation, meaning simulation with all weight-based corrections applied. The agreement of the red line is greatly improved with respect to the green line. While the low `nTracks` region is compatible across the years, the weights in the high `nTracks` region are quite different across the years.

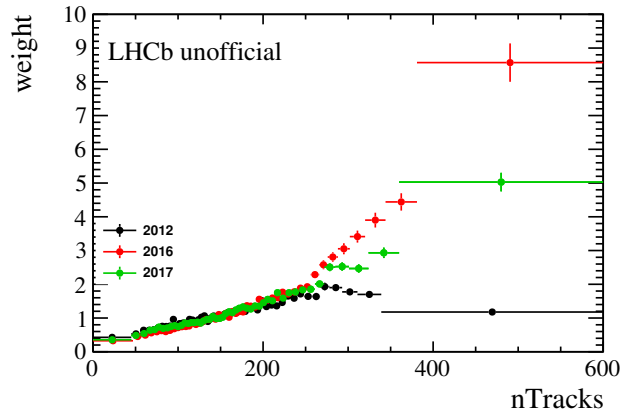


Figure 5.8: Weights to correct the `nTracks` distributions in the simulated samples for 2012 conditions in black, 2016 conditions in red and 2017 conditions in green.

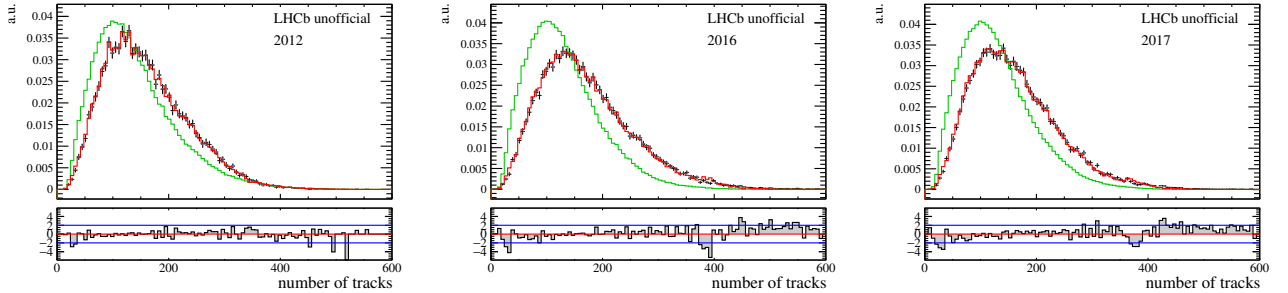


Figure 5.9: Comparison between *sWeighted*  $B_s^0 \rightarrow J/\psi \phi$  data and simulation for (left) 2012, (middle) 2016 and (right) 2017 data taking conditions. The data (black markers), raw (green line) and fully corrected simulation (red line) are shown, and the pulls indicate the difference between data and corrected simulation.

### 5.5.2 Corrections to the $B_s^0$ kinematics

As the transverse momentum of the  $B_s^0$  meson ( $p_T(B_s^0)$ ) is used in the BDT, good agreement between data and simulation in this variable is especially important. For this reason, the

deviation in  $p_T(B_s^0)$  from the control mode simulation with respect to the control mode data is corrected for using weights. The weights are derived using the full selection except for the BDT by comparing simulated  $B_s^0 \rightarrow J/\psi \phi$  events with the **nTracks** correction applied to *sWeighted* control mode data, in full analogy to the **nTracks** correction. The  $p_T(B_s^0)$  weights are determined in 60 equally populated bins in the range of  $0 \text{ MeV}/c^2 < p_T(B_s^0) < 60000 \text{ MeV}/c^2$ . The histograms of the weights for the three data sets can be found in Fig. 5.10, and the corresponding comparisons between raw or fully corrected simulation and *sWeighted*  $B_s^0 \rightarrow J/\psi \phi$  data in Fig. 5.11. The agreement of the corrected simulation and *sWeighted* control mode data is greatly improved.

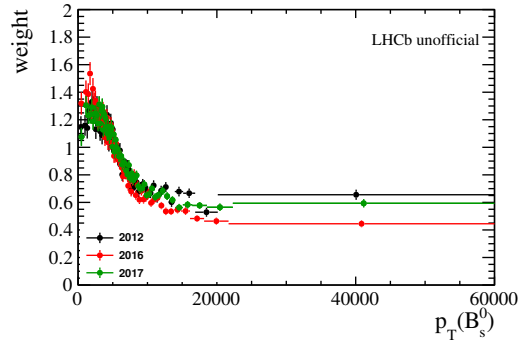


Figure 5.10: Weights to reweight the  $p_T(B_s^0)$  distributions for 2012 simulation in black, 2016 simulation in red and 2017 simulation in red.

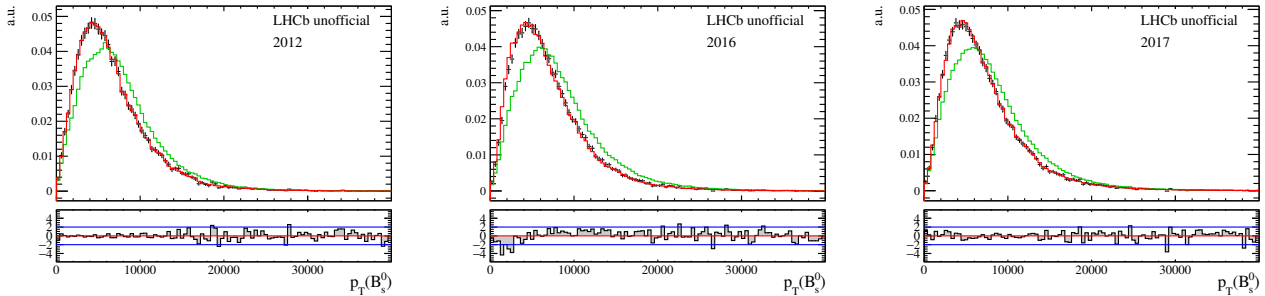


Figure 5.11: Comparison of *sWeighted*  $B_s^0 \rightarrow J/\psi \phi$  data and simulated samples for (left) 2012, (middle) 2016 and (right) 2017 data taking conditions. The data (black markers), raw (green line) and fully corrected simulation (red line) is shown, and the pulls indicate the difference between data and corrected simulation.

### 5.5.3 Correction to the L0 trigger

The third variable corrected for using weights is the L0 trigger response. Considering that data is only available after the trigger selection, one needs to carefully choose the definition of the trigger efficiency. Therefore, one cannot use the naive approach of defining the trigger efficiency of all events in the LHCb acceptance  $\epsilon_{\text{Trig}|\text{Acc}}$ , namely

$$\epsilon_{\text{Trig}|\text{Acc}} = \frac{N_{\text{Trig}|\text{Acc}}}{N_{\text{Acc}}},$$

where  $N_{\text{Trig|Acc}}$  is the number of events passing the trigger given they are in the LHCb acceptance and  $N_{\text{Acc}}$  is the number of events in the LHCb acceptance, as  $N_{\text{Acc}}$  is not directly observable. For this reason the **TISTOS** method [228] has been developed. In this method, the trigger efficiency is determined after applying a selection and is given as

$$\epsilon_{\text{Trig|Sel}} = \frac{N_{\text{Trig|Sel}}}{N_{\text{Sel}}},$$

where  $N_{\text{Sel}}$  is the number of events passing a given selection and  $N_{\text{Trig|Sel}}$  corresponds to the number of events passing the trigger requirements in addition to the given selection. Assuming the **TOS** trigger decision to have little to no correlation with the **TIS** trigger decision, the **TOS** efficiency can be determined on a sample of events triggered by the **TIS** trigger. It is then given as

$$\epsilon_{\text{TOS}} = \epsilon^{\text{TISTOS}} = \frac{N^{\text{TIS\&\&TOS}}}{N^{\text{TIS}}}.$$

The **TOS** efficiency determined on the **TIS** sample will be referred to as the **TISTOS** efficiency from hereon.

When determining the **TISTOS** efficiency each event is required to have passed the **TOS** requirement for the other trigger stages. For example the **L0** trigger sample requires both **HLT** trigger **TOS** selections to be passed. The **TIS** sample used to derive the **L0** trigger efficiencies is constructed by requiring a positive **TIS** trigger decision of the **L0Electron**, **L0Hadron**, **L0Photon** or **L0Muon** trigger lines. These trigger lines have been chosen to give a large sample for the efficiency determination. For the efficiencies of the **HLT** stages, the **TIS** sample is constructed by requiring a positive **TIS** decision of the corresponding **HLT** trigger lines given in Tab. 5.4. Considering that the **HLT** selection used in this analysis is already fairly inclusive, the sample is sufficiently large and does not require the addition of additional trigger lines.

The comparison of the **TISTOS** efficiency is done using control mode data and simulation. For both samples, the preselection and **BDT** are required to have been passed. Given that the efficiency is determined from events that need to fulfill the **TIS** decision compared to those fulfilling the **TIS** and **TOS** decision, the *sPlot* technique can not be used here. Therefore, the  $B_s^0$  candidates have to additionally be located in the invariant mass region of  $\pm 50 \text{ MeV}/c^2$  around the known mass of the  $B_s^0$  meson to get a clean data sample. After this set of selections, the data set is considered to be sufficiently clean and as such background subtraction is not needed. The simulated sample is also truth matched and has the corrections for the **nTracks** and  $p_{\text{T}}(B_s^0)$  applied. As the stripping contains an **HLT** filter for Run 1 data, as outlined in Tab. 5.5, it is also applied to the simulated sample. The **TISTOS** efficiencies, and subsequently the weights, are determined as a function of  $\max(p_{\text{T}}(\mu^+), p_{\text{T}}(\mu^-))$ . This choice allows the trigger response to be compared between data and simulation for different kinematic regions of the muons, even though the muon  $p_{\text{T}}$  distribution is different across the rare and control modes.

While the **TISTOS** efficiency is well simulated for the **HLT1** and **HLT2** trigger stages, as shown in Fig. 5.12 and Fig. 5.13, respectively, the **L0** trigger is not simulated sufficiently well and is, therefore, corrected.

The comparison of the binned **L0** **TISTOS** efficiencies ( $\epsilon^{\text{TISTOS}, \text{L0}}$ ) in data and simulation exhibits deviations, most notably in Run 2 as shown in Fig. 5.14. The disagreement originates from simulated events always being generated with a single trigger configuration (**TCK**), representing the configuration used for the majority of data events, while data comprises a mix of different configurations. When comparing only data that was taken with the **TCK** modelled

in simulation, the disagreement vanishes, as shown in Fig. 5.15. This difference in  $\epsilon^{\text{TISTOS}, \text{LO}}$  is corrected for using weights. They are calculated such that the L0 efficiency curve of the simulated sample aligns with the one found in data. The weights are calculated for all years independently and can be found in Fig. 5.16. The weights are close to unity for Run 1 as the agreement in Fig. 5.14 is already good.

The final agreement of the full TISTOS efficiency is shown in Fig. 5.17 between  $B_s^0 \rightarrow J/\psi \phi$  simulation with all weight-based corrections applied, shown as red markers, and data, shown as black markers. No significant deviations are found over the full maximum muon  $p_T$  distribution.

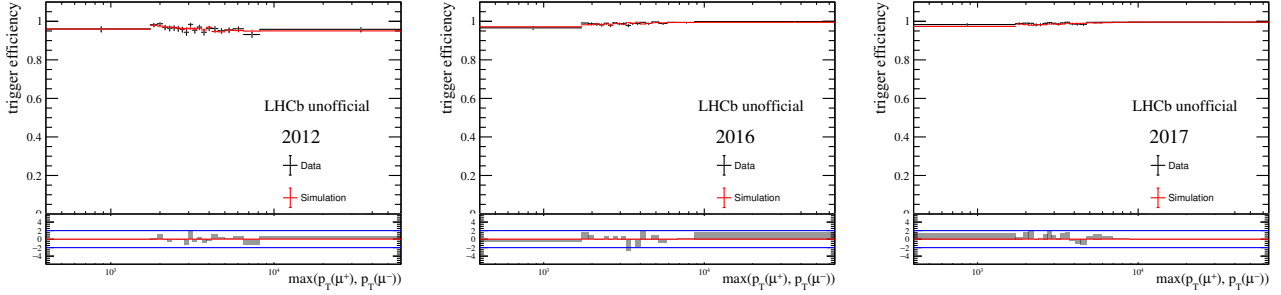


Figure 5.12: Comparison of HLT1 TISTOS efficiencies from  $B_s^0 \rightarrow J/\psi \phi$  data (black) and fully corrected simulation (red) for (left) 2012, (middle) 2016 and (right) 2017 data taking conditions in bins of the maximum  $p_T$  of the two muons. Good agreement is found between the distributions.

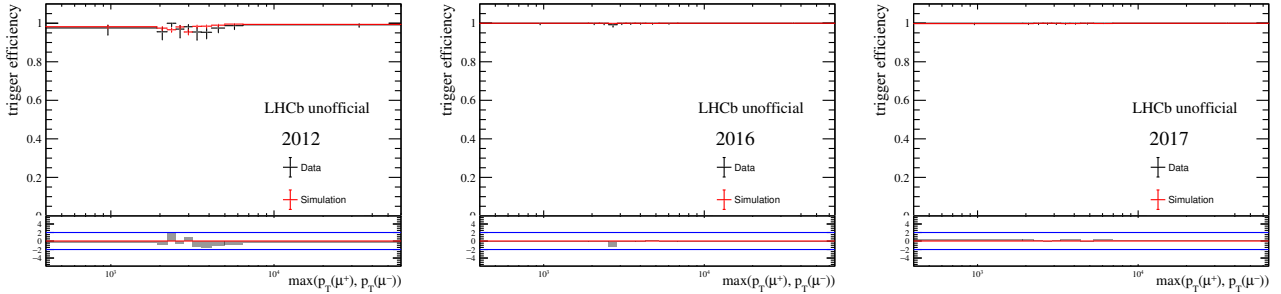


Figure 5.13: Comparison of HLT2 TISTOS efficiencies from  $B_s^0 \rightarrow J/\psi \phi$  data (black) and fully corrected simulation (red) for (left) 2012, (middle) 2016 and (right) 2017 data taking conditions in bins of the maximum  $p_T$  of the two muons. Good agreement is found between the distributions.

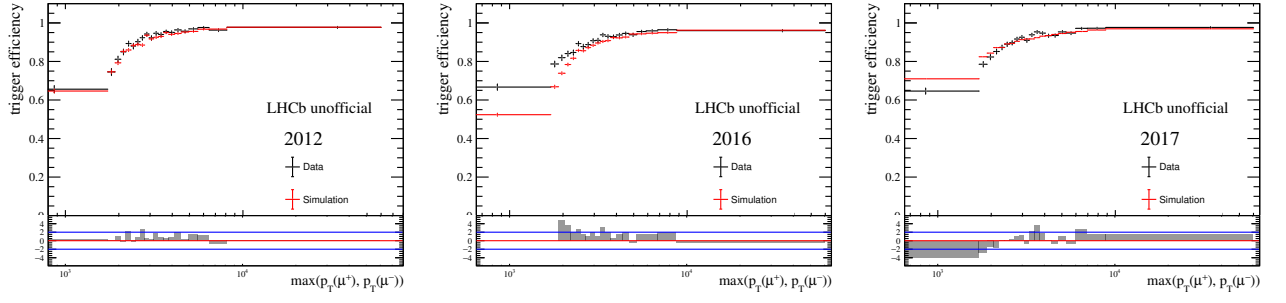


Figure 5.14: Comparison of L0 TISTOS efficiencies from  $B_s^0 \rightarrow J/\psi \phi$  data (black) and fully corrected simulation (red) for (left) 2012, (middle) 2016 and (right) 2017 data taking conditions in bins of the maximum  $p_T$  of the two muons. The agreement is very good for Run 1, while Run 2 exhibits sizeable differences originating from different TCK compositions used in simulation and data.

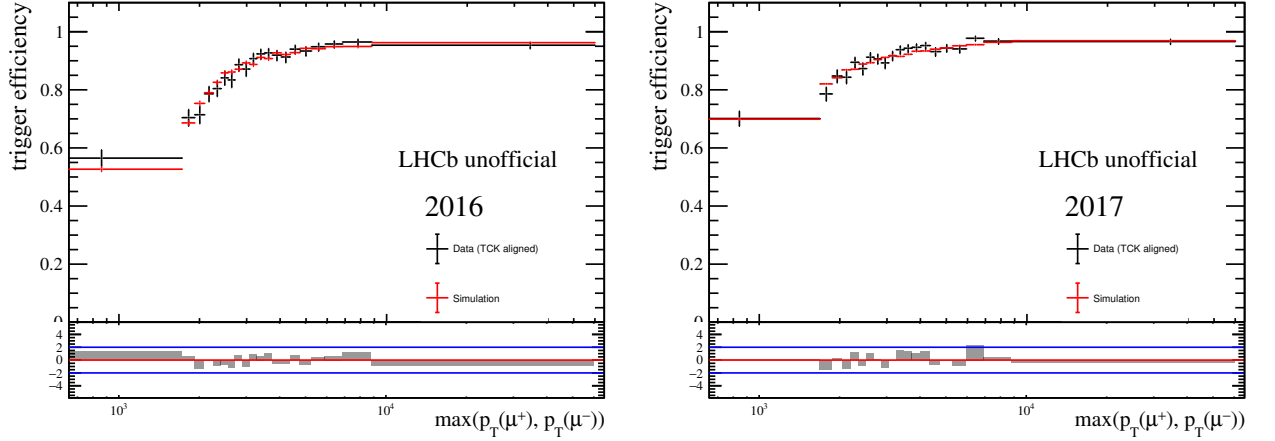


Figure 5.15: The L0 TISTOS efficiencies for data and simulation for (left) 2016 and (right) 2017 data taking conditions when data and simulation are restricted to the same TCK.

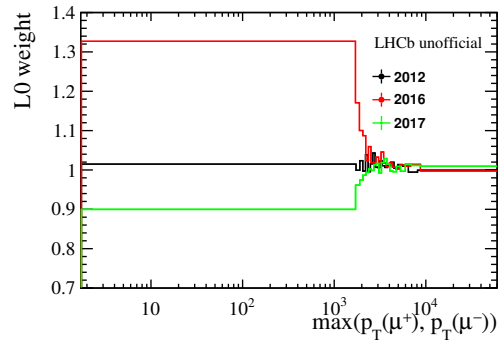


Figure 5.16: Weights determined to correct the L0 trigger efficiency for the 2012, 2016 and 2017 simulation conditions in black, red and green, respectively.

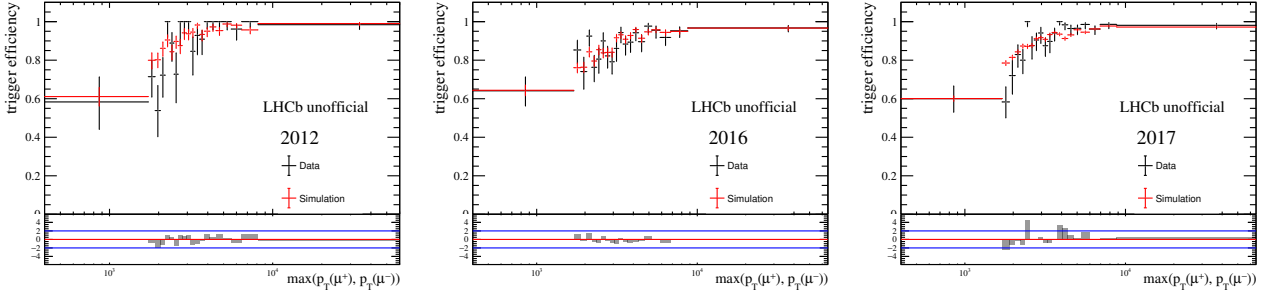


Figure 5.17: Comparison of total TISTOS efficiencies from  $B_s^0 \rightarrow J/\psi \phi$  data (black) and fully corrected simulation (red) for (left) 2016 and (right) 2017 data taking conditions in bins of the maximum  $p_T$  of the two muons for the different years. Good agreement is found between the distributions.

### 5.5.4 Correction of the PID response using MEERKAT

The PID distributions are not modeled perfectly in the simulation and can be corrected using clean calibration samples taken from data. These calibration samples are processed LHCb wide in the `PIDCalib` package [229] and are detailed in Tab. 5.10. From the available methods to correct the PID distributions in the `PIDCalib` package, the `MEERKAT` tool [225–227] is used in this thesis. The name `MEERKAT` is an abbreviation for “Multidimensional Efficiency Estimation using the Relative Kernel Approximation Technique” essentially describing the approach used in the tool. The particle identification variables of the calibration samples are parameterised in four dimensions using the kernel density estimation (KDE) technique. The four dimensions are the particle identification variable, the number of tracks of the event (`nTracks`) as well as the pseudorapidity ( $\eta$ ) and the transverse momentum ( $p_T$ ) of the particle of interest. The parameterisation is done in two steps, first each of the four dimensions probability density functions (PDF) are estimated using a KDE in a single dimension. In the second step, the one-dimensional PDFs are iteratively combined into a four-dimensional PDF. This approach to parameterise the four-dimensional distribution is shown to be able to describe fine features in Ref. [225].

Table 5.10: The decay modes used to correct the PID response of each particle type in the `PIDCalib` package.

| Particle | Decay mode                                             |
|----------|--------------------------------------------------------|
| $K$      | $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ |
| $\pi$    | $D^{*+} \rightarrow (D^0 \rightarrow K^- \pi^+) \pi^+$ |
| $p$      | $\Lambda_c^+ \rightarrow p K^- \pi^+$                  |
| $\mu$    | $J/\psi \rightarrow \mu^+ \mu^-$                       |
| $e$      | $B^+ \rightarrow (\psi \rightarrow e^+ e^-) K^+$       |

The `MEERKAT` tool supports two different subpackages to correct the PID distributions, the `PIDGen` and `PIDCorr` packages. The `PIDGen` package uses a resampling approach, in which the corrected PID variable is randomly drawn from a one-dimensional PID distribution in the range of the lower bound of the distribution up to the original value of the PID variable in the event. The one-dimensional PID distribution is generated by evaluating the four-dimensional phasespace at the point in  $p_T$ ,  $\eta$  and `nTracks` for the given candidate being resampled. This approach does not capture the correlation between PID variables for the same candidate and consequently combinations of corrected variables are often not correctly described.

To correctly describe the particle identification variable and combinations of different variables, the `PIDCorr` package can be used. In the `PIDCorr` package a simulated sample is used in addition to the calibration sample from data and a variable transformation method is employed instead. The distributions of the particle identification variable from data and simulation are obtained in the same way as in the `PIDGen` package. To determine the corrected value for the particle identification variable  $x_{\text{corr}}$ , first the point  $x_{\text{MC}}$  needs to be determined, which is the point at which the cumulative probability density for the particle identification variable in the simulated sample is same as in the data sample for a given point in  $\eta$ ,  $p_T$  and `nTracks`. The corrected value  $x_{\text{corr}}$  is then given by the transformation

$$x_{\text{corr}} = P_{\text{exp}}^{-1} (P_{\text{MC}}(x_{\text{MC}} | p_T, \eta, \text{nTracks}) | p_T, \eta, \text{nTracks}) ,$$

where  $P$  is the cumulative PDF.

Which of the two methods is used to correct the PID response depends on the PID variable and the type of the track at generator-level. While only using **PIDCorr** is preferable, this method is typically not available for tracks generated as leptons. Therefore, PID variables of leptons are corrected for with the **PIDGen** package, whereas hadron PID variables are corrected for with the **PIDCorr**, whenever available. The used **MEERKAT** subpackage for the PID variables used in this analysis is shown in Tab. 5.11 for each of the different generated particle species.

Table 5.11: The **MEERKAT** subpackage used for the correction of the particle identification variables used in this analysis is shown for the different generated particle species.

| particle identification variable | particle species    | MEERKAT subpackage |
|----------------------------------|---------------------|--------------------|
| ProbNNk                          | $K, \pi, p$         | PIDCorr            |
| ProbNNk                          | $\mu, e$            | PIDGen             |
| ProbNNp                          | $K, \pi, p$         | PIDCorr            |
| ProbNNp                          | $\mu, e$            | PIDGen             |
| PIDmu                            | $\mu, e, K, \pi, p$ | PIDGen             |
| ProbNNmu                         | $\mu, e, K, \pi, p$ | PIDGen             |

### 5.5.5 Comparison of control mode data and simulation

In this section the effects of the corrections are shown using  $B_s^0 \rightarrow J/\psi \phi$  data and  $B_s^0 \rightarrow J/\psi \phi$  simulation generated according to the conditions in 2017. Figures for the remaining data taking periods are shown in Figs. B.1-B.6 located in App. B. For both samples the stripping and trigger selections have been applied. Furthermore, the dikaon mass is required to be within  $\pm 12 \text{ MeV}/c^2$  of the known mass of the  $\phi$  meson, and  $q^2$  has to be within  $[8.0 - 11.0] \text{ GeV}^2/c^4$ . The control mode simulation is shown in green without corrections and in red with corrections for **nTracks**,  $p_T(B_s^0)$ , the L0 trigger response and the PID response. After all corrections, data and simulation agree sufficiently well. There are minor discrepancies in the  $B_s^0$  meson's impact parameter as well as the vertex  $\chi^2$  remaining when comparing 2017 data to simulation, as shown Fig. 5.18. This is the most likely reason for the slight difference in the BDT response.

In order to estimate the effect of this difference, the BDT efficiency for  $B_s^0 \rightarrow J/\psi \phi$  events is compared between *sWeighted* data and simulation. In data the BDT efficiency is found to be 96.700% for 2017, and to be 96.670% in simulation. This is a relative difference of less than 0.1% and will be neglected in the following, as it is sufficiently small and well-covered by the systematic uncertainty assigned to the corrections of the simulation, which is discussed in Sec. 8.6.



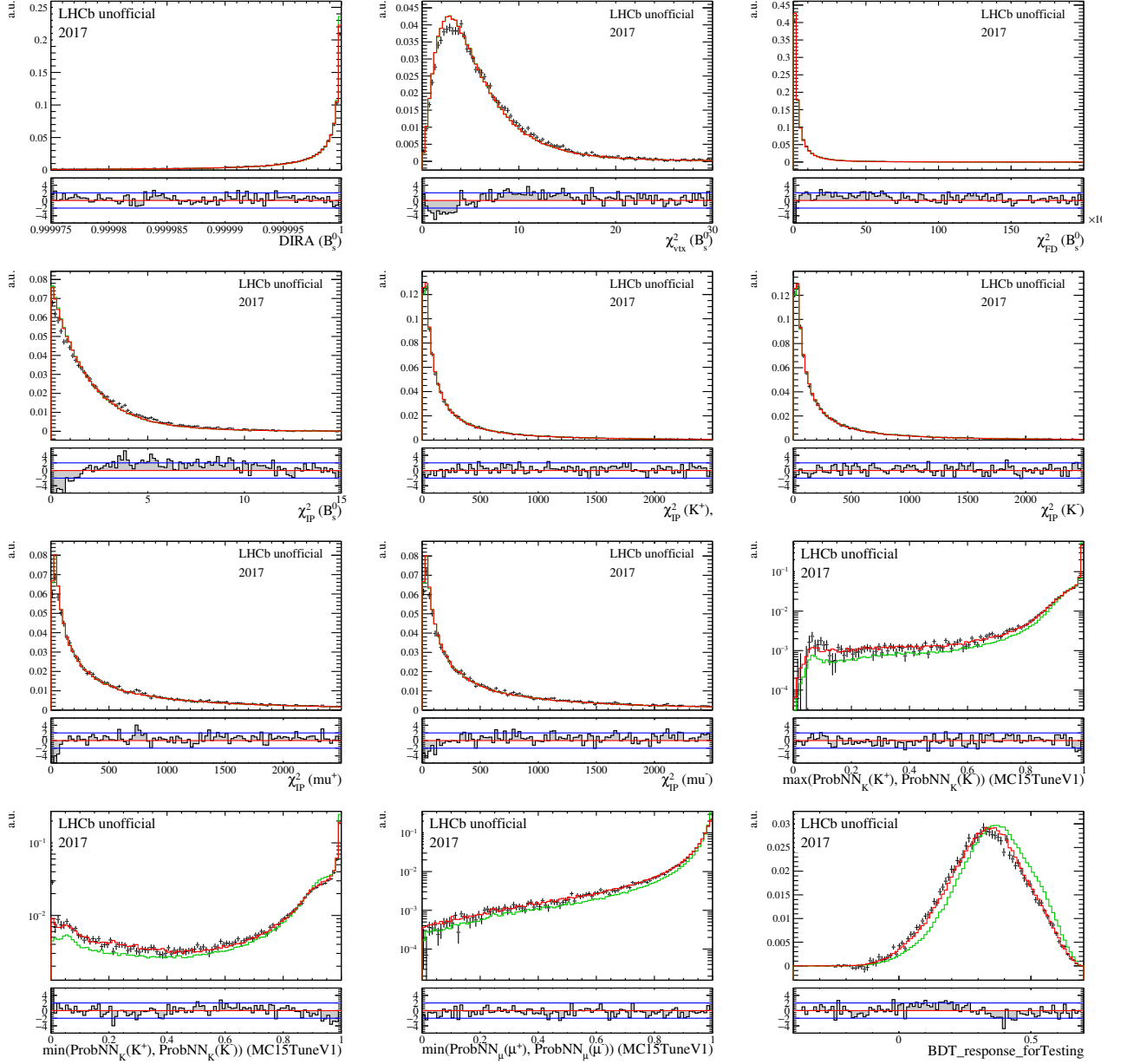


Figure 5.18: Data and simulation comparison of the BDT input variables and the BDT response (bottom right) for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2017 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

# Chapter 6

## Signal and background yields

As the selection for this analysis is developed with the signal window taken as  $\pm 50 \text{ MeV}/c^2$  around the known  $B_s^0$  meson mass in the rare mode blinded, rare mode data cannot be used to determine the signal and background yields and, therefore, the relative contribution of background sources. As a consequence the signal and background yields must be estimated. This chapter details how the expected signal yields and contributions from peaking backgrounds are estimated.

The total number of events produced at the LHCb experiment,  $N_{\text{prod}}$ , can be calculated using the production cross section,  $\sigma_{\text{prod}}$ , and the integrated luminosity,  $\mathcal{L}_{\text{int}}$ , using the formula

$$N_{\text{prod}} = \sigma_{\text{prod}} \times \mathcal{L}_{\text{int}} .$$

The production cross section for a given  $B$  hadron of type  $B_i$  decaying into a specific final state is given by

$$(6.1) \quad \sigma_{\text{prod}} = \sigma(p\bar{p} \rightarrow b\bar{b}X) \times f_i \times \mathcal{B}(B_i \rightarrow \text{final state}) ,$$

where  $\sigma(p\bar{p} \rightarrow b\bar{b}X)$  is the production cross section of a  $b\bar{b}$  pair in a  $p\bar{p}$  collision,  $i$  indicates the type of the  $B$  hadron,  $f_i$  is the hadronisation fraction, which describes the probability for a  $b$  quark to hadronise into a  $B$  hadron of type  $i$ , and  $\mathcal{B}(B_i \rightarrow \text{final state})$  is the branching fraction of the  $B$  hadron of type  $i$  to decay into the final state of interest. When considering the decay of a  $\Lambda_b^0$  baryon the same approach and formula is used with the corresponding constants. Given that not every produced event is selected in the final data set, the selection efficiency ( $\epsilon$ ) is taken into account and the number of selected events ( $N_{\text{sel}}$ ) is determined as

$$(6.2) \quad N_{\text{sel}} = \epsilon \times N_{\text{prod}} = \epsilon \times \sigma_{\text{prod}} \times \mathcal{L}_{\text{int}} .$$

While this equation can in principle be used to determine the expected yield for a given decay, in practise a normalisation channel is used to determine the expected yield. The advantage of using a normalisation channel is the cancellation of factors with potentially large uncertainties, such as the  $b\bar{b}X$  production cross section with a relative uncertainty of roughly 10% [89]. Using Eq. 6.1 and Eq. 6.2 the ratio of yields of the decay of interest ( $N_{B_i \rightarrow \text{final state}}$ ) with respect to the yield of the normalisation channel ( $N_{\text{norm}}$ ) is then given by

$$\frac{N_{B_i \rightarrow \text{final state}}}{N_{\text{norm}}} = \frac{\epsilon_{B_i \rightarrow \text{final state}}}{\epsilon_{\text{norm}}} \times \frac{f_i}{f_{\text{norm}}} \times \frac{\mathcal{B}(B_i \rightarrow \text{final state})}{\mathcal{B}(\text{norm})} \times \frac{\mathcal{L}_{\text{int}}^{B_i \rightarrow \text{final state}}}{\mathcal{L}_{\text{int}}^{\text{norm}}} .$$

Rearranging this equation and using the control mode  $B_s^0 \rightarrow J/\psi \phi$  as the normalisation channel from the same data set, the yield of the decay of interest can be calculated by

$$(6.3) \quad N_{B_i \rightarrow \text{final state}} = \frac{N_{B_s^0 \rightarrow J/\psi \phi}^{\text{data}}}{\epsilon_{B_s^0 \rightarrow J/\psi \phi}} \times \frac{f_i}{f_s} \times \epsilon_{B_i \rightarrow \text{final state}} \times \frac{\mathcal{B}(B_i \rightarrow \text{final state})}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)},$$

where the superscript data indicates that the yield of the normalisation channel is taken from data.

## 6.1 Considered decay modes

In this thesis, two different type of backgrounds are considered: peaking backgrounds and combinatorial background. The peaking backgrounds can originate either from a single final state particle misidentification or from a double misidentification of final state particles. The considered peaking backgrounds will be introduced in the following. The combinatorial background will not be discussed in this Chapter as they are included in the fit model as discussed in Sec. 7.2.

The final states of the decays  $B^0 \rightarrow (K^{*0} \rightarrow K^+ \pi^-) \mu^+ \mu^-$  and  $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$  can be misidentified as the  $K^+ K^- \mu^+ \mu^-$  final state with a single pion or proton misidentification, respectively. They are single misidentification backgrounds in this analysis and considering their related topologies and comparably large branching fractions can contribute to the  $K^+ K^- \mu^+ \mu^-$  spectrum. Contributions from the charmonium tree-level decays  $B_s^0 \rightarrow J/\psi \phi$  ( $B^0 \rightarrow J/\psi K^{*0}$ ) occur via the double misidentification background. Here, a kaon (pion) is misidentified as a muon and a muon of the same charge is misidentified as a kaon of  $B_s^0 \rightarrow J/\psi \phi$  ( $B^0 \rightarrow J/\psi K^{*0}$ ) decays. If this double misidentification occurs these decays can also be misreconstructed as signal as the invariant dimuon mass can fall into the rare mode  $q^2$  region. Given their large branching fractions with respect to the signal branching fraction, these double misidentification backgrounds can also constitute a sizeable source of background events. All considered decay modes and their branching and hadronisation fractions are given in Tab. 6.1, where the values for the branching fractions are mostly taken from Ref. [65] and the hadronisation fractions are taken from Ref. [230]. The branching fraction for  $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$  decays has not been measured yet and has, therefore, been taken from the SM prediction [231] for the branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-$  decays and the fit fraction of  $\Lambda_b^0 \rightarrow J/\psi p K^-$  decays in the full angular fit of  $\Lambda_b^0 \rightarrow J/\psi \Lambda^0(1520)$  decays [232]. The  $\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-$  SM prediction depends on the choice of form factors. There are two sets of form factors [233] available, the SCA form factors, based on a non-relativistic reduction of the quark spinors and computed using a single component wave function, and the MCN form factors, based on the full relativistic form for the quark spinors and the multi-component quark model wave function. The SM predictions based on these form factors are found to be [233]

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-) = 1.3 \times 10^{-7}(\text{SCN}) \text{ or } 2.1 \times 10^{-7}(\text{MCN}).$$

Given the MCN form factors make use of fewer assumptions, they are expected to be more reliable. Furthermore, their prediction for the branching fraction yields the larger value such that using this value can be considered a more conservative approach. The distance between the two predictions is taken as the uncertainty on the branching fraction. Therefore, the value

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-) = (2.1 \pm 0.8) \times 10^{-7}$$

is used as the branching fraction for the calculation. This prediction then needs to be scaled by the branching fraction of a  $\Lambda^0(1520)$  baryon decaying into the  $pK^-$  final state. This is given in Ref. [65] as

$$\mathcal{B}(\Lambda^0(1520) \rightarrow pK^-) = 22.5\%.$$

Lastly, the ratio of  $\Lambda_b^0 \rightarrow J/\psi \Lambda^0(1520)$  decays excluding decays with pentaquark intermediary states to  $\Lambda_b^0 \rightarrow J/\psi (\Lambda^0(1520) \rightarrow pK^-)$  decays has been measured in data [232] to be

$$\frac{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi pK^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi (\Lambda^0(1520) \rightarrow pK^-))} = 6.05.$$

Assuming the hadronic structure in the  $J/\psi$  mode and the rare mode to be the same, the SM prediction is scaled by this factor. Using these components, the branching fraction of the  $\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-$  decay is given as

$$\mathcal{B}(\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-) = (2.9 \pm 1.1) \times 10^{-7},$$

with the uncertainty being propagated from the uncertainty on  $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-)$ .

Table 6.1: Numerical values for branching and hadronisation fractions used to calculate the expected number of signal events. If applicable, the branching fraction of a meson such as the  $\phi$  or  $K^{*0}$  into the hadronic final state particles or the  $J/\psi$  into a pair of muons is given as the hadronic (had.) and leptonic (lep.) branching fraction, respectively.

| Decay                                      | $\mathcal{B}$           | had. $\mathcal{B}$ | lep. $\mathcal{B}$ | $f_i/f_S$ |
|--------------------------------------------|-------------------------|--------------------|--------------------|-----------|
| $B^0 \rightarrow J/\psi K^{*0}$            | $1.27(5) \cdot 10^{-3}$ | 66.6%              | 0.05961(33)        | 3.86(30)  |
| $B^0 \rightarrow K^{*0} \mu^+ \mu^-$       | $9.4(5) \cdot 10^{-7}$  | 66.6%              | 1                  | 3.86(30)  |
| $\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-$ | $2.9(11) \cdot 10^{-7}$ | -                  | 1                  | 1.85(33)  |
| $B_s^0 \rightarrow J/\psi \phi$            | $1.08(8) \cdot 10^{-3}$ | 49.2(5)%           | 0.05961(33)        | 1         |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$       | $8.2(12) \cdot 10^{-7}$ | 49.2(5)%           | 1                  | 1         |

## 6.2 Determination of selection efficiencies

The selection efficiencies of the control, rare and background modes needed in Eq. 6.3 are determined using simulation. For both, the signal mode and the control mode, the simulation is corrected as described in Chap. 5.5, whereas only the PID response is corrected for background modes as the weight-based corrections are not easily transferable between different final states. However, as will be seen in Sec. 6.3, the contributions from peaking background are found to be negligible and not taking the small effect from neglecting the corrections into account in the estimation is in turn a negligible effect. The efficiency of a given decay is determined by

$$(6.4) \quad \epsilon_{B_i \rightarrow \text{final state}} = \frac{N_{\text{all sel}}^{\text{sim}}}{N_{\text{DST}}^{\text{sim}}} \times \epsilon_{\text{acc}},$$

where  $N_{\text{sel}}^{\text{sim}}$  is the number of events selected in the simulation for the decay  $B_i \rightarrow \text{final state}$ ,  $N_{\text{DST}}^{\text{sim}}$  is the number of events generated within the LHCb acceptance and saved on tape and  $\epsilon_{\text{acc}}$  is the efficiency of the LHCb acceptance.

The numerical values for the number of events on tape within the LHCb acceptance, the number of events after all selection as well as the LHCb acceptance efficiency for the 2012, 2016 and 2017 data taking conditions are given in Tab. 6.2. The simulated samples for  $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$  decays are generated with a phase-space model for the kaon-proton and the dimuon spectra and are, therefore, not expected to correctly model the decay. These sample are, therefore, reweighted to match the shape of the  $q^2$  distribution given by the theory prediction for  $\Lambda_b^0 \rightarrow \Lambda^0\mu^+\mu^-$  [234] and the invariant mass distribution of the  $pK^-$  system found in the full angular analysis of  $\Lambda_b^0 \rightarrow J/\psi\Lambda^0(1520)$  decays [232]. The distributions before and after reweighting are shown in Fig. 6.1.

Table 6.2: LHCb acceptance efficiency and number of events on tape ( $N_{\text{DST}}^{\text{sim}}$ ) of the simulated background decays. No filtering is applied to any sample.

| Decay                                    | year | $\epsilon_{\text{acc.}} [\%]$ | $N_{\text{DST}}^{\text{sim}}$ | $N_{\text{all sel}}^{\text{sim}}$ |
|------------------------------------------|------|-------------------------------|-------------------------------|-----------------------------------|
| $B^0 \rightarrow K^{*0}\mu^+\mu^-$       | 2012 | 16.009(33)                    | 2520250                       | 61                                |
| $B^0 \rightarrow K^{*0}\mu^+\mu^-$       | 2016 | 17.045(39)                    | 1365262                       | 35                                |
| $B^0 \rightarrow K^{*0}\mu^+\mu^-$       | 2017 | 17.127(59)                    | 4014543                       | 117                               |
| $B^0 \rightarrow J/\psi K^{*0}$          | 2012 | 15.7(1)                       | 4422426                       | 0                                 |
| $B^0 \rightarrow J/\psi K^{*0}$          | 2016 | 16.8(1)                       | 2449607                       | 0                                 |
| $B^0 \rightarrow J/\psi K^{*0}$          | 2017 | 16.701(57)                    | 4127863                       | 0                                 |
| $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$ | 2012 | 16.300(35)                    | 2019539                       | 179                               |
| $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$ | 2016 | 17.303(46)                    | 7974997                       | 681                               |
| $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$ | 2017 | 17.297(70)                    | 2225233                       | 304                               |
| $B_s^0 \rightarrow J/\psi\phi$           | 2012 | 16.744(31)                    | 2014253                       | 0                                 |
| $B_s^0 \rightarrow J/\psi\phi$           | 2016 | 17.770(43)                    | 3975972                       | 0                                 |
| $B_s^0 \rightarrow J/\psi\phi$           | 2017 | 17.751(25)                    | 20035067                      | 0                                 |
| $B_s^0 \rightarrow \phi\mu^+\mu^-$       | 2012 | 16.933(23)                    | 2046275                       | 77200                             |
| $B_s^0 \rightarrow \phi\mu^+\mu^-$       | 2016 | 17.938(23)                    | 3917516                       | 153046                            |
| $B_s^0 \rightarrow \phi\mu^+\mu^-$       | 2017 | 17.939(34)                    | 7992126                       | 350728                            |

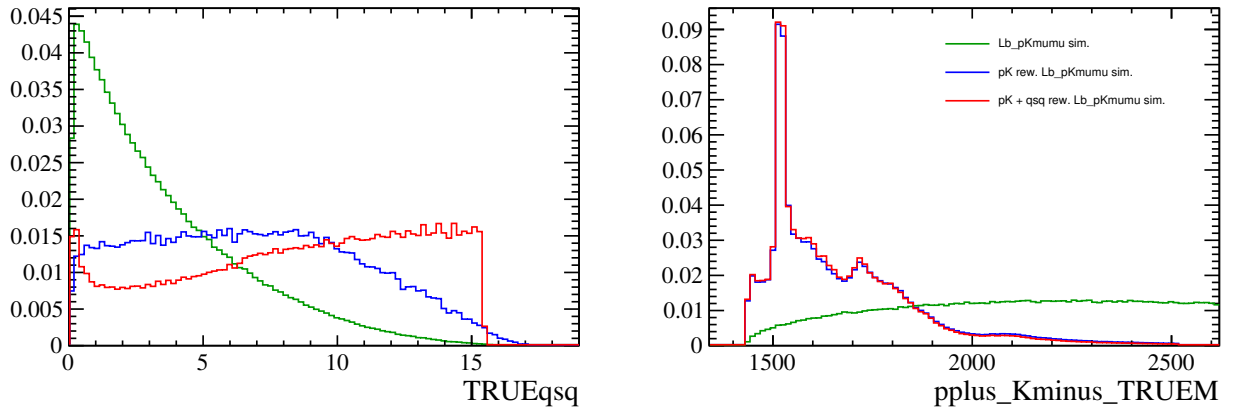


Figure 6.1: Comparison of unweighted (green) and weighted (red) generator-level simulation of  $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$  decays. First, the weights to align the  $pK^-$ -mass spectrum to what is found in  $\Lambda_b^0 \rightarrow J/\psi pK^-$  data, are applied (blue), then also the weights to align the  $q^2$  distribution to the theory prediction (red). The (left)  $q^2$  and the (right)  $pK^-$ -mass spectra are shown.

The yields of the control mode as found in fits to data as well as the respective selection efficiencies needed in Eq. 6.3 are summarised in Tab. 6.3. The yields and efficiencies are combined

into an efficiency-corrected number of  $B_s^0 \rightarrow J/\psi \phi$  events according to the data sets. They are given in Tab. 6.4.

Table 6.3: Yields in data and corresponding efficiencies for  $B_s^0 \rightarrow J/\psi \phi$  decays across the years.

|      | $N_{B_s^0 \rightarrow J/\psi \phi}^{\text{data}}$ | $\epsilon_{B_s^0 \rightarrow J/\psi \phi} [\%]$ |
|------|---------------------------------------------------|-------------------------------------------------|
| 2011 | $19570 \pm 150$                                   | $1.2410 \pm 0.0041$                             |
| 2012 | $43510 \pm 230$                                   | $1.1566 \pm 0.0037$                             |
| 2016 | $61500 \pm 270$                                   | $1.3347 \pm 0.0040$                             |
| 2017 | $66980 \pm 280$                                   | $1.4363 \pm 0.0050$                             |
| 2018 | $81850 \pm 310$                                   | $1.3533 \pm 0.0047$                             |

Table 6.4: Combination of efficiency-corrected  $B_s^0 \rightarrow J/\psi \phi$  events for the different runs.

|         | $N_{B_s^0 \rightarrow J/\psi \phi} / \epsilon_{B_s^0 \rightarrow J/\psi \phi}$ |
|---------|--------------------------------------------------------------------------------|
| Run 1   | $(5.338 \pm 0.026) \cdot 10^6$                                                 |
| Run 2p1 | $(4.607 \pm 0.025) \cdot 10^6$                                                 |
| Run 2p2 | $(10.711 \pm 0.040) \cdot 10^6$                                                |

### 6.3 Expected signal and background yields

The signal and background yields in Run 1, Run 2p1 or Run 2p2 are calculated with Eq. 6.3. As inputs the efficiency-corrected normalisation mode yields given in Tab. 6.4 are used along the branching and hadronisation fractions given in Tab. 6.1 and the efficiencies for a given final state calculated according to Eq. 6.4 using Tab. 6.2. For the double misidentification backgrounds, no events pass the full selection in the simulated samples. Therefore, limits at 90% confidence level are determined and found to be less than 1.03, 0.56 and 0.22 events for  $B_s^0 \rightarrow J/\psi \phi$  decays in the region of  $\pm 50 \text{ MeV}/c^2$  around the known  $B_s^0$  mass for Run 1, Run 2p1 and Run 2p2, respectively, and less than 1.12, 0.95 and 2.23  $B^0 \rightarrow J/\psi K^{*0}$  decays for Run 1, Run 2p1 and Run 2p2, respectively. To have a better estimate for the background from  $B_s^0 \rightarrow J/\psi \phi$  decays, the efficiency to select a background event is assumed to be independent of the year and the three simulated samples are added to determine a single efficiency for all three years. Furthermore, the effect of the charmonium vetoes on the  $B_s^0 \rightarrow J/\psi \phi$  and  $B^0 \rightarrow J/\psi K^{*0}$  decays is assumed to factorise with the rest of the analysis specific preselection. These assumptions reduce the limits on the background contribution from  $B_s^0 \rightarrow J/\psi \phi$  decays to 0.03, 0.03 and 0.07 for the three Run periods, respectively, and to less than 0.74, 0.64 and 1.42 for  $B^0 \rightarrow J/\psi K^{*0}$  decays in the three Run periods.

The expected amount of signal and background in the rare  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  channel is summarised in Tab. 6.5. No single source of peaking backgrounds is expected to contribute more than 0.5% in any data set and will as such not be modelled in the fit. The effect of this choice on the angular observables is determined as a  $q^2$ -dependent systematic uncertainty in Sec. 8.2.

Table 6.5: Total number of expected events in the rare signal region, with applied  $q^2$  vetoes and in a  $B_s^0$  mass window of  $\pm 50 \text{ MeV}/c^2$  around the known value for the mass of the  $B_s^0$  meson, for background modes and the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  signal.

| Decay                                           | Run 1           | Run 2p1         | Run 2p2         |
|-------------------------------------------------|-----------------|-----------------|-----------------|
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$            | $434 \pm 63$    | $411 \pm 60$    | $1070 \pm 160$  |
| $B^0 \rightarrow K^{*0} \mu^+ \mu^-$            | $1.58 \pm 0.25$ | $1.53 \pm 0.29$ | $4.08 \pm 0.54$ |
| $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$     | $1.71 \pm 0.73$ | $1.50 \pm 0.62$ | $5.6 \pm 2.4$   |
| $B_s^0 \rightarrow J/\psi \phi$ (double mis-id) | $< 0.03$        | $< 0.03$        | $< 0.07$        |
| $B^0 \rightarrow J/\psi K^{*0}$ (double mis-id) | $< 0.74$        | $< 0.64$        | $< 1.42$        |

# Chapter 7

## Angular analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$

In this Chapter, the angular analysis of the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  will be discussed. The angular acceptance modelling the efficiency shape in the decay angles and  $q^2$  is introduced in Sec. 7.1. The model used to perform the four-dimensional fit to the decay angles and the mass of the  $B_s^0$  meson candidate is discussed in Sec. 7.2. In Sec. 7.3 the validation of the fit using simulated events,  $B_s^0 \rightarrow J/\psi \phi$  control mode decays, and pseudoexperiments is detailed.

### 7.1 Angular acceptance

The angular and  $q^2$  distributions of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay are significantly distorted by the detector geometry, as well as the trigger, reconstruction and selection used in this analysis. These effects are referred to as acceptance effects and it is crucial that they are properly accounted for. The efficiency shape is parameterised in four dimensions, the three decay angles and  $q^2$ , and is denoted as  $\epsilon(\cos \theta_\ell, \cos \theta_K, \phi, q^2)$ . For the parameterisation, a product of Legendre polynomials  $L$  of order  $i, j, k$  and  $l$  are used. The efficiency shape is thus parameterised as

$$\epsilon(\cos \theta_K, \cos \theta_\ell, \phi, q^2) = \sum_{i,j,k,l} c_{ijkl} L_i(\cos \theta_\ell) L_j(\cos \theta_K) L_k(\phi) L_l(q^2),$$

where the  $c_{ijkl}$  are coefficients determined on simulated events using the method of moments [235, 236], taking advantage of the orthogonality of the Legendre polynomials. The coefficients are determined by

$$c_{i,j,k,l} = \frac{1}{N'} \sum_{e=1}^N w_e^{\text{corr}} \left[ \left( \frac{2i+1}{2} \right) \left( \frac{2j+1}{2} \right) \left( \frac{2k+1}{2} \right) \left( \frac{2l+1}{2} \right) \right. \\ \left. \times L_i(\cos \theta_\ell) L_j(\cos \theta_K) L_k(\phi) L_l(q^2) \right],$$

where  $N$  is the number of simulated events,  $w_e^{\text{corr}}$  will be weighting factors to correct for the non-flat  $q^2$  distribution of the phasespace simulation as discussed below and additionally absorb the corrections to the simulation discussed in Sec. 5.5, and  $N'$  is a normalization factor defined as  $N' = \sum_{e=1}^N w_e^{\text{corr}}$ . The parameterisation is calculated up-to and including orders  $i \leq 4$ ,  $j \leq 2$  (4 for Run 2),  $k \leq 6$  and  $l \leq 7$ , which was found to be the lowest set of orders necessary to

sufficiently describe the acceptance shape. Considering the flavour-symmetrical final state, only even orders in the decay angles are used in the parameterisation. A systematic uncertainty to this specific choice of orders for the acceptance is assigned, as discussed in Sec. 8.3.

The acceptance is determined on fully reconstructed and selected simulated events of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay. The decay angles are described using a phasespace model, meaning they are generated flat in all three angles, whereas the distribution of  $q^2$  is not generated flat. This choice of generation model is advantageous, as the efficiency shape in the fully reconstructed and selected sample in the decay angles is the shape of the acceptance effect and can be parameterised directly. As the  $q^2$  spectrum is not flat at the generator-level, weights are determined as a function of  $q^2$  on a large sample of 50M events at generator-level to correct for the non-flatness. These weights are then applied to the fully reconstructed and selected sample, such that also the  $q^2$  efficiency shape corresponds directly to the shape induced by the acceptance effect. The effect of the weights on the  $q^2$  distribution of the generator-level sample is shown in Fig. 7.1. As the angular distribution generated under a phasespace model is independent of  $q^2$ , the angular distributions also remain flat at generator level after the application of the weights to correct for the non-flatness in  $q^2$ .

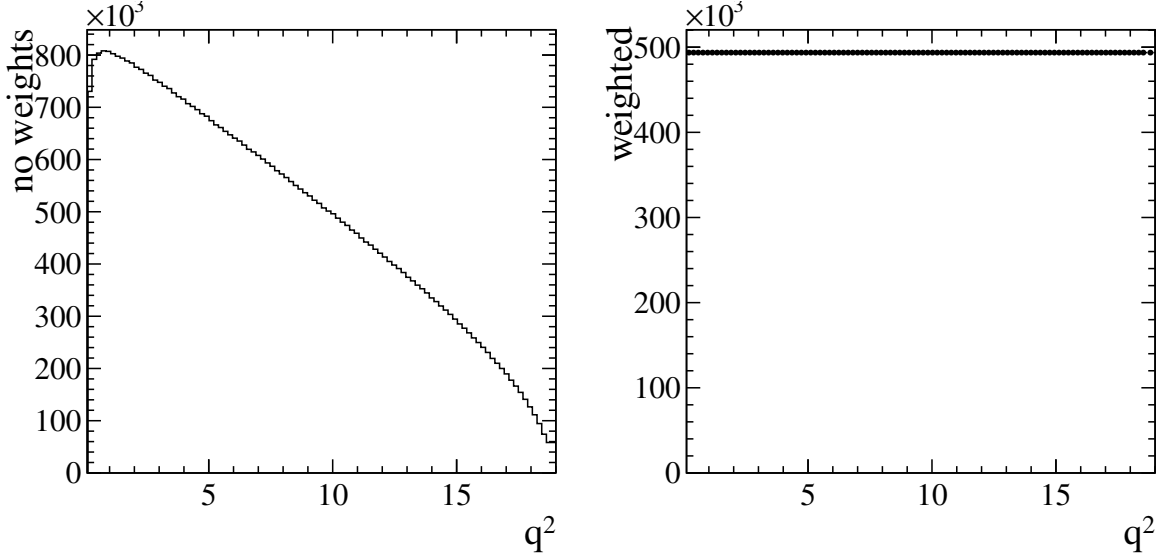


Figure 7.1: The  $q^2$  distribution for the phasespace MC at generator-level is shown before (left) and after (right) applying weights to force the  $q^2$  distribution to be flat.

The sample used to determine the acceptance for the combined 2011 and 2012 data sample, is simulated using only 2012 data taking conditions. This is possible as the reconstruction and trigger conditions are similar enough to allow the 2012 simulation to be used for both years. In contrast, the data taken in Run 2 of the LHC need a separate parameterisation of the acceptance effects, as the data taking conditions and the performance of the PID variables are significantly different from the conditions in Run 1. The Run 2 data set is further subdivided into data taken in 2016 (Run 2p1) and data taken in 2017 and 2018 (Run 2p2). This is motivated by different trigger thresholds between the two subsets and a bug being present in the trigger configuration causing the `DiMuonDetached` and `DiMuonDetachedHeavy` lines to be virtually identical and subsequently introducing a step like behaviour in both data and simulated events in the 2016 sample. To determine the acceptance for 2016 data, the simulated



sample use the running conditions from 2016, while for 2017 and 2018 data, a simulated sample with 2017 data taking conditions is used. Only using a simulated sample with 2017 data taking conditions is again possible due to the near-identical data taking conditions in 2017 and 2018. The corrections described in Sec. 5.5 are applied to each of the simulated samples used to calculate the acceptance. These corrections are the reweighting of the track multiplicity, the  $B_s^0 p_T$  spectrum and the L0 response using the TISTOS method, as well as the PID resampling using the MEERKAT tool.

The one-dimensional projections of the angular acceptance in the decay angles and  $q^2$  are shown in Figs. 7.2-7.4 for Run 1, Run 2p1 and Run 2p2, respectively. The distribution of events as found in simulation are indicated as crosses while the solid line indicates the parameterisation. The  $q^2$  distribution shown in Fig. 7.3 exhibits a step-like feature around  $9.5 \text{ GeV}^2/c^4$  for the Run 2p1 simulation caused by the omission of the `DiMuonDetached` trigger in the 2016 selection. The omission originates from the aforementioned bug incorrectly setting the threshold in  $q^2$  such that the `DiMuonDetached` line has almost 100% overlap with the `DiMuonDetachedHeavy` trigger. As a consequence of this incorrect threshold, the low  $q^2$  region is not as populated as it should be, which causes this step-like feature. However, this step can be described by the acceptance parameterisation using the chosen orders in  $q^2$ . The distributions of the three angles are well described for each year, which confirms the use of only even orders to be sufficient.

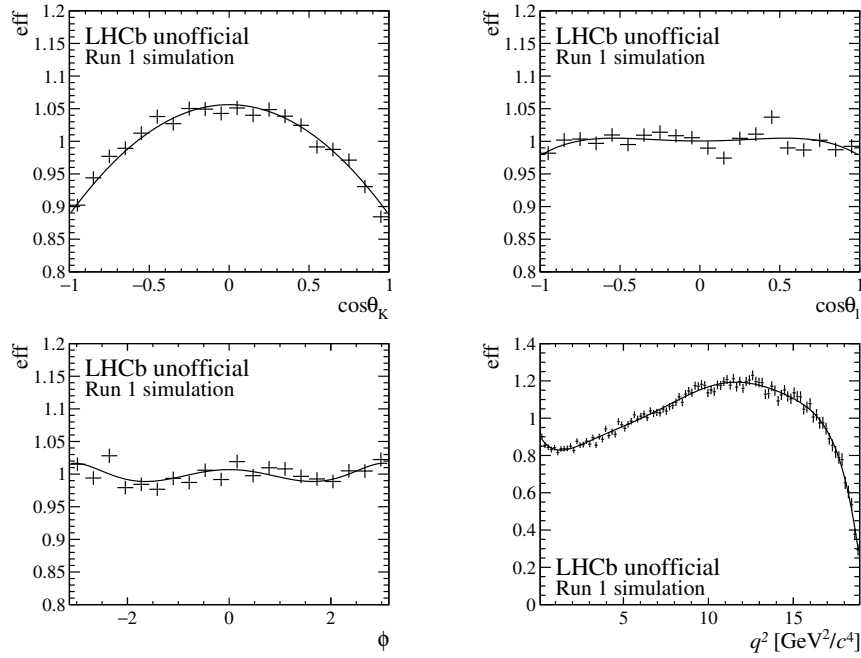


Figure 7.2: One-dimensional projections of the angular acceptance as found in simulated events overlaid with the parameterisation for Run 1 simulation. From left to right and top to bottom, the distributions in  $\cos\theta_K$ ,  $\cos\theta_l$ ,  $\phi$  and  $q^2$  are shown.

In principle, the acceptance also affects the decay-time distribution ( $t$ ) of the  $B_s^0$  meson, however, as this analysis is performed time-integrated, this  $t$ -dependency is not taken into account by default. A systematic uncertainty for this choice is assigned in Sec. 8.4. Additionally, a systematic uncertainty is assigned to the choice of using a simulated sample generated with  $\Delta\Gamma_s = 0$ , as discussed in Sec. 8.5.

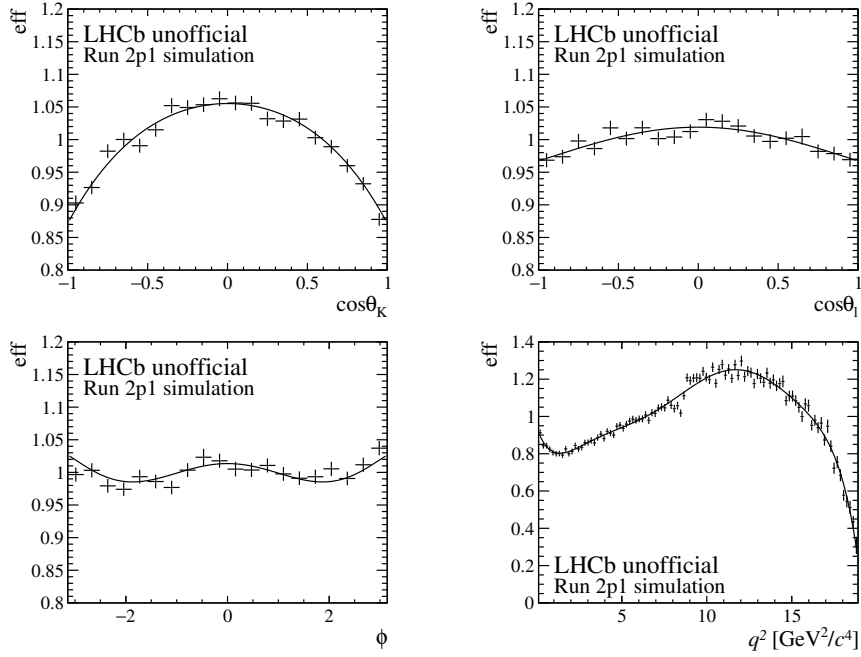


Figure 7.3: One-dimensional projections of the angular acceptance as found in simulated events overlaid with the parameterisation for Run 2p1 simulation. From left to right and top to bottom, the distributions in  $\cos\theta_K$ ,  $\cos\theta_l$ ,  $\phi$  and  $q^2$  are shown. The step-like feature in the  $q^2$ -distribution around  $9.5$  GeV<sup>2</sup>/c<sup>4</sup> is caused by the DiMuonDetached and discussed in the text.

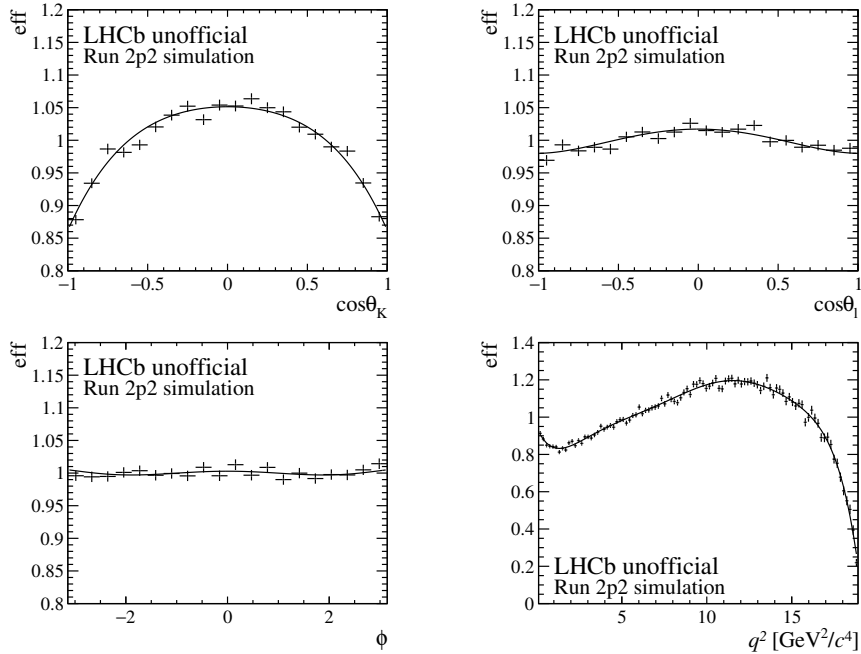


Figure 7.4: One-dimensional projections of the angular acceptance as found in simulated events overlaid with the parameterisation for Run 2p2 simulation. From left to right and top to bottom, the distributions in  $\cos\theta_K$ ,  $\cos\theta_l$ ,  $\phi$  and  $q^2$  are shown.

For the corrections using weights, namely to correct the **nTracks** and the  $B_s^0 p_T$  distributions, and the **L0 TISTOS** efficiency shape a systematic uncertainty is assigned, as outlined in Sec. 8.6. For the correction of the **PID** response using the **MEERKAT** tool a separate systematic uncertainty is assigned, which is the topic of Sec. 8.7. In addition, a systematic uncertainty estimating the impact of the phasespace MC sample having limited statistics is assigned. This systematic uncertainty is discussed in Sec. 8.8.

## 7.2 Fit model

An unbinned maximum likelihood fit to the three decay angles ( $\vec{\Omega} = (\cos \theta_l, \cos \theta_K, \phi)$ ) and the reconstructed mass of the  $B_s^0$  candidate ( $m(K^+ K^- \mu^+ \mu^-)$ ) is performed in regions of  $q^2$  to extract the angular observables  $F_L$ ,  $S_3$ ,  $S_4$ ,  $A_5$ ,  $A_{\text{FB}}^{CP}$ ,  $S_7$ ,  $A_8$  and  $A_9$ , which were introduced in Sec. 2.3.4. The data set used for this fit does not exclusively consist of signal events, but also contains background events originating from combinatorial background. Other sources of backgrounds such as single and double misidentification backgrounds contribute less than 0.5% in any data set and will not be modeled as discussed in Sec. 6. The likelihood function ( $\mathcal{L}$ ) is then given by

$$\begin{aligned} -\ln \mathcal{L}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) &= - \sum_{\text{events}} \ln \left[ \mathcal{P}_{\text{tot}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) \right] \\ &= - \sum_{\text{events}} \ln \left[ f_{\text{sig}} \mathcal{P}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) \right. \\ &\quad \left. + (1 - f_{\text{sig}}) \mathcal{P}_{\text{bkg}}(m, \vec{\Omega} | \vec{\lambda}_{\text{nuisance}}) \right], \end{aligned}$$

where  $\vec{\lambda}_{\text{phys}}$  and  $\vec{\lambda}_{\text{nuisance}}$  denote the physics and nuisance parameters,  $f_{\text{sig}}$  is the fraction of signal events,  $\mathcal{P}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}})$  is the signal PDF and  $\mathcal{P}_{\text{bkg}}(m, \vec{\Omega} | \vec{\lambda}_{\text{nuisance}})$  the background PDF. The mass and angular components of the PDF are assumed to factorise in both signal and background, such that

$$\mathcal{P}_{\text{sig (bkg)}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) = \mathcal{P}_{\text{sig (bkg)}}(m | \vec{\lambda}_{\text{nuisance}}) \cdot \mathcal{P}_{\text{sig (bkg)}}(\vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}).$$

Furthermore, the maximum likelihood fit is performed simultaneously to the data taken in Run 1, Run 2p1 and Run 2p2 by combining the logarithmic likelihoods according to

$$\begin{aligned} -\ln \mathcal{L}_{\text{tot}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) &= - \sum_{\text{Run 1 events}} \ln \left[ \mathcal{P}_{\text{tot}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}^{\text{Run 1}}) \right] \\ &\quad - \sum_{\text{Run 2p1 events}} \ln \left[ \mathcal{P}_{\text{tot}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}^{\text{Run 2p1}}) \right] \\ &\quad - \sum_{\text{Run 2p2 events}} \ln \left[ \mathcal{P}_{\text{tot}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}^{\text{Run 2p2}}) \right]. \end{aligned}$$

The angular component of the signal PDF ( $\mathcal{P}_{\text{sig}}(\vec{\Omega})$ ), is given by Eq. 2.8, taking the angular acceptance additionally into account. This is done either by performing a weighted fit, where each event is weighted by  $1/\epsilon(\vec{\Omega}, q^2)$  or directly as an additional term  $\epsilon(\vec{\Omega}, q^2)$  in the PDF. Details on the implementation of the angular acceptance will be given in Sec. 7.2.2. The

mass component of the signal PDF ( $\mathcal{P}_{\text{sig}}(m)$ ), is given by a sum of two Crystal Ball functions (CB) [224], each of which is Gaussian function with a power-law tail. The parameters of the CBs are the mean and width ( $\sigma_{1,2}$ ) of the Gaussian and the tail parameters ( $n_{1,2}$  and  $\alpha_{1,2}$ ). Furthermore, a fraction (frac) relating the first and second CB is used as a parameter in for the mass component of the signal PDF. The power-law tails of the two CBs are chosen to point to opposite sides of the mean of the Gaussian.

The background PDF,  $\mathcal{P}_{\text{bkg}}(m, \vec{\Omega})$ , is given by an exponential function for the mass component, where the slope of the exponential function ( $\alpha_m$ ) is the only parameter. A product of first order Chebychev polynomials ( $P_1^{Ch}$ ) is used for the angular component. The background PDF is, therefore, given as

$$\begin{aligned} \mathcal{P}_{\text{bkg}}(m, \cos \theta_l, \cos \theta_K, \phi) &= \mathcal{P}_{\text{bkg}}(m) \cdot \mathcal{P}_{\text{bkg}}(\cos \theta_l) \cdot \mathcal{P}_{\text{bkg}}(\cos \theta_K) \cdot \mathcal{P}_{\text{bkg}}(\phi) \\ &= \frac{1}{\mathcal{N}} \cdot \exp(-\alpha_m \cdot m) \cdot [1 + c_1^{\cos \theta_l} P_1^{Ch}(\cos \theta_l)] \\ &\quad \cdot [1 + c_1^{\cos \theta_K} P_1^{Ch}(\cos \theta_K)] \\ &\quad \cdot [1 + c_1^\phi P_1^{Ch}(\phi)], \end{aligned}$$

where  $\mathcal{N}$  is the normalisation of the background PDF, the  $c_1^j$  are the coefficients for the Chebychev polynomials of the decay angle  $j$  and are free parameters in the fit.

Several systematic uncertainties are assigned to specific choices made in the fit model: The size of the effect induced by omitting peaking backgrounds in the fit model is covered by a systematic uncertainty, as described in Sec. 8.2. In Sec. 8.9, a systematic uncertainty is assigned to the choice of a double CB as the signal mass model. A further systematic uncertainty is assigned to the use of a product of first order Chebychev polynomials instead of using higher orders, as discussed in Sec. 8.10.

### 7.2.1 Fit model parameter determination

The fit model for the simultaneous fit consists of 17 free parameters for the three data sets in the rare mode. Eight out of these 17 free parameters are the angular observables, which are shared between the data sets, and three parameters are used to describe the signal fraction in each of the data sets. The remaining six free parameters are connected to the description of the combinatorial background, with three parameters encoding the slope of the exponential describing the mass component in each data set and the other three free parameters correspond to the coefficients of the Chebychev polynomials. The coefficients of the Chebychev polynomials are shared across the data sets to stabilise the fit. This stabilisation of the fit is especially important in the  $q^2$  region  $q^2 \in [0.1, 0.98] \text{ GeV}^2/c^4$ , in which only very few background events are available.

The parameters of the signal mass model are fixed from simulation or control mode data. The mean and the width of the CBs in each data set are taken from control mode data and the tail parameters as well as the ratio between the CBs are taken from simulated  $B_s^0 \rightarrow J/\psi \phi$  events. To account for  $q^2$  dependent resolution effects in the rare mode, a scale factor ( $s$ ) is introduced in the rare mode for the widths. The scale factor is determined for each  $q^2$  region on simulated  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  events generated according to the BTOSLLBALL v6 model by fitting the reconstructed mass of the  $B_s^0$  candidate. The scale factors are shown in Fig. 7.5.

The ranges for the angular observables are set to the range  $[-1, 1]$ , except for  $F_L$  where the lower bound is set to the physical boundary of zero. All angular background parameters

are allowed to float in the fit in the range of  $[-2, 2]$ .

A systematic uncertainty is assigned for the effect sharing the angular background parameters negligible has on the angular observables. This is discussed in Sec. 8.11.

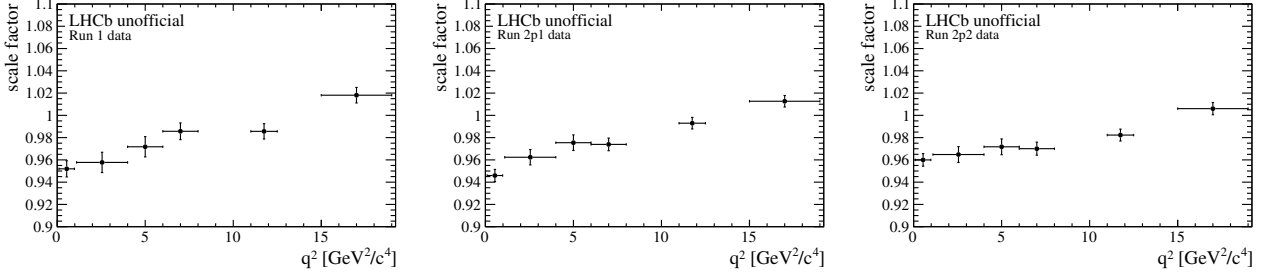


Figure 7.5: Factors  $s$  to scale the fit model widths from control to rare mode in order to account for the  $q^2$ -dependent mass resolution.

## 7.2.2 Implementation of the angular acceptance in the likelihood

The effects of selection, reconstruction and resolution distort the distributions of the decay angles and  $q^2$ . Therefore, it is crucial to account for the acceptance effect in fits to data or simulated events. The acceptance effect can be either included by performing a weighted fit, in which the events are weighted by  $1/\epsilon(\vec{\Omega}, q^2)$ , or by directly including the effect in the signal PDF.

In the first option, the per-event weight ( $w_e$ ) is included in the likelihood as follows

$$\begin{aligned} -\ln \mathcal{L} &= - \sum_{\text{event } e} w_e \times \ln \mathcal{P}_{\text{tot}}(\vec{\Omega}_e, m_e) \\ &= - \sum_{\text{event } e} \frac{1}{\epsilon(\vec{\Omega}_e, q_e^2)} \times \ln \mathcal{P}_{\text{tot}}(\vec{\Omega}_e, m_e). \end{aligned}$$

In this approach, all events will be weighted, meaning not only events originating from  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays but also those originating from background candidates. However, the background events are still well described with a first order polynomial even when correcting for the angular acceptance. The weighted fit approach is the preferred for the wide  $q^2$  regions,  $q^2 \in [1.1, 6] \text{ GeV}^2/c^4$  and  $q^2 \in [15.0, 18.9] \text{ GeV}^2/c^4$ , considering that this method can account for possible variation of the acceptance within a  $q^2$  region. Furthermore, the expected signal yield in these  $q^2$  regions is sufficiently large to reduce possible fit stability issues arising from the use of a weighted fit. Special care needs to be taken when estimating the parameter uncertainties since weighted fits in general need special methods to give correct coverage [190]. Obtaining uncertainty estimates via bootstrapping has been shown to generally be a reliable method of guaranteeing good coverage in weighted fits [190]. As such, the bootstrapping approach described in Sec. 4.2 is used in this analysis. The coverage is validated on pseudoexperiments as outlined in Sec. 7.3.3.

The second option to include the acceptance effects is the inclusion of the efficiency in the signal PDF. The signal PDF is then given by

$$-\ln \mathcal{L}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) = - \sum_{\text{events}} \ln \frac{\epsilon(\vec{\Omega}, q^2) \mathcal{P}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}})}{\int \epsilon(\vec{\Omega}, q^2) \mathcal{P}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) d\vec{\Omega} dq^2}$$

$$\begin{aligned}
&= - \sum_{\text{events}} \ln \frac{\epsilon(\vec{\Omega}, q^2) \mathcal{P}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}})}{\mathcal{N}_{\text{sig}}} \\
&= - \sum_{\text{events}} \left[ \ln \left( \epsilon(\vec{\Omega}, q^2) \right) + \ln \left( \mathcal{P}_{\text{sig}}(m, \vec{\Omega} | \vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) \right) \right] \\
&\quad + N \mathcal{N}_{\text{sig}},
\end{aligned}$$

where  $N$  is the number of events and  $\mathcal{N}_{\text{sig}}$  is the normalisation of the signal PDF. The term  $\ln \left( \epsilon(\vec{\Omega}, q^2) \right)$  is a constant for a given data set and can, therefore, be neglected in the minimisation without changing the physics parameters at the minimum. In contrast, the normalisation of the signal PDF needs to be taken into account in the minimisation. It can be written as

$$\begin{aligned}
\mathcal{N}_{\text{sig}} &= \iint \epsilon(\vec{\Omega}, q^2) \mathcal{P}_{\text{sig}}(\vec{\Omega}) d\vec{\Omega} \\
&= \iint \epsilon(\vec{\Omega}, q^2) \frac{9}{32\pi} \left( \sum_i S_i f_i(\vec{\Omega}) + \sum_j A_j f_j(\vec{\Omega}) \right) d\vec{\Omega} dq^2 \\
&= \frac{9}{32\pi} \left( \sum_i \int S_i \xi_i(q^2) dq^2 + \sum_j \int A_j \xi_j(q^2) dq^2 \right),
\end{aligned}$$

where the index  $i$  is taken for the  $CP$ -averages  $S_i$  from [1c, 3, 4, 7], the index  $j$  is taken for the  $CP$ -asymmetries from [5, 6s, 8, 9]. The angular terms  $f_{i(j)}(\vec{\Omega})$  are defined by Eq. 2.6 and  $\xi_{i(j)}$  encodes the integral of the efficiencies and projected onto the angular terms and is defined as  $\xi_{i(j)} = \int \epsilon(\vec{\Omega}, q^2) f_{i(j)}(\vec{\Omega}) d\vec{\Omega}$ . This approach of including the angular acceptance in the signal PDF is the preferred approach for the narrow  $q^2$  regions, as it is more stable than a weighted fit for low event yields. As the acceptance is not expected to vary significantly over the narrow  $q^2$  region, the acceptance is evaluated at a single point in  $q^2$ . This evaluation point is the median value for the  $q^2$  distribution in a given  $q^2$  region. The median value is determined on simulated events of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay generated according to the BTOSLLBALL v6 model. A systematic uncertainty for this choice is assigned in Sec. 8.12. For the narrow  $q^2$  regions, the statistical uncertainties are determined using the Hessian matrix, as implemented in the ROOT framework in the HESSE function. Before running HESSE, the limited parameter range imposed during the fitting procedure is removed. This results in good uncertainty coverage, as discussed in Sec. 7.3.3.

## 7.3 Fit validation

When performing an angular analysis, it is crucial to validate the fit behaviour before performing a fit to the rare mode data. In this analysis three avenues are used to perform this validation. The validation is performed with simulated events, as described in Sec. 7.3.1, using control mode events in data, as discussed in Sec. 7.3.2, and with pseudoexperiments, as introduced in Sec. 7.3.3. These three options have different advantages and can be used to probe different effects. Simulated events can be used to validate the acceptance, by comparing the extracted observables corrected for acceptance effects to the observables used in generation. Each of the three acceptance parameterisations is validated independently of each other with this method. As the values for the  $B_s^0 \rightarrow J/\psi \phi$  control mode have been measured before, the values extracted

from fits to control mode data can also be used to validate the fit behaviour and the angular acceptance. The control mode data can further be used to validate the simultaneous fit, by validating that the angular observables extracted from fits to the individual data sets agree and that the results of the individual data sets agree with those of the simultaneous fit. The third option, namely the use of pseudoexperiments, allows to test the coverage and fit bias for the expected event yields in the rare mode fits.

### 7.3.1 Fit validation using simulation

The parameterisation of the acceptance is validated by comparing the angular observables extracted from fits to generator-level simulation to those from fits to fully reconstructed and selected simulation. If the acceptance correctly describes all reconstruction, selection and resolution effects, the angular observables in the fit to the fully reconstructed and selected simulation will be compatible with those extracted from the generator-level sample within the statistical uncertainty.

The values for the observables on generator-level are extracted by performing a fit to a privately generated, high yield sample. This sample contains 50 000 000 events that have been generated according to the **BTOSLLBALL v6** model [212,213]. Considering the size of the simulated sample, a fast simulation technique known as particle gun [237,238] is used, where only the signal decay is generated and simulated instead of the full proton-proton interaction. For this sample, **PHOTOS** [206] is not used in the generation, as the acceptance includes corrections for the emissions of photons. In the fit to this sample, the angular acceptance is not taken into account considering the generator-level simulation does not suffer from detector effects.

To extract the observables from fits to fully reconstructed and selected simulation, samples centrally generated by the LHCb collaboration according to the **BTOSLLBALL v6** model are used. These fits are performed separately under Run 1, Run 2p1 and Run 2p2 data taking conditions. The angular acceptance is implemented as described in Sec. 7.2.2. A visual comparison of the angular observables extracted from fully reconstructed and selected simulation and the comparison with generator-level values can be found in Fig. 7.6. The shown uncertainties are only of statistical nature. The values extracted from the fits to reconstructed simulation typically agree with the values from the fit to generator-level simulation at the level of 0.01 or lower. The most significant deviation found is 0.023 for 2016 simulation in  $F_L$  in the  $q^2$  region  $[0.1, 0.98] \text{ GeV}^2/c^4$ . The significance of this deviation depends on what is considered the appropriate uncertainty. One can consider just the statistical uncertainty from the fit, the statistical uncertainty from the fit plus relevant systematic uncertainties (*i.e.* involving the acceptance) or the expected sensitivity from data. These options correspond to the significances of  $4.4 \sigma_{\text{MC stats}}$ ,  $2.6 \sigma_{\text{MC statistics+relevant systematic}}$  or  $0.5 \sigma_{\text{expected sensitivity}}$ , respectively. The second of these options is the most reliable as the relevant effects are considered and should, therefore, be used to evaluate the deviation. Taking the number of observables and number of  $q^2$  regions used in this validation into account, the deviation of 0.023 with a significance of  $2.6 \sigma_{\text{MC statistics+relevant systematic}}$  is considered to be acceptable. As a consequence, the overall agreement is considered to be good.

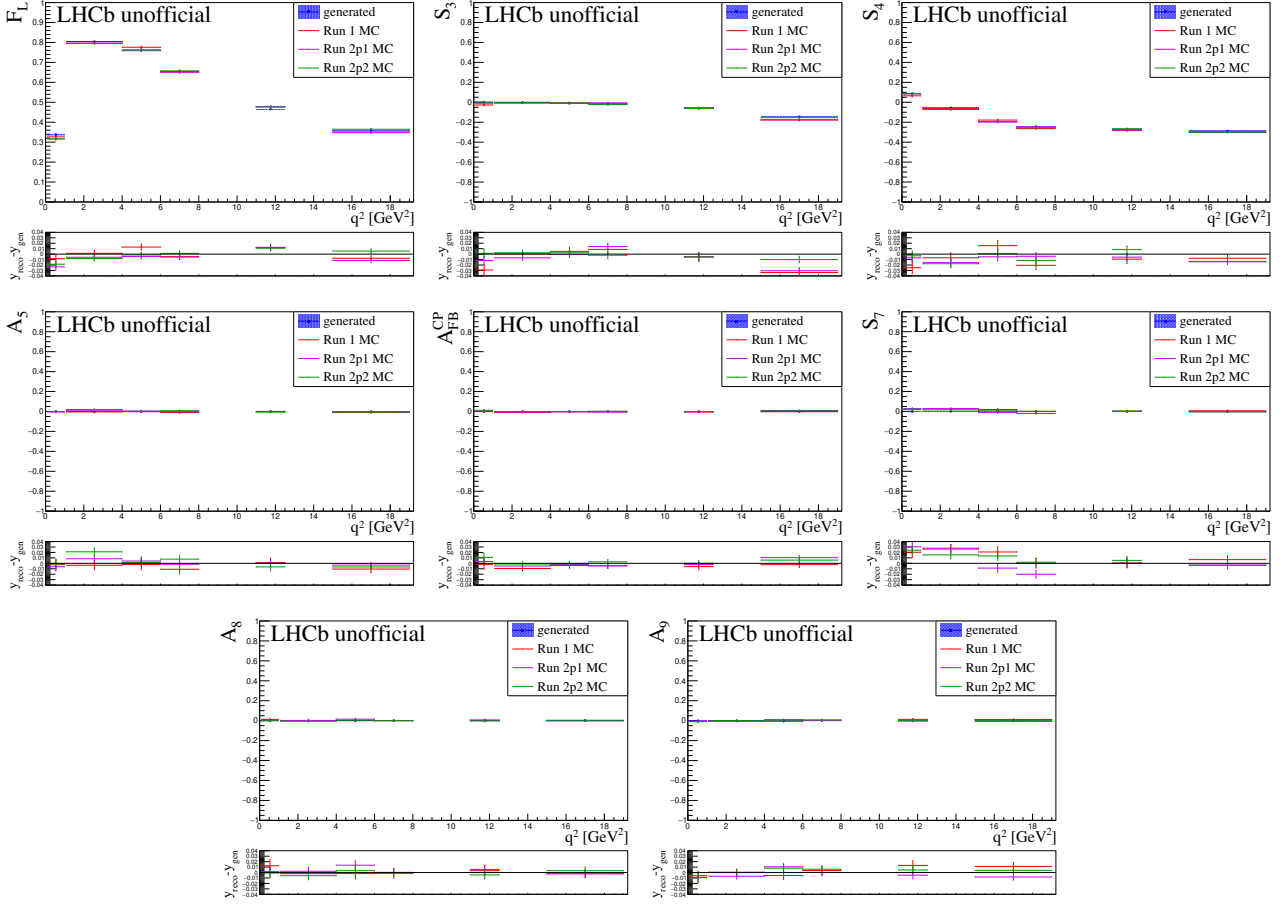


Figure 7.6: The extracted values for the angular observables are shown as a function of  $q^2$ . From left to right (first row)  $F_L$ ,  $S_3$  and  $S_4$ , (second row)  $A_5$ ,  $A_{FB}^{cp}$  and  $S_7$ , and (third row)  $A_8$  and  $A_9$  are shown. In red, violet and green observables extracted from the fully reconstructed simulation generated under Run 1, Run 2p1 and Run 2p2 data taking conditions are shown, respectively. The blue hatched regions indicate the  $1\sigma$  region of the values from the generator-level simulation.



### 7.3.2 Fit validation using control mode data

Several aspects of the fit can be validated with a fit to control mode  $B_s^0 \rightarrow J/\psi \phi$  data. These aspects are the general ability of the PDF to describe the data, the compatibility of the three control mode data sets with each other, the correction for the angular acceptance and the behaviour of the simultaneous fit. These validation steps include fitting each individual data set and comparing the extracted values to each other and to previously published results. The results of these validation steps will be presented in the following.

#### 7.3.2.1 Results from control model fits

In order to validate the full simultaneous angular fit, the result from the simultaneous fit to control mode  $B_s^0 \rightarrow J/\psi \phi$  data across all data-taking periods is compared with the results to individual fits of each data-taking period. These fits use the fit model described in Sec. 7.2. In Figs. 7.7-7.9 the projections of the fits to the independent data sets onto the mass and the three decay angles are shown for the Run 1, Run 2p1 and Run 2p2 data sets, respectively. The pulls comparing the data distribution and PDF projections for all data sets statistically fluctuate around 0 and are, therefore, considered to be well described.

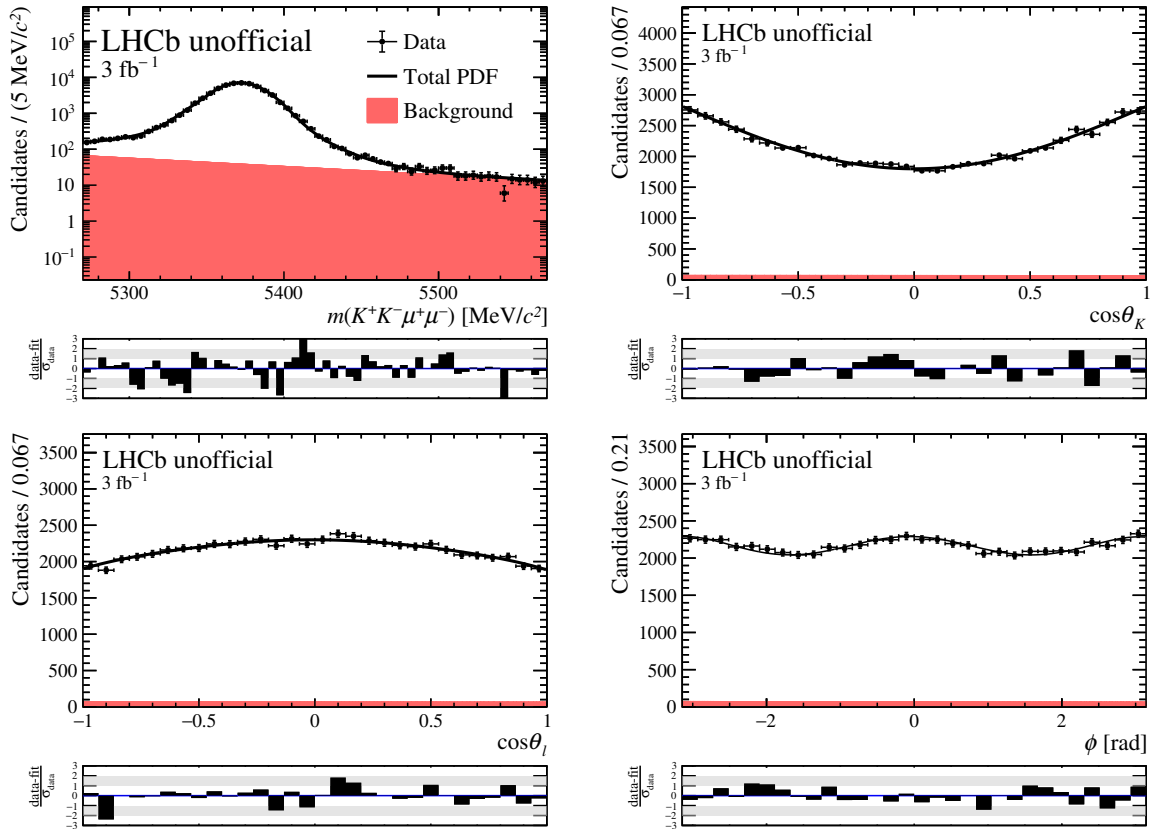


Figure 7.7: The projections of the angular fit on  $B_s^0 \rightarrow J/\psi \phi$  control mode data using Run 1 data are shown. From top left to bottom right, the distribution of the mass (logarithmically),  $\cos \theta_K$ ,  $\cos \theta_l$  and  $\phi$  are shown. Each pull fluctuates around 0 and does not show any sign of obvious misbehaviour of the fit.

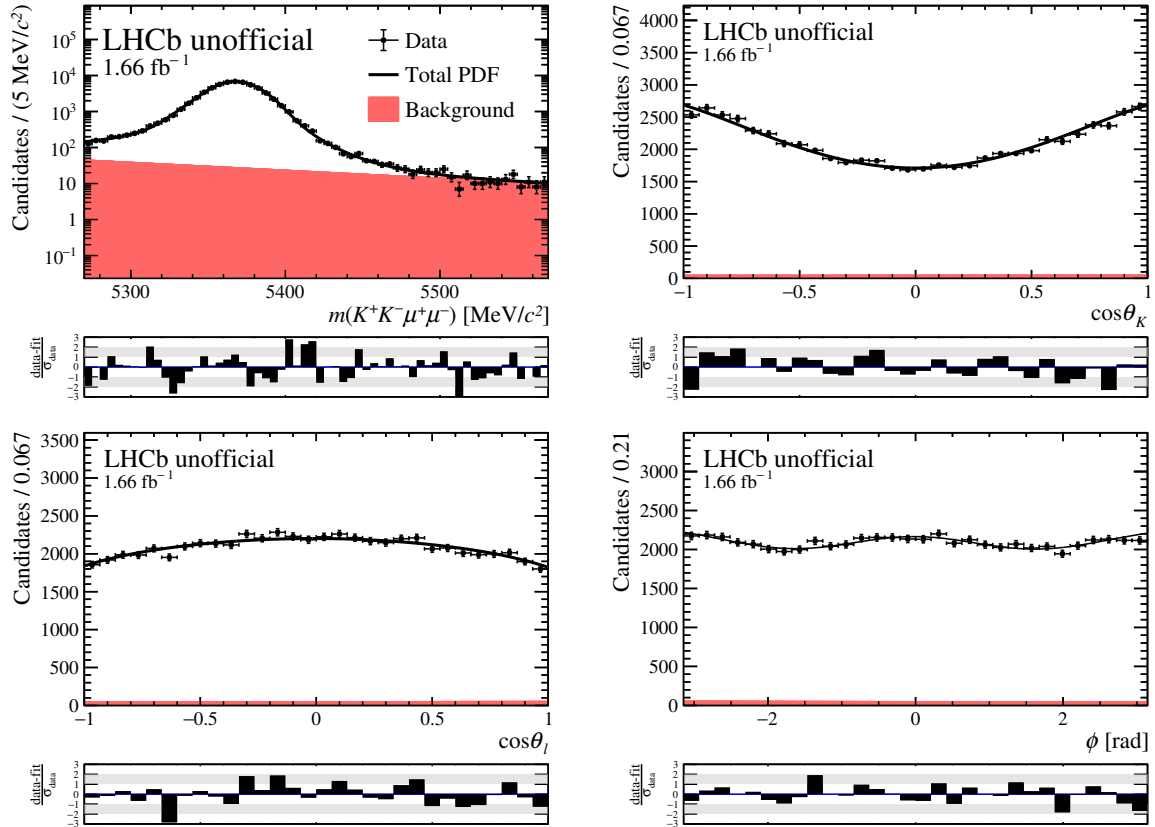


Figure 7.8: The projections of the angular fit on  $B_s^0 \rightarrow J/\psi \phi$  control mode data using Run 2p1 data are shown. From top left to bottom right, the distribution of the mass (logarithmically),  $\cos \theta_K$ ,  $\cos \theta_l$  and  $\phi$  are shown. Each pull fluctuates around 0 and does not show any sign of obvious misbehaviour of the fit.

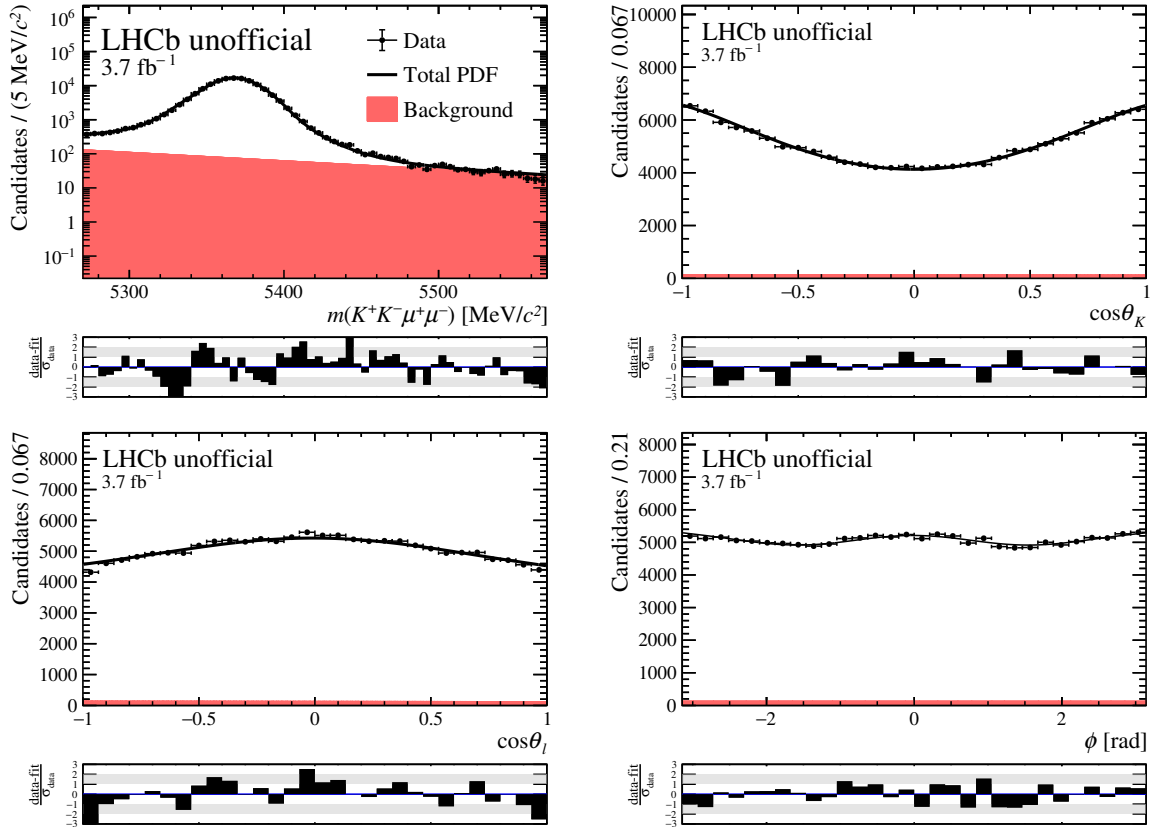


Figure 7.9: The projections of the angular fit on  $B_s^0 \rightarrow J/\psi \phi$  control mode data using Run 2p2 data are shown. From top left to bottom right, the distribution of the mass (logarithmically),  $\cos \theta_K$ ,  $\cos \theta_l$  and  $\phi$  are shown. Each pull fluctuates around 0 and does not show any sign of obvious misbehaviour of the fit.

In Fig. 7.10 the differences for the obtained angular observables between the simultaneous fit of  $B_s^0 \rightarrow J/\psi \phi$  control mode data and the fits to the individual data sets are given. The uncertainties for the angular observables are only of statistical nature and expected to be systematically dominated. For the fits to the individual data sets, a mild tension in  $S_3$  of  $2.3 \sigma_{\text{stat.}}$  can be seen between the fits to Run 1 and Run 2 data, however, this is unremarkable given the number of observables being compared. Furthermore, the absolute size of this deviation (0.013) is negligible compared to the expected precision on the rare mode, which depends on the  $q^2$  region and is in the range of  $[0.042, 0.089]$  for  $S_3$ . The three data sets are, therefore, considered to be in good agreement with each other.

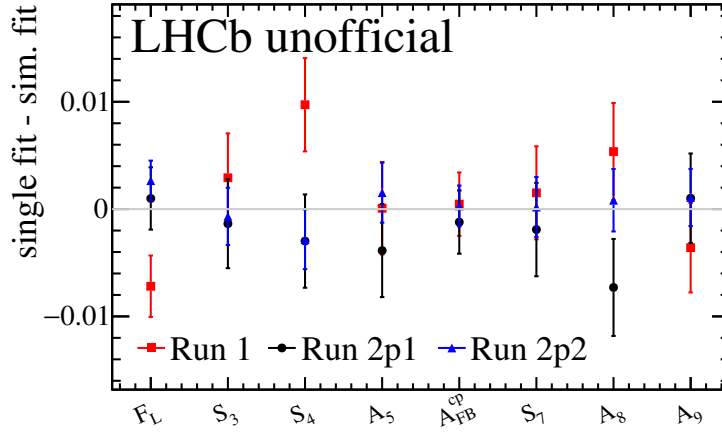


Figure 7.10: The difference in the angular observables between the simultaneous fit of  $B_s^0 \rightarrow J/\psi \phi$  control mode data and the fits to the individual data sets are given in red, black and blue for Run 1, Run 2p1 and Run 2p2 data, respectively.

### 7.3.2.2 Comparison with dedicated $B_s^0 \rightarrow J/\psi \phi$ analyses

Comparing the measured observables to the ones from a dedicated  $B_s^0 \rightarrow J/\psi \phi$  analysis performed at LHCb [239, 240], is not straightforward as dedicated  $B_s^0 \rightarrow J/\psi \phi$  angular analyses typically measure transversity amplitudes as opposed to the angular observables measured in this analysis. However, the angular observables and the transversity amplitudes are linked to another as the angular coefficients ( $I_i$ ) depend on the transversity amplitudes as shown in Eq. 2.7. Furthermore, the comparison is complicated due to the fit strategy employed in this analysis differing from the strategy used in the dedicated  $B_s^0 \rightarrow J/\psi \phi$  analyses. The latter typically are time-dependent, flavour-tagged analyses that take the spin-0 component of the  $K^+ K^-$  system ( $S$ -wave) explicitly into account. The differences in fit strategy arise from the branching fractions of the two decay modes being vastly different. The branching fraction of the  $B_s^0 \rightarrow J/\psi \phi$  mode is roughly a factor 100 larger than that of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  mode over the complete  $q^2$  range. As a consequence the yields are roughly 100 times as large in the  $B_s^0 \rightarrow J/\psi \phi$  mode and the dedicated  $B_s^0 \rightarrow J/\psi \phi$  analyses are able to be performed time-dependent and flavour-tagged, but are also required to consider the approximately 2% of  $S$ -wave events. Therefore, the measured angular observables illustrated in Fig. 7.10 can only be meaningfully compared to the transversity amplitudes measured in the dedicated analyses [239, 240], which are performed

at LHCb using Run 1 and Run 2p1 data, by converting the angular observables to transversity amplitudes while the differences in the fit strategies are corrected for.

The conversion of the  $CP$ -averaged observables to the transversity amplitudes is done using Eq. 2.7. Following these equations, the observables  $F_L$  and  $S_3$  are given as

$$F_L(t) = \frac{|A_0(t)|^2 + |\bar{A}_0(t)|^2}{\mathcal{N}(t)}$$

$$\text{and } S_3(t) = \frac{1}{2} \frac{|A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2 - |A_\parallel(t)|^2 - |\bar{A}_\parallel(t)|^2}{\mathcal{N}(t)},$$

where  $\mathcal{N}(t)$  is a shorthand for the normalisation, defined as

$$\mathcal{N}(t) = \sum_{i=0,\perp,\parallel} \left( |A_i(t)|^2 + |\bar{A}_i(t)|^2 \right).$$

As the amplitudes are time-dependent, the angular observables exhibit a dependence on the decay time as well. Following Ref. [241], the time dependence of the amplitudes  $A_i(t)$  has the form

$$|A_i(t)|^2 = |A_i(0)|^2 e^{-\Gamma_s t} \left( \cosh \left( \frac{\Delta\Gamma_s}{2} t \right) - \xi_i \cos(\phi_s) \sinh \left( \frac{\Delta\Gamma_s}{2} t \right) + \xi_i \sin(\phi_s) \sin(\Delta m_s t) \right),$$

where the index  $i$  indicates the polarisation and is taken from  $[0, \perp, \parallel]$ , the factor  $\xi_i$  is  $+1$  for the polarisations  $0$  and  $\parallel$  or  $-1$  for the polarisation  $\perp$ ,  $\Delta\Gamma_s$  denotes the decay width difference of the mass eigenstates of the two  $B_s^0$  meson and  $\phi_s$  denotes the  $CP$ -violating phase in the  $B_s^0$  system. The barred amplitudes  $\bar{A}_i(t)$  have the same general dependency but the sign of the term  $\xi_i \sin(\phi_s) \sin(\Delta m_s t)$  is flipped. Therefore, this last term vanishes when summing the amplitude and the barred amplitude.

In the dedicated amplitude analyses, the values of the amplitudes are taken at  $t = 0$ . In this case, the amplitudes can be expressed as

$$(7.1) \quad |A_0(0)|^2 + |\bar{A}_0(0)|^2 = F_L(0)$$

$$(7.2) \quad \text{and } |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 = \left( S_3(0) - \frac{1 - F_L(0)}{2} \right).$$

This expression even holds true for all  $t$  if the decay width difference is negligible, *i.e.*  $\Delta\Gamma_s = 0$ , or if the efficiency is nearly independent of  $t$ . For  $\Delta\Gamma_s \neq 0$  and a significant time-dependence of the efficiency, the time-integrated angular observables measured in data include the integral over the time-dependent selection and reconstruction efficiency,  $\epsilon(t)$ . The angular observables  $F_L$  and  $S_3$  as measured from data are given by

$$F_L = \frac{\int_0^\infty \epsilon(t) \left( |A_0(t)|^2 + |\bar{A}_0(t)|^2 \right) dt}{\int_0^\infty \epsilon(t) \left( |A_0(t)|^2 + |\bar{A}_0(t)|^2 + |A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2 + |A_\parallel(t)|^2 + |\bar{A}_\parallel(t)|^2 \right) dt}$$

$$\text{and } S_3 = \frac{\int_0^\infty \epsilon(t) \left( |A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2 - |A_\parallel(t)|^2 - |\bar{A}_\parallel(t)|^2 \right) dt}{2 \int_0^\infty \epsilon(t) \left( |A_0(t)|^2 + |\bar{A}_0(t)|^2 + |A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2 + |A_\parallel(t)|^2 + |\bar{A}_\parallel(t)|^2 \right) dt}$$

These two equations can now be solved for the transversity amplitudes. The detailed calculation is given in App. C. The final equations for  $|A_0(0)|^2 + |\bar{A}_0(0)|^2$  and  $|A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2$  are given by

$$(7.3) \quad |A_0(0)|^2 + |\bar{A}_0(0)|^2 = F_L \left( 1 + \frac{(2S_3 - F_L + 1)}{F_+(1 - 2S_3 + F_L) + F_-(1 + 2S_3 - F_L)} (F_+ - F_-) \right)$$

$$(7.4) \quad |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 = \frac{F_-(2S_3 - F_L + 1)}{F_+(1 - 2S_3 + F_L) + F_-(1 + 2S_3 - F_L)},$$

where  $F_\pm$  are terms absorbing the integral of the product of the time-dependent efficiency  $\epsilon(t)$  with the time-dependent components of the amplitudes. They are defined as

$$F_\pm = \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \left( \cosh \left( \frac{\Delta\Gamma_s}{2} t \right) \pm \cos(\phi_s) \sinh \left( \frac{\Delta\Gamma_s}{2} t \right) \right) dt.$$

The time-dependent efficiency is determined using fully reconstructed  $B_s^0 \rightarrow J/\psi \phi$  simulation and a privately produced generator-level  $B_s^0 \rightarrow J/\psi \phi$  simulation, generated with the fast simulation technique known as particle gun [237, 238].

Since the  $S$ -wave component is neglected in this analysis, but not in Refs. [239] and [240], the amplitudes are corrected for the effect of including the  $S$ -wave in the angular decay rate. There are two key effects to consider: The effect of the  $S$ -wave on the overall normalisation of the  $P$ -wave PDF, and the effect of the  $S$ -wave angular terms on the measured  $P$ -wave angular observables. The change in normalisation is an effect on the percent level and is trivial to correct for. To account for the new normalisation, an additional factor of  $\frac{1}{1 - F_S}$ , where  $F_S$  is the fraction of the  $S$ -wave, is included in the  $P$ -wave PDF. Thus each angular observable is rescaled by this factor. In contrast, the effect of including the full PDF for the  $S$ -wave in the fit is expected to be small and non-trivial to correct for. As such, this effect is neglected

The dedicated analyses measure the  $S$ -wave contamination in 6 bins in  $m_{KK}$ . To obtain the expected value of  $F_S$  for the single  $m_{KK}$  bin, the values from the internal LHCb notes for the dedicated analyses are averaged over accounting for the relative contributions. The resulting  $S$ -wave fractions are  $F_S^{\text{av, Run1}} = 0.024 \pm 0.005$  and  $F_S^{\text{av, Run2p1}} = 0.021 \pm 0.003$ .

The results of the fits to  $B_s^0 \rightarrow J/\psi \phi$  events in the three data sets are compared to the dedicated analyses [239, 240] in Tab. 7.1. The results that are corrected for the effects of the time-integration and the  $S$ -wave agree well with those from the dedicated analyses. The largest difference observed is a  $2.6 \sigma_{\text{stat}}$  shift between the  $|A_0|^2$  value found in Run 2p2 and the one found in the dedicated analysis of Run 2p1 data [240].<sup>1</sup> This corresponds to an absolute difference of 0.0114, which is substantially smaller than the expected statistical uncertainty in the rare mode and, therefore, considered to agree sufficiently well. The central values for Run 1, Run 2p1 and Run 2p2 agree at the level of  $1.5 \sigma_{\text{stat}}$  or better with each other.

The effect neglecting the  $S$ -wave component has on the rare mode is assigned as a systematic uncertainty, as discussed in Sec. 8.13. Similarly, effects due to the lifetime dependent efficiency are considered as a systematic uncertainty, as discussed in Sec. 8.4.

---

<sup>1</sup>The significance takes both statistical uncertainties into account and assumes them to be independent of another.

Table 7.1: Summary table for the comparison of the transversity amplitudes  $|A_0|^2$  and  $|A_\perp|^2$  determined from fits to  $B_s^0 \rightarrow J/\psi \phi$  data in each Run period are compared to those measured in Ref. [239] and Ref. [240]. The fit results are stated both with the corrections concerning the time-integration and the  $S$ -wave. The corrected fit results are given without uncertainties as no systematical uncertainties are estimated and the coverage of the statistical uncertainties has not been tested.

|               | PRL 114 (2015) 4    | Run 1<br>corrected | EPJC 80, 601 (2020) | Run 2p1<br>corrected | Run 2p2<br>corrected |
|---------------|---------------------|--------------------|---------------------|----------------------|----------------------|
| $ A_0 ^2$     | $0.5241 \pm 0.0075$ | 0.5235             | $0.5186 \pm 0.0037$ | 0.5283               | 0.5300               |
| $ A_\perp ^2$ | $0.2504 \pm 0.0061$ | 0.2515             | $0.2456 \pm 0.0044$ | 0.2459               | 0.2457               |

### 7.3.3 Fit validation using pseudoexperiments

To further validate the simultaneous fit of the rare signal mode, the bias and coverage of the angular observables extracted by the fit are determined using pseudoexperiments for each  $q^2$  region. For each  $q^2$  region 5000 pseudoexperiments are performed. A simultaneous fit to three separately generated pseudo data sets, representing the Run 1, Run 2p1 and Run 2p2 data sets, is performed for each pseudoexperiment. The event yields for each of the data set is taken from blind fits to the mass distribution of the corresponding data set. Furthermore, the values for the angular observables used for the generation are taken from blind fits to data (best-fit values). Due to potential statistical fluctuations in the fit parameters there is no guarantee that the signal PDF calculated using observables from data will be positive across the entire angular phasespace. Pseudoexperiments generated flat across the decay angles are used to determine whether the signal PDF in a given  $q^2$  region has negative PDF regions for the angular observables taken from data. For each of the pseudoexperiments, 100 000 pseudoevents are generated. Using these pseudoexperiments, the best-fit values for the angular observables are found to have a signal PDF which is positive across the entire angular phasespace for all  $q^2$  regions, with the exception of the  $q^2$  regions  $[4.0, 6.0] \text{ GeV}^2/c^4$  and  $[11.0, 12.5] \text{ GeV}^2/c^4$ . For these  $q^2$  regions, the fraction of pseudoevents with a negative signal PDF are 1.43 % and 0.22 %, respectively, and the fraction of data events with a negative signal PDF 0.3 % and 0 %, respectively. To avoid any bias from the negative PDF regions due to the generation of pseudoevents, in the two affected  $q^2$  regions, new sets of angular observables with a purely positive PDF are used to generate the pseudoexperiments instead. The new sets of angular observables are ensured to be the ones closest to the best-fit set with a purely positive signal PDF by minimising over the Euclidean distance  $d$ . The distance defines how close an alternative set of observables is and is given as

$$d = \sqrt{\sum_{\alpha \in \text{angular observables}} \left( \frac{\alpha_{\text{new}} - \alpha_{\text{orig.}}}{\sigma_\alpha} \right)^2},$$

where  $\sigma$  is the uncertainty on the angular observable as taken from HESSE. To find the minimum for  $d$ , 1 000 000 sets of angular observables are generated. The set with the lowest value for  $d$  for which none of the 100 000 generated pseudoevents in a pseudoexperiment, generated flat in the decay angles, has a negative signal PDF is chosen to generate the pseudoexperiments for the  $q^2$  regions  $[4.0, 6.0] \text{ GeV}^2/c^4$  and  $[11.0, 12.5] \text{ GeV}^2/c^4$ .

To determine the bias and coverage, the value and pull distributions of each angular observable are fit with Gaussian distributions, where the pull is defined as the difference of the fitted and generated values divided by the uncertainty of the fit ( $\frac{x_{\text{fit}} - x_{\text{gen.}}}{\sigma}$ ). The results from the fit validation pseudoexperiments are given in Tabs. 7.2-7.8. The proportion of fits that successfully converged, defined as **MIGRAD** having found a minimum with sufficient precision within 40 000 iterations (return code 0) and for **HESSE** to be positive definite (return code 3), for each  $q^2$  region is given in the captions of the tables.

The fits to the pseudoexperiments exhibit good behaviour, with most biases being less than 10% of the statistical uncertainty. The largest bias is found to be approximately 15% of the statistical uncertainty in the  $q^2$  region  $[4.0, 6.0] \text{ GeV}^2/c^4$ . The remaining bias in the value distribution is assigned as a systematic uncertainty, as discussed in Sec. 8.14. The pull widths in general indicate slight undercoverage but are mostly contained within the region of  $[0.9, 1.1]$ . In the final fit to data, the coverage is corrected by multiplying the uncertainty given by **HESSE** ( $\sigma_{\text{HESSE}}$ ) by the pull width found in the pseudoexperiments as follows

$$\sigma_{\text{value}} = \sigma_{\text{HESSE}} \times \text{pull width},$$

where  $\sigma_{\text{value}}$  represents the true error. All distributions of the pulls can be found in App. D.

Table 7.2: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 0.1 to 0.98  $\text{GeV}^2/c^4$  are shown. 4818/5000 pseudoexperiments converged successfully.

| variable             | bias               | pull mean          | pull width        |
|----------------------|--------------------|--------------------|-------------------|
| $F_L$                | $0.002 \pm 0.001$  | $-0.004 \pm 0.015$ | $1.048 \pm 0.011$ |
| $S_3$                | $0.002 \pm 0.001$  | $0.030 \pm 0.014$  | $1.005 \pm 0.010$ |
| $S_4$                | $0.004 \pm 0.001$  | $0.079 \pm 0.016$  | $1.084 \pm 0.011$ |
| $A_5$                | $-0.000 \pm 0.001$ | $-0.003 \pm 0.015$ | $1.037 \pm 0.011$ |
| $A_{\text{FB}}^{CP}$ | $0.009 \pm 0.001$  | $0.139 \pm 0.014$  | $1.001 \pm 0.010$ |
| $S_7$                | $0.000 \pm 0.001$  | $-0.017 \pm 0.015$ | $1.057 \pm 0.011$ |
| $A_8$                | $-0.004 \pm 0.001$ | $-0.055 \pm 0.015$ | $1.044 \pm 0.011$ |
| $A_9$                | $0.001 \pm 0.001$  | $0.015 \pm 0.015$  | $1.054 \pm 0.011$ |

Table 7.3: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 1.1 to 4.0  $\text{GeV}^2/c^4$  are shown. 5000/5000 pseudoexperiments converged successfully.

| variable             | bias               | pull mean          | pull width        |
|----------------------|--------------------|--------------------|-------------------|
| $F_L$                | $0.002 \pm 0.001$  | $0.045 \pm 0.014$  | $1.022 \pm 0.010$ |
| $S_3$                | $-0.000 \pm 0.001$ | $-0.006 \pm 0.015$ | $1.063 \pm 0.011$ |
| $S_4$                | $-0.000 \pm 0.001$ | $-0.017 \pm 0.015$ | $1.031 \pm 0.010$ |
| $A_5$                | $0.002 \pm 0.001$  | $0.030 \pm 0.014$  | $1.021 \pm 0.010$ |
| $A_{\text{FB}}^{CP}$ | $0.000 \pm 0.001$  | $0.015 \pm 0.015$  | $1.052 \pm 0.011$ |
| $S_7$                | $-0.001 \pm 0.001$ | $-0.015 \pm 0.015$ | $1.028 \pm 0.010$ |
| $A_8$                | $0.001 \pm 0.001$  | $0.017 \pm 0.015$  | $1.028 \pm 0.010$ |
| $A_9$                | $0.001 \pm 0.001$  | $0.016 \pm 0.015$  | $1.055 \pm 0.011$ |



Table 7.4: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 4.0 to 6.0 GeV<sup>2</sup>/c<sup>4</sup> are shown. 4999/4999 pseudoexperiments converged successfully.

| variable      | bias               | pull mean          | pull width        |
|---------------|--------------------|--------------------|-------------------|
| $F_L$         | $0.002 \pm 0.001$  | $0.056 \pm 0.015$  | $1.063 \pm 0.011$ |
| $S_3$         | $-0.010 \pm 0.001$ | $-0.152 \pm 0.016$ | $1.140 \pm 0.011$ |
| $S_4$         | $-0.009 \pm 0.001$ | $-0.127 \pm 0.016$ | $1.098 \pm 0.011$ |
| $A_5$         | $-0.003 \pm 0.001$ | $-0.046 \pm 0.015$ | $1.090 \pm 0.011$ |
| $A_{FB}^{CP}$ | $-0.002 \pm 0.001$ | $-0.036 \pm 0.015$ | $1.065 \pm 0.011$ |
| $S_7$         | $0.003 \pm 0.001$  | $0.044 \pm 0.015$  | $1.040 \pm 0.010$ |
| $A_8$         | $0.001 \pm 0.001$  | $0.015 \pm 0.015$  | $1.054 \pm 0.011$ |
| $A_9$         | $0.001 \pm 0.001$  | $0.021 \pm 0.016$  | $1.098 \pm 0.011$ |

Table 7.5: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 6.0 to 8.0 GeV<sup>2</sup>/c<sup>4</sup> are shown. 5000/5000 pseudoexperiments converged successfully.

| variable      | bias               | pull mean          | pull width        |
|---------------|--------------------|--------------------|-------------------|
| $F_L$         | $-0.000 \pm 0.001$ | $0.004 \pm 0.015$  | $1.028 \pm 0.010$ |
| $S_3$         | $-0.000 \pm 0.001$ | $-0.001 \pm 0.015$ | $1.055 \pm 0.011$ |
| $S_4$         | $-0.001 \pm 0.001$ | $-0.033 \pm 0.015$ | $1.040 \pm 0.010$ |
| $A_5$         | $0.001 \pm 0.001$  | $0.014 \pm 0.015$  | $1.034 \pm 0.010$ |
| $A_{FB}^{CP}$ | $0.001 \pm 0.001$  | $0.016 \pm 0.015$  | $1.046 \pm 0.010$ |
| $S_7$         | $0.002 \pm 0.001$  | $0.026 \pm 0.015$  | $1.031 \pm 0.010$ |
| $A_8$         | $-0.002 \pm 0.001$ | $-0.041 \pm 0.014$ | $1.021 \pm 0.010$ |
| $A_9$         | $0.001 \pm 0.001$  | $0.007 \pm 0.015$  | $1.060 \pm 0.011$ |

Table 7.6: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 11.0 to 12.5 GeV<sup>2</sup>/c<sup>4</sup> are shown. 5000/5000 pseudoexperiments converged successfully

| variable      | bias               | pull mean          | pull width        |
|---------------|--------------------|--------------------|-------------------|
| $F_L$         | $0.001 \pm 0.001$  | $0.005 \pm 0.014$  | $1.014 \pm 0.010$ |
| $S_3$         | $0.001 \pm 0.001$  | $-0.011 \pm 0.015$ | $1.049 \pm 0.010$ |
| $S_4$         | $-0.004 \pm 0.001$ | $-0.094 \pm 0.015$ | $1.073 \pm 0.011$ |
| $A_5$         | $0.001 \pm 0.001$  | $0.018 \pm 0.015$  | $1.049 \pm 0.010$ |
| $A_{FB}^{CP}$ | $-0.000 \pm 0.001$ | $-0.004 \pm 0.014$ | $1.008 \pm 0.010$ |
| $S_7$         | $-0.000 \pm 0.001$ | $-0.012 \pm 0.015$ | $1.037 \pm 0.010$ |
| $A_8$         | $0.002 \pm 0.001$  | $0.032 \pm 0.015$  | $1.029 \pm 0.010$ |
| $A_9$         | $0.000 \pm 0.001$  | $0.009 \pm 0.015$  | $1.044 \pm 0.010$ |

Table 7.7: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 1.1 to 6.0 GeV<sup>2</sup>/c<sup>4</sup> are shown. 5000/5000 pseudoexperiments converged successfully.

| variable      | bias               | pull mean          | pull width        |
|---------------|--------------------|--------------------|-------------------|
| $F_L$         | $0.001 \pm 0.001$  | $0.051 \pm 0.014$  | $1.003 \pm 0.010$ |
| $S_3$         | $0.001 \pm 0.001$  | $-0.004 \pm 0.015$ | $1.027 \pm 0.010$ |
| $S_4$         | $-0.001 \pm 0.001$ | $-0.044 \pm 0.015$ | $1.027 \pm 0.010$ |
| $A_5$         | $-0.000 \pm 0.001$ | $-0.016 \pm 0.014$ | $1.002 \pm 0.010$ |
| $A_{FB}^{CP}$ | $-0.001 \pm 0.001$ | $-0.031 \pm 0.014$ | $1.016 \pm 0.010$ |
| $S_7$         | $-0.001 \pm 0.001$ | $-0.017 \pm 0.014$ | $1.009 \pm 0.010$ |
| $A_8$         | $0.000 \pm 0.001$  | $0.001 \pm 0.014$  | $1.000 \pm 0.010$ |
| $A_9$         | $-0.000 \pm 0.001$ | $-0.018 \pm 0.014$ | $1.001 \pm 0.010$ |

Table 7.8: Absolute bias, pull means and pull widths for the angular observables from blind fits to data in the bin from 15.0 to 18.9 GeV<sup>2</sup>/c<sup>4</sup> are shown. 5000/5000 pseudoexperiments converged successfully.

| variable      | bias               | pull mean          | pull width        |
|---------------|--------------------|--------------------|-------------------|
| $F_L$         | $-0.003 \pm 0.000$ | $-0.138 \pm 0.015$ | $1.044 \pm 0.010$ |
| $S_3$         | $-0.001 \pm 0.001$ | $-0.078 \pm 0.014$ | $1.020 \pm 0.010$ |
| $S_4$         | $-0.000 \pm 0.001$ | $-0.046 \pm 0.014$ | $1.016 \pm 0.010$ |
| $A_5$         | $0.001 \pm 0.001$  | $0.001 \pm 0.015$  | $1.028 \pm 0.010$ |
| $A_{FB}^{CP}$ | $0.000 \pm 0.000$  | $-0.006 \pm 0.014$ | $1.011 \pm 0.010$ |
| $S_7$         | $0.001 \pm 0.001$  | $0.011 \pm 0.014$  | $1.000 \pm 0.010$ |
| $A_8$         | $0.001 \pm 0.001$  | $0.016 \pm 0.015$  | $1.039 \pm 0.010$ |
| $A_9$         | $0.000 \pm 0.001$  | $-0.002 \pm 0.014$ | $1.023 \pm 0.010$ |

# Chapter 8

## Systematic uncertainties

In this Chapter, the approach to determine the systematic uncertainties will be introduced in Sec. 8.1 and all considered sources of systematic uncertainty will be discussed, subsequently. These include the systematic uncertainties associated with the omission of peaking backgrounds in the fit model (Sec. 8.2), the choice of the acceptance orders (Sec. 8.3), the effect of neglecting time-dependent efficiencies (Sec. 8.4), the effect of using simulated  $B_s^0 \rightarrow \phi\mu^+\mu^-$  events with  $\Delta\Gamma_s = 0$  (Sec. 8.5), the corrections to simulation excluding PID (Sec. 8.6), the PID correction (Sec. 8.7), the effect of limited size of simulated samples on the angular acceptance (Sec. 8.8), the choice of the signal mass model (Sec. 8.9), the background modeling (Sec. 8.10), sharing the angular background parameters across data sets (Sec. 8.11), using an acceptance evaluation point in  $q^2$  (Sec. 8.12), neglecting the  $S$ -wave contribution (Sec. 8.13), and the remaining fit bias (Sec. 8.14). The combination of the systematic uncertainties is given in Sec. 8.15.

### 8.1 Approach to the determination of systematic uncertainties

Systematic uncertainties are determined for specific assumptions taken in this analysis, such as the maximum order of the acceptance or signal mass model amongst others. These choices can bias the measured angular observables. The sources of systematic uncertainty can be divided into two categories: (i) effects arising from mismodelling of the angular acceptance, (ii) effects due to assumptions imposed on the fit model and effects originating from the composition of the events (*e.g.* additional background events). The procedure used to determine systematic uncertainties relies on pseudoexperiments.

For systematic uncertainties arising through case (i), an alternative acceptance is calculated and used to generate pseudoexperiments in each  $q^2$  region. The generated pseudoexperiments are then fit once using the varied acceptance and once using the default acceptance. These fit results are generally not identical and exhibit shifts in the angular observables. The resulting distribution of the differences between the ensembles of observables extracted from the fits using the altered and default acceptances is fit with a Gaussian function. The mean of this Gaussian function gives the parameter shift and is quoted as the systematic uncertainty. By fitting the pseudoexperiments twice, once with the alternative and once with the default model, fluctuations introduced by finite pseudoexperiment statistics are reduced.

For the majority of systematic uncertainties falling under case (ii) a very similar approach

is followed, but the pseudoexperiments are generated using an alternative fit model. The experiments are then fit with both the default and alternative model, and the differences are assessed in the same way.

An exception are systematic uncertainties relating to neglecting certain backgrounds in the fit model. Here, it is not feasible to alter the fit model and instead only the composition of the pseudoevents in the pseudoexperiments is altered to reflect the expected contamination from given backgrounds. The systematic uncertainty is taken as the difference between the angular observables used to generate the pseudoexperiments and the values for the angular observables extracted from the fits.

For each pseudoexperiment, 50 000 events are generated for the Run 1 data set. The number of events generated for the Run 2p1 and Run 2p2 data sets corresponds to 50 000 events scaled with the ratio of the yields observed in blind fits to that  $q^2$  region in Run 1 and Run 2p1 or Run 2p2, respectively. For each systematic uncertainty investigated, 500 pseudoexperiments are generated per  $q^2$  region. For the high statistics pseudoexperiments used for the evaluation of the systematic uncertainties, the success rate of the fit is 100%.

## 8.2 Peaking backgrounds

Only background events originating from combinatorial background are considered in the fit model used in the fit to data. In contrast, peaking backgrounds, namely,  $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$  and  $B^0 \rightarrow K^{*0}\mu^+\mu^-$  decays, are neglected. The effect of neglecting these peaking backgrounds is covered by assigning a systematic uncertainty. The systematic uncertainty is evaluated by injecting events representative of the relevant neglected background into the pseudoexperiment. The background shapes for  $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$  and  $B^0 \rightarrow K^{*0}\mu^+\mu^-$  decays for the angular and mass distributions are modelled using a kernel density estimation obtained from simulation. The number of events injected into each pseudoexperiment is determined by the ratio of expected background to signal events as given in Tab. 6.5. The pseudoexperiments are fitted and the difference of the fitted value and the generated value for each angular observable is taken as a systematic uncertainty.

The resulting absolute values for the systematic uncertainty are shown for every variable in every  $q^2$  region in Tab. 8.1. The systematic uncertainty is found to be 4.2% of the statistical uncertainty or less.

Table 8.1: Values for the systematic uncertainty originating from neglecting the peaking backgrounds in the fit model.

| $q^2$ region                                  | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-----------------------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| [0.1, 0.98] GeV <sup>2</sup> /c <sup>4</sup>  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000                | -0.000 | 0.000  | -0.000 |
| [1.1, 4.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.000 | 0.000  | -0.000 | -0.000 | -0.000               | -0.000 | 0.000  | 0.000  |
| [4.0, 6.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.000 | -0.000 | 0.000  | -0.000 | 0.000                | -0.000 | -0.000 | 0.000  |
| [6.0, 8.0] GeV <sup>2</sup> /c <sup>4</sup>   | 0.000  | -0.000 | 0.000  | -0.000 | 0.000                | 0.000  | -0.000 | 0.000  |
| [11.0, 12.5] GeV <sup>2</sup> /c <sup>4</sup> | -0.000 | 0.000  | -0.000 | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |
| [1.1, 6.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.000 | -0.000 | 0.000  | -0.000 | -0.000               | -0.000 | -0.000 | 0.000  |
| [15.0, 18.9] GeV <sup>2</sup> /c <sup>4</sup> | -0.001 | -0.000 | 0.000  | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |

### 8.3 Acceptance orders

The choice of the orders for the Legendre polynomials used to parameterise the acceptance can potentially influence the fit results for the angular observables. Therefore, a systematic uncertainty is assigned to the specific choice of orders by determining a new acceptance with the next highest, allowed order. For the angles the orders are increased by 2 as only even orders are taken into account and 1 for  $q^2$  as also odd orders are used. With these new acceptances the systematic uncertainty is evaluated according to approach (i). The resulting systematic uncertainty is shown for every variable in every  $q^2$  region in Tab. 8.2. The size of the systematic uncertainty is below 2.3% of the statistical uncertainty in all  $q^2$  regions except for the region of  $q^2 \in [15.0, 18.9] \text{ GeV}^2/c^4$  where it has a size of 31.7% of the statistical uncertainty. This  $q^2$  region is especially sensitive to small changes in the modelling of the acceptance given the low number of events in simulation close to the kinematically allowed border of  $q^2 = 18.9 \text{ GeV}^2/c^4$ .

Table 8.2: Values for the systematic uncertainty originating from increasing the acceptance orders to the next higher order.

| $q^2$ region                     | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|----------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| $[0.1, 0.98] \text{ GeV}^2/c^4$  | -0.001 | -0.002 | -0.000 | -0.000 | -0.000               | -0.000 | -0.000 | 0.000  |
| $[1.1, 4.0] \text{ GeV}^2/c^4$   | -0.001 | 0.000  | 0.000  | 0.000  | -0.000               | -0.000 | 0.000  | -0.000 |
| $[4.0, 6.0] \text{ GeV}^2/c^4$   | 0.001  | 0.000  | 0.000  | -0.000 | 0.000                | 0.000  | 0.000  | -0.000 |
| $[6.0, 8.0] \text{ GeV}^2/c^4$   | 0.001  | 0.000  | 0.000  | 0.000  | -0.000               | 0.000  | -0.000 | 0.000  |
| $[11.0, 12.5] \text{ GeV}^2/c^4$ | 0.000  | -0.000 | 0.000  | -0.000 | 0.000                | -0.000 | 0.000  | -0.000 |
| $[1.1, 6.0] \text{ GeV}^2/c^4$   | -0.000 | 0.000  | -0.000 | -0.000 | 0.000                | 0.000  | -0.000 | 0.000  |
| $[15.0, 18.9] \text{ GeV}^2/c^4$ | -0.009 | -0.004 | 0.000  | 0.001  | -0.001               | 0.000  | 0.001  | -0.000 |

### 8.4 Effect of neglecting time-dependent efficiencies

The combined angular coefficients  $J_i \pm \bar{J}_i$  used to construct the angular observables  $S_i$  and  $A_i$  are time-dependent as can be seen in Eq. 2.9 and Eq. 2.10. In this analysis, the angular observables are measured integrated over time ignoring the time-dependent efficiency ( $\epsilon(t)$ ). The efficiency  $\epsilon(t)$  is determined from centrally produced  $B_s^0 \rightarrow J/\psi \phi$  simulation, where the normalisation is taken from a privately generated sample on generator level with particle gun. The efficiencies for the three data sets are shown in Fig. 8.1. The difference between the time-integrated observables using  $\epsilon(t) = 1$  and  $\epsilon(t)$  as taken from simulated  $B_s^0 \rightarrow J/\psi \phi$  events is assigned as the systematic uncertainty and can be found in Tab. 8.3. The time-dependent terms have the form

$$\begin{aligned}
& \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) dt, \\
& \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) dt, \\
& \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cos(\Delta m_s t) dt \\
& \text{and } \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \sin(\Delta m_s t) dt
\end{aligned}$$

and the time integration is performed numerically in steps of size 0.2 ps in the region of  $t \in [0, 20]$  ps. Furthermore, the constants needed for the integral are taken as  $\Delta\Gamma_s/\Gamma_s = 0.166$ ,  $x = \Delta m_s/\Gamma_s = 26.89$  and  $\Gamma_s = 65.6895 \cdot 10^{10} s^{-1}$  [65], whereas the angular coefficients  $J_i$ ,  $\bar{J}_i$  and  $h_i$  are taken from the **flavio** software package. The  $s_i$  coefficients are not implemented in the **flavio** software package and are, therefore, calculated following Ref. [57].

This systematic uncertainty is below 19.7% of the statistical uncertainty everywhere and is one of the dominant systematic uncertainties.

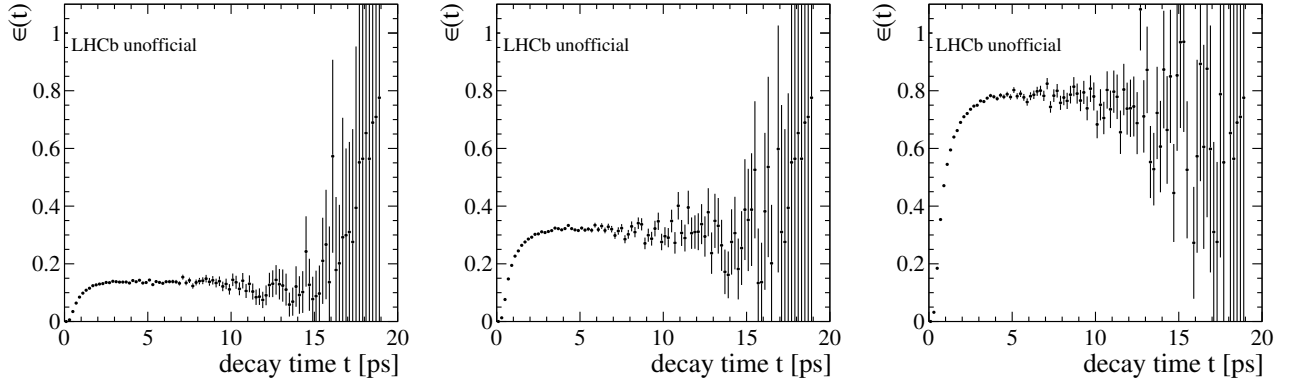


Figure 8.1: The efficiency of the selection for Run 1 and Run 2p1 and Run 2p2 from left to right, respectively, as a function of the time in units picoseconds.

Table 8.3: Values for the systematic uncertainty originating from the effect of time-dependent efficiencies in the time integration.

| $q^2$ region                     | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{FB}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
|----------------------------------|-------|-------|-------|-------|---------------|-------|-------|-------|
| $[0.1, 0.98] \text{ GeV}^2/c^4$  | 0.009 | 0.010 | 0.002 | 0.000 | 0.000         | 0.001 | 0.000 | 0.001 |
| $[1.1, 4.0] \text{ GeV}^2/c^4$   | 0.006 | 0.003 | 0.001 | 0.000 | 0.000         | 0.000 | 0.000 | 0.000 |
| $[4.0, 6.0] \text{ GeV}^2/c^4$   | 0.006 | 0.004 | 0.002 | 0.001 | 0.000         | 0.000 | 0.001 | 0.000 |
| $[6.0, 8.0] \text{ GeV}^2/c^4$   | 0.008 | 0.006 | 0.003 | 0.001 | 0.000         | 0.000 | 0.001 | 0.001 |
| $[11.0, 12.5] \text{ GeV}^2/c^4$ | 0.007 | 0.009 | 0.005 | 0.001 | 0.001         | 0.000 | 0.001 | 0.001 |
| $[1.1, 6.0] \text{ GeV}^2/c^4$   | 0.006 | 0.004 | 0.001 | 0.000 | 0.000         | 0.000 | 0.000 | 0.000 |
| $[15.0, 18.9] \text{ GeV}^2/c^4$ | 0.004 | 0.007 | 0.003 | 0.000 | 0.001         | 0.000 | 0.001 | 0.001 |

## 8.5 Effect of using simulated $B_s^0 \rightarrow \phi \mu^+ \mu^-$ events with $\Delta\Gamma_s = 0$

Simulated  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays are generated assuming  $\Delta\Gamma_s = 0$ . This means that the  $B_s^0$  lifetime distribution is incorrectly modelled and the correlation between the lifetime of the  $B_s^0$  and the angular terms is lost. As a consequence, the resulting effect on the angular and  $q^2$ -dependent efficiency of any selection shaping the lifetime distribution may not be properly modelled. To study the size of this effect, a data set is generated from a five-dimensional

PDF which correctly accounts for the dependence of the decay rate on  $q^2$ , the angles and the  $B_s^0$  lifetime. The five-dimensional PDF is taken from Eqs. 2.9-2.10 following Ref. [57]. The numerical values for  $I_i(q^2)$ ,  $\bar{I}_i(q^2)$  and  $h_i$  are taken from the **flavio** software package [109], and the lifetime  $\Gamma_s$  and lifetime difference  $\Delta\Gamma_s$  are set to the values used in the **EvtGen** simulation, namely  $\Gamma_s = 65.6895 \times 10^{10} s^{-1}$  and  $\Delta\Gamma_s/\Gamma_s = 0.166$  [65]. Given that the  $CP$ -asymmetries are approximately zero in SM, the time dependence of these terms are ignored.

Weights are then derived for the phasespace simulation used for the determination of the angular acceptance by comparing the phasespace simulation at generator-level with the five-dimensional distribution simulated using the PDF given by the **flavio** software package. Crucially, the application of these weights to the phasespace simulation means that the angular and  $q^2$  distributions now follow a physics model, and are no longer flat at generator-level. To ensure the efficiencies still have the correct shape, a second set of four-dimensional weights is applied to guarantee flatness in the angles and  $q^2$  at generator level. These weights correspond to a division of the shape after all selection by the shape at the generator-level. This set of two separate weights is used to determine an altered acceptance which is compared with the default acceptance following approach (i). The resulting systematic uncertainty is shown for every variable in every  $q^2$  region in Tab. 8.4. This systematic uncertainty is found to be 25.0% of the statistical uncertainty or less and constitutes the dominant systematic uncertainty in most  $q^2$  regions.

Table 8.4: Values for the systematic uncertainty originating from the effect of using simulated  $B_s^0 \rightarrow \phi\mu^+\mu^-$  events with  $\Delta\Gamma_s = 0$ .

| $q^2$ region                    | $F_L$ | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|---------------------------------|-------|--------|--------|--------|---------------|--------|--------|--------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | 0.011 | -0.010 | -0.000 | -0.000 | 0.000         | -0.000 | -0.000 | -0.000 |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | 0.007 | -0.003 | 0.001  | -0.000 | -0.000        | 0.000  | -0.000 | 0.000  |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | 0.008 | -0.004 | 0.002  | -0.000 | 0.000         | -0.000 | 0.000  | -0.000 |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | 0.008 | -0.007 | 0.002  | -0.000 | 0.000         | -0.000 | -0.000 | -0.000 |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 0.007 | -0.009 | 0.002  | 0.000  | 0.000         | -0.000 | -0.000 | 0.000  |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 0.009 | -0.005 | 0.001  | -0.000 | 0.000         | 0.000  | -0.000 | 0.000  |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 0.006 | -0.008 | 0.001  | 0.000  | 0.000         | 0.000  | -0.000 | -0.000 |

## 8.6 Corrections to simulation based on weights

The response of the L0 trigger, the **nTracks** distribution and the  $B_s^0 p_T$  distribution in simulation are corrected for mismodelling by using weights derived from the  $B_s^0 \rightarrow J/\psi\phi$  channel. These weights can impact the efficiency shape in the decay angles and  $q^2$ . To estimate the size of the systematic uncertainty arising from these corrections, angular acceptances are recalculated without these corrections. This is performed for each of the three corrections and the largest difference for a given angular observable is taken as systematic uncertainty. The resulting systematic uncertainty is shown for every variable in every  $q^2$  region in Tab. 8.5. This systematic uncertainty is found to be 10.7% of the statistical uncertainty or less except for the  $q^2$  region of [15.0, 18.9]  $\text{GeV}^2/c^4$ , where its size is 43.6% in  $F_L$ . The large effect of the correction on the angular observables in this  $q^2$  region is due to a weighted fit being performed in this region. The weighted fit is sensitive to large weights, which can occur in this  $q^2$  region following to the sparse population in simulation as a consequence of the kinematic boundary at  $q^2 = 18.9 \text{ GeV}^2/c^4$ .

Table 8.5: Values for the systematic uncertainty originating from the corrections to simulation.

| $q^2$ region                    | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
|---------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | 0.005 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | 0.003 | 0.000 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | 0.000 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | 0.002 | 0.000 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 0.004 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 0.003 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 0.013 | 0.002 | 0.000 | 0.001 | 0.000                | 0.001 | 0.001 | 0.000 |

## 8.7 Correction of the PID response

The PID response in the simulated samples is corrected for using the **MEERKAT** tool. Following the recommendations given in the **MEERKAT** twiki [227], a systematic uncertainty for the correction of the PID response is determined by using alternative PID response templates to obtain the corrected PID response. There are separate templates covering two different effects. The first alternative template estimates the effect associated with using a given kernel width when parameterising the PID control samples. Here, the kernel density estimates in **MEERKAT** are recalculated using a 50% wider kernel width. The second considered effect is the limited size of the PID calibration samples used when correcting the PID variables. Here, five templates are derived to estimate the size of this effect. While using only five different templates gives limited precision, using five templates is sufficient given the small size of the effect. For each new template the selection is then rerun with the new PID response variables. Using this new selection a new acceptance is determined and the systematic uncertainty is determined following approach (i). As these effects are independent of one another, they are added in quadrature. The resulting systematic uncertainty is shown for every variable in all  $q^2$  regions in Tab. 8.6. This systematic uncertainty is found to be 6.6% of the statistical uncertainty or less.

Table 8.6: Values for the systematic uncertainty originating from the correction of the PID variables.

| $q^2$ region                    | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
|---------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | 0.001 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | 0.000 | 0.000 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | 0.001 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | 0.001 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 0.000 | 0.001 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 0.000 | 0.000 | 0.000 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 0.002 | 0.002 | 0.001 | 0.000 | 0.000                | 0.000 | 0.000 | 0.000 |

## 8.8 Limited size of simulated samples for acceptance

A systematic uncertainty is assigned to account for the limited size of the simulated samples used to calculate the angular acceptance. For each pseudoexperiment a new acceptance is determined by bootstrapping the simulation samples with a unique seed. Using these 500 altered acceptances, the systematic uncertainty is determined following approach (i). Here, however, the spread of the pseudoexperiments instead of the mean shift is taken as the value for the systematic uncertainty. The resulting systematic uncertainty is shown in Tab. 8.7. This systematic uncertainty is found to be 0.3% of the statistical uncertainty or less except for the  $q^2$  region of  $[15.0, 18.9] \text{ GeV}^2/c^4$ , where its size is 19.9% in  $F_L$ . The increase of size in this  $q^2$  region originates in the angular acceptance being sensitive to small variations in the number of available events given the sparse population in this  $q^2$  region in simulation.

Table 8.7: Values for the systematic uncertainty originating from the statistics of the phase-space simulation.

| $q^2$ region                     | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|----------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| $[0.1, 0.98] \text{ GeV}^2/c^4$  | 0.000  | 0.000  | 0.000  | 0.000  | -0.000               | 0.000  | -0.000 | -0.000 |
| $[1.1, 4.0] \text{ GeV}^2/c^4$   | -0.000 | -0.000 | -0.000 | 0.000  | -0.000               | -0.000 | -0.000 | 0.000  |
| $[4.0, 6.0] \text{ GeV}^2/c^4$   | 0.000  | -0.000 | 0.000  | -0.000 | -0.000               | -0.000 | 0.000  | -0.000 |
| $[6.0, 8.0] \text{ GeV}^2/c^4$   | 0.000  | -0.000 | 0.000  | -0.000 | -0.000               | -0.000 | -0.000 | -0.000 |
| $[11.0, 12.5] \text{ GeV}^2/c^4$ | 0.000  | -0.000 | 0.000  | 0.000  | 0.000                | -0.000 | 0.000  | -0.000 |
| $[1.1, 6.0] \text{ GeV}^2/c^4$   | -0.000 | -0.000 | 0.000  | -0.000 | 0.000                | 0.000  | 0.000  | -0.000 |
| $[15.0, 18.9] \text{ GeV}^2/c^4$ | -0.006 | -0.002 | 0.000  | 0.000  | -0.000               | -0.000 | -0.000 | 0.000  |

## 8.9 Signal mass model

The reconstructed  $B_s^0$  mass is used to separate signal from combinatorial background events. The choice of signal mass model can have an influence on the angular observables. A systematic uncertainty is assigned to the specific choice of the signal mass model by comparing the default model to an alternative model, namely the Hypatia function [242]. The parameters for the Hypatia function are determined in a similar approach as is chosen for the double Crystal Ball function. The tail parameters are fixed from fits to simulated  $B_s^0 \rightarrow J/\psi \phi$  events and the width from  $B_s^0 \rightarrow J/\psi \phi$  data. A width scale-factor is determined for each  $q^2$  region using rare mode simulation. The systematic uncertainty is then determined based on approach (ii). The resulting systematic uncertainty is shown for every variable and all  $q^2$  regions in Tab. 8.8. This systematic uncertainty is found to be 0.4% of the statistical uncertainty or less in the narrow  $q^2$  regions and 16.7% or less in the wide  $q^2$  regions. The wide  $q^2$  regions are affected the most by the change in the fit model due to the scale factors for the width being  $q^2$  dependent. As a consequence the larger a  $q^2$  region is, the larger the effect is expected to be. Nevertheless, the systematic uncertainty assigned to the signal mass model is not dominant in any  $q^2$  region.



Table 8.8: Values for the systematic uncertainty originating from using different mass models.

| $q^2$ region                    | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|---------------------------------|--------|--------|--------|--------|---------------|--------|--------|--------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000         | -0.000 | -0.000 | -0.000 |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | 0.000  | 0.000  | -0.000 | 0.000  | -0.000        | -0.000 | 0.000  | -0.000 |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | -0.000 | 0.000  | 0.000  | -0.000 | -0.000        | 0.000  | -0.000 | 0.000  |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | -0.000 | 0.000  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 0.000  | -0.000 | -0.000 | 0.000  | 0.000         | -0.000 | 0.000  | -0.000 |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 0.002  | 0.001  | -0.002 | -0.000 | 0.000         | 0.000  | 0.000  | -0.000 |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | -0.005 | 0.000  | -0.000 | -0.000 | 0.000         | -0.000 | -0.000 | -0.000 |

## 8.10 Background modeling

The angular background is modeled with Chebyshev polynomials of order one. A systematic uncertainty is assigned to this choice by investigating the effect of using polynomials of order two in the fit. The coefficients describing the combinatorial background are determined by fitting the mass sideband in the ranges [5100, 5270]  $\text{MeV}/c^2$  and [5570, 5700]  $\text{MeV}/c^2$  with the increased order. Then pseudoexperiments following approach (ii) are generated with these coefficients and fit back with the altered and default fit models. The resulting systematic uncertainty is shown for every variable in every  $q^2$  region in Tab. 8.9. This systematic uncertainty is found to be 21.2% of the statistical uncertainty or less. The regions most affected by the modelling of the background are the  $q^2$  regions of [1.1, 4.0] and [4.0, 6.0]. These  $q^2$  regions suffer from low event yields and are, therefore, more sensitive to small changes in the modelling of the background.

Table 8.9: Values for the systematic uncertainty originating from increasing the order of the background Chebychev polynomials from one to two.

| $q^2$ region                    | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|---------------------------------|--------|--------|--------|--------|---------------|--------|--------|--------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | -0.000 | -0.001 | 0.001  | -0.000 | 0.000         | 0.000  | 0.000  | 0.000  |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | -0.006 | 0.000  | -0.001 | 0.000  | -0.000        | 0.000  | 0.000  | 0.000  |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | 0.010  | 0.001  | -0.003 | 0.000  | 0.001         | 0.000  | 0.000  | -0.000 |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | -0.003 | -0.002 | -0.001 | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 0.004  | -0.002 | 0.002  | -0.000 | 0.000         | -0.000 | -0.000 | 0.000  |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 0.002  | 0.000  | -0.002 | 0.000  | 0.000         | 0.000  | 0.000  | -0.000 |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 0.004  | -0.001 | 0.001  | 0.000  | 0.000         | -0.000 | -0.000 | 0.000  |

## 8.11 Shared angular background parameters

The angular background parameters are shared across data sets to improve fit stability. In particular, the HESSE algorithm fails to converge for roughly 20% of pseudoexperiment fits in the first  $q^2$  region if the angular background parameters are not shared. However, the background parameters are not guaranteed to be identical across the different data taking periods. To assess the systematic uncertainty associated with sharing the angular background description, the upper mass side band in data is fitted for each year independently, and the resulting angular coefficients are used to generate high statistics pseudoexperiments. Following approach (ii), these

pseudoexperiments are then fitted twice, once with a shared background, once without, and the resulting difference on the angular observables is taken as a systematic uncertainty. The resulting effect on the angular observables is given in Tab. 8.10. This systematic uncertainty is found to be 3.2% of the statistical uncertainty or less except for the  $q^2$  region of  $[0.1, 0.98] \text{ GeV}^2/c^4$ , where its size is 12.0% in  $F_L$ . The increase of size in this  $q^2$  region is expected as this region is the only region to suffer from HESSE not converging when not sharing the angular background parameters.

Table 8.10: Values for the systematic uncertainty originating from the shared angular background parameters.

| $q^2$ region                     | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|----------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| $[0.1, 0.98] \text{ GeV}^2/c^4$  | 0.006  | -0.001 | 0.000  | -0.000 | -0.001               | -0.000 | 0.000  | -0.000 |
| $[1.1, 4.0] \text{ GeV}^2/c^4$   | -0.001 | 0.000  | -0.000 | -0.000 | 0.000                | -0.000 | -0.000 | -0.000 |
| $[4.0, 6.0] \text{ GeV}^2/c^4$   | 0.000  | 0.000  | 0.000  | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |
| $[6.0, 8.0] \text{ GeV}^2/c^4$   | 0.000  | 0.000  | 0.000  | 0.000  | 0.000                | 0.000  | -0.000 | -0.000 |
| $[11.0, 12.5] \text{ GeV}^2/c^4$ | 0.000  | 0.000  | -0.000 | 0.000  | -0.000               | 0.000  | 0.000  | 0.000  |
| $[1.1, 6.0] \text{ GeV}^2/c^4$   | 0.000  | -0.000 | -0.000 | -0.000 | -0.000               | -0.000 | -0.000 | -0.000 |
| $[15.0, 18.9] \text{ GeV}^2/c^4$ | -0.000 | -0.000 | 0.000  | -0.000 | 0.000                | -0.000 | -0.000 | -0.000 |

## 8.12 Acceptance evaluation point in $q^2$

For the narrow  $q^2$  regions the acceptance is evaluated at a specific point in  $q^2$ . The choice of the evaluation point can impact the value for the angular observables as the  $q^2$  efficiency shape is not perfectly flat across the  $q^2$  regions. To estimate the size of the effect of the  $q^2$  variation inside a  $q^2$  region, events are generated with the  $q^2$  shape from the SM prediction obtained by the `flavio` software package. Only signal events are generated for this systematic. These events are fit twice, once with the acceptance entering as a per-event weight, which is sensitive to the variation of both the  $q^2$  efficiency and distribution of  $q^2$  inside a  $q^2$  region, and once with the default approach. The difference between the fit results is assigned as a systematic uncertainty. The resulting systematic uncertainty is shown for every variable in every  $q^2$  region in Tab. 8.11. It should be noted, that the  $q^2$  regions of  $[1.1, 6.0] \text{ GeV}^2/c^4$  and  $[15.0, 18.9] \text{ GeV}^2/c^4$  are performed as weighted fits, which already account for the  $q^2$  dependence of the acceptance and this systematic uncertainty is, therefore, not applicable in these  $q^2$  regions. The systematic uncertainty is found to be 6.7% of the statistical uncertainty or less.

Table 8.11: Values for the systematic uncertainty originating from neglecting the  $q^2$  dependence.

| $q^2$ region                     | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|----------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| $[0.1, 0.98] \text{ GeV}^2/c^4$  | -0.003 | 0.001  | -0.000 | -0.000 | -0.000               | 0.000  | -0.000 | -0.000 |
| $[1.1, 4.0] \text{ GeV}^2/c^4$   | 0.002  | -0.000 | 0.000  | -0.000 | 0.000                | 0.000  | 0.000  | 0.000  |
| $[4.0, 6.0] \text{ GeV}^2/c^4$   | -0.001 | -0.000 | -0.000 | -0.000 | 0.000                | -0.000 | 0.000  | 0.000  |
| $[6.0, 8.0] \text{ GeV}^2/c^4$   | -0.000 | -0.000 | -0.000 | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |
| $[11.0, 12.5] \text{ GeV}^2/c^4$ | -0.000 | -0.000 | 0.000  | -0.000 | -0.000               | 0.000  | 0.000  | 0.000  |

## 8.13 Neglecting $S$ -wave contribution

In this analysis, the contribution from events originating from the two kaons in the spin 0 configuration ( $S$ -wave) is neglected in the fit. This is motivated by the low amount of expected  $S$ -wave events in the narrow  $\phi$  window. In the dedicated LHCb  $B_s^0 \rightarrow J/\psi \phi$  analyses using the Run 1 and Run 2p1 data sets, an  $S$ -wave fraction of approximately 2% has been measured [239, 240]. To estimate the effect that neglecting a 2%  $S$ -wave component would have on the angular observables, 500 pseudoexperiments are generated from a PDF in which the  $S$ -wave component has been included. The PDF including the  $S$ - and  $P$ -wave has the form

$$\begin{aligned} \frac{1}{d(\Gamma + \bar{\Gamma}/dq^2} \frac{d^4(\Gamma + \bar{\Gamma})}{dq^2 d^3\vec{\Omega}}|_{S+P} &= (1 - F_S) \frac{1}{d(\Gamma + \bar{\Gamma}/dq^2} \frac{d^4(\Gamma + \bar{\Gamma})}{dq^2 d^3\vec{\Omega}}|_P \\ &+ \frac{3}{16\pi} F_S \sin^2 \theta_l \\ &+ \frac{9}{32\pi} (S_{11} + S_{13} \cos 2\theta_l) \cos \theta_K \\ &+ \frac{9}{32\pi} (S_{14} \sin 2\theta_l + S_{13} \sin \theta_l) \sin \theta_K \cos \phi \\ &+ \frac{9}{32\pi} (S_{16} \sin \theta_l + S_{13} \sin 2\theta_l) \sin \theta_K \sin \phi. \end{aligned}$$

For this systematic uncertainty the interference terms are set to zero and the  $S$ -wave fraction ( $F_S$ ) is set to 2%. These generated events are then fit with the default PDF and the difference of the values for the angular observables to the equivalent generated value is taken as the uncertainty. The resulting systematic uncertainty is shown for every variable in every  $q^2$  region in Tab. 8.12. This systematic uncertainty is found to be 19.9% of the statistical uncertainty or less and, therefore, one of the dominating systematic uncertainties.

Table 8.12: Values for the systematic uncertainty originating from neglecting the  $S$ -wave contribution.

| $q^2$ region                                  | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-----------------------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| [0.1, 0.98] GeV <sup>2</sup> /c <sup>4</sup>  | -0.000 | -0.001 | -0.002 | 0.000  | 0.000                | 0.000  | 0.000  | -0.000 |
| [1.1, 4.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.009 | 0.000  | 0.001  | -0.000 | -0.000               | 0.000  | 0.000  | 0.000  |
| [4.0, 6.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.006 | 0.001  | 0.003  | 0.000  | 0.000                | 0.000  | -0.000 | 0.000  |
| [6.0, 8.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.004 | 0.001  | 0.004  | -0.000 | 0.000                | 0.000  | -0.000 | 0.000  |
| [11.0, 12.5] GeV <sup>2</sup> /c <sup>4</sup> | -0.000 | 0.003  | 0.004  | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |
| [1.1, 6.0] GeV <sup>2</sup> /c <sup>4</sup>   | -0.006 | 0.000  | 0.002  | -0.000 | -0.000               | 0.000  | 0.000  | 0.000  |
| [15.0, 18.9] GeV <sup>2</sup> /c <sup>4</sup> | 0.000  | 0.005  | 0.005  | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |

## 8.14 Remaining fit bias

In Sec. 7.3 the fit validation using pseudoexperiments was detailed. A remaining bias of up to roughly 15% of the expected statistical uncertainty was found in pseudoexperiments using the angular observables from blind fits to data. The bias in the value distribution of each observable in each  $q^2$  region is assigned as a systematic uncertainty to cover this effect. The resulting systematic uncertainty is shown for every variable and all  $q^2$  regions in Tab. 8.13.

This systematic uncertainty is found to be 15.9% of the statistical uncertainty or less. This systematic uncertainty often constitutes the dominant systematic uncertainty for all angular observables except  $F_L$  and  $S_3$ .

Table 8.13: Values for the systematic uncertainty originating from the remaining bias in the fit validation using the angular observables from blind fits to data.

| $q^2$ region                    | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|---------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | -0.002 | -0.002 | -0.004 | 0.000  | -0.009               | -0.000 | 0.004  | -0.001 |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | -0.002 | 0.000  | 0.000  | -0.002 | -0.000               | 0.001  | -0.001 | -0.001 |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | -0.002 | 0.010  | 0.009  | 0.003  | 0.002                | -0.003 | -0.001 | -0.001 |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | 0.000  | 0.000  | 0.001  | -0.001 | -0.001               | -0.002 | 0.002  | -0.001 |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | -0.001 | -0.001 | 0.004  | -0.001 | 0.000                | 0.000  | -0.002 | -0.000 |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | -0.001 | -0.001 | 0.002  | -0.001 | -0.001               | -0.002 | -0.001 | 0.001  |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 0.003  | 0.001  | 0.000  | -0.001 | -0.000               | -0.001 | -0.001 | -0.000 |

## 8.15 Combination of systematic uncertainties

The systematic uncertainties are combined by summation in quadrature. The quadratic sum of the systematic uncertainties relative to the expected statistical uncertainty is shown for each  $q^2$  region separately in Tabs. 8.14-8.20 where each source is given explicitly in the table, again relative to the expected statistical uncertainty. In App. E the absolute values for the systematic uncertainties are also given for each  $q^2$  region separately. In Tab. 8.21 a summary table showing the absolute values of the combined systematic uncertainty for all  $q^2$  regions. The equivalent table showing the total systematic uncertainty relative to the statistical uncertainty in % is given in Tab. 8.22. This analysis is still statistically dominated, as the combined systematic uncertainty for any variable is always smaller than or equal to 0.02, while the statistical uncertainty is always larger than 0.03. The systematic uncertainty of  $F_L$  is typically largest at the level of 30% of the statistical uncertainty. The largest value for the systematic uncertainty of  $F_L$  is found to be at the 60% level in the  $q^2$  region [15.0, 18.9]  $\text{GeV}^2/c^4$ .

Table 8.14: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [0.1, 0.98]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | 0.0        | 0.1        | 0.0        | 0.1        | 0.0                | -0.1       | 0.1        | -0.0       |
| signal mass model             | 0.2        | 0.0        | 0.1        | 0.0        | 0.0                | -0.0       | -0.0       | -0.0       |
| angular background model      | -0.2       | -1.9       | 1.1        | -0.0       | 0.2                | 0.4        | 0.0        | 0.3        |
| acceptance order              | -1.9       | -2.3       | -0.1       | -0.0       | -0.0               | -0.0       | -0.0       | 0.0        |
| simulation correction         | 10.7       | 1.3        | 0.1        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| PID correction                | 1.7        | 1.3        | 0.0        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| simulation statistics         | 0.2        | 0.6        | 0.0        | 0.0        | -0.0               | 0.0        | -0.0       | -0.0       |
| $q^2$ dependency              | -6.7       | 0.7        | -0.2       | -0.0       | -0.0               | 0.0        | -0.0       | -0.0       |
| S-wave                        | -0.2       | -0.9       | -2.2       | 0.1        | 0.0                | 0.7        | 0.2        | -0.0       |
| lifetime mismodelling         | 22.3       | -14.2      | -0.1       | -0.0       | 0.0                | -0.4       | -0.0       | -0.0       |
| time integration              | 19.7       | 14.0       | 3.1        | 0.4        | 0.3                | 0.9        | 0.4        | 1.2        |
| shared backgrounds            | 12.0       | -0.9       | 0.1        | -0.0       | -1.4               | -0.1       | 0.0        | -0.1       |
| remaining Bias                | -3.8       | -2.9       | -5.2       | 0.6        | -14.4              | -0.1       | 5.0        | -1.6       |
| quadratic sum                 | 34.8       | 20.5       | 6.6        | 0.8        | 14.5               | 1.3        | 5.0        | 2.0        |
| quadratic sum, absolute value | 0.017      | 0.014      | 0.005      | 0.001      | 0.009              | 0.001      | 0.004      | 0.001      |
| absolute expected statistical | 0.048      | 0.068      | 0.080      | 0.068      | 0.063              | 0.067      | 0.081      | 0.068      |

Table 8.15: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [1.1, 4.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | -0.2       | 0.0        | -0.0       | -0.1       | -0.0               | -0.1       | 0.3        | 0.0        |
| signal mass model             | 0.4        | 0.0        | -0.0       | 0.0        | -0.0               | -0.0       | 0.0        | -0.0       |
| angular background model      | -13.5      | 0.2        | -1.0       | 0.0        | -0.0               | 0.1        | 0.0        | 0.0        |
| acceptance order              | -2.1       | 0.4        | 0.0        | 0.0        | -0.0               | -0.0       | 0.0        | -0.0       |
| simulation correction         | 6.6        | 0.6        | 0.5        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| PID correction                | 0.6        | 0.7        | 0.0        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| simulation statistics         | -0.2       | -0.0       | -0.0       | 0.0        | -0.0               | -0.0       | -0.0       | 0.0        |
| $q^2$ dependency              | 4.1        | -0.2       | 0.3        | -0.0       | 0.0                | 0.0        | 0.1        | 0.0        |
| S-wave                        | -19.9      | 0.2        | 1.6        | -0.1       | -0.0               | 0.7        | 0.3        | 0.0        |
| lifetime mismodelling         | 14.3       | -4.7       | 0.8        | -0.0       | -0.0               | 0.0        | -0.0       | 0.0        |
| time integration              | 12.4       | 6.5        | 0.8        | 0.3        | 0.4                | 0.3        | 0.3        | 0.6        |
| shared backgrounds            | -3.2       | 0.0        | -0.3       | -0.0       | 0.2                | -0.1       | -0.0       | -0.2       |
| remaining Bias                | -4.1       | 0.0        | 0.5        | -3.0       | -1.1               | 0.8        | -1.4       | -1.7       |
| quadratic sum                 | 32.1       | 8.1        | 2.3        | 3.0        | 1.2                | 1.2        | 1.5        | 1.8        |
| quadratic sum, absolute value | 0.015      | 0.004      | 0.002      | 0.002      | 0.001              | 0.001      | 0.001      | 0.001      |
| absolute expected statistical | 0.047      | 0.054      | 0.079      | 0.071      | 0.045              | 0.071      | 0.080      | 0.054      |

Table 8.16: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [4.0, 6.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | -0.0       | -0.3       | 0.2        | -0.0       | 0.2                | -0.4       | -0.4       | 0.2        |
| signal mass model             | -0.2       | 0.0        | 0.1        | -0.0       | -0.0               | 0.0        | -0.0       | 0.0        |
| angular background model      | 21.2       | 2.1        | -4.0       | 0.0        | 1.1                | 0.3        | 0.1        | -0.3       |
| acceptance order              | 1.0        | 0.2        | 0.0        | -0.0       | 0.0                | 0.0        | 0.0        | -0.0       |
| simulation correction         | 0.4        | 0.8        | 0.1        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| PID correction                | 1.0        | 0.9        | 0.1        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| simulation statistics         | 0.0        | -0.0       | 0.0        | -0.0       | -0.0               | -0.0       | 0.0        | -0.0       |
| $q^2$ dependency              | -1.8       | -0.2       | -0.2       | -0.0       | 0.0                | -0.0       | 0.0        | 0.0        |
| S-wave                        | -12.1      | 1.1        | 4.4        | 0.0        | 0.0                | 0.1        | -0.2       | 0.3        |
| lifetime mismodelling         | 15.7       | -5.8       | 2.4        | -0.0       | 0.0                | -0.0       | 0.0        | -0.0       |
| time integration              | 13.1       | 6.7        | 2.4        | 0.8        | 0.2                | 0.1        | 0.7        | 0.5        |
| shared backgrounds            | 0.8        | 0.0        | 0.1        | 0.0        | -0.0               | -0.1       | -0.0       | -0.3       |
| remaining Bias                | -4.2       | 15.9       | 10.8       | 4.1        | 4.3                | -3.5       | -1.7       | -2.1       |
| quadratic sum                 | 32.2       | 18.4       | 12.8       | 4.2        | 4.4                | 3.5        | 1.9        | 2.2        |
| quadratic sum, absolute value | 0.016      | 0.012      | 0.010      | 0.003      | 0.002              | 0.003      | 0.002      | 0.001      |
| absolute expected statistical | 0.050      | 0.066      | 0.080      | 0.078      | 0.047              | 0.080      | 0.083      | 0.066      |

Table 8.17: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [6.0, 8.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | 0.2        | -0.1       | 0.1        | -0.3       | 0.0                | 0.1        | -0.1       | 0.0        |
| signal mass model             | -0.2       | 0.0        | 0.1        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| angular background model      | -5.9       | -2.8       | -0.9       | 0.1        | 0.0                | 0.0        | 0.0        | 0.0        |
| acceptance order              | 1.4        | 0.3        | 0.1        | 0.0        | -0.0               | 0.0        | -0.0       | 0.0        |
| simulation correction         | 3.2        | 0.3        | 0.4        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| PID correction                | 1.4        | 1.0        | 0.1        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| simulation statistics         | 0.2        | -0.0       | 0.0        | -0.0       | -0.0               | -0.0       | -0.0       | -0.0       |
| $q^2$ dependency              | -0.6       | -0.0       | -0.0       | 0.0        | -0.0               | -0.0       | -0.0       | -0.0       |
| S-wave                        | -7.1       | 1.7        | 5.4        | -0.3       | 0.2                | 0.4        | -0.1       | 0.1        |
| lifetime mismodelling         | 17.2       | -9.4       | 2.7        | -0.0       | 0.0                | -0.0       | -0.0       | -0.0       |
| time integration              | 15.6       | 9.3        | 4.0        | 0.8        | 0.6                | 0.1        | 0.9        | 0.7        |
| shared backgrounds            | 0.4        | 0.0        | 0.0        | 0.0        | 0.0                | 0.0        | -0.0       | -0.0       |
| remaining Bias                | 0.6        | 0.3        | 1.6        | -1.6       | -2.3               | -2.4       | 2.1        | -0.9       |
| quadratic sum                 | 25.3       | 13.7       | 7.5        | 1.8        | 2.4                | 2.5        | 2.3        | 1.1        |
| quadratic sum, absolute value | 0.012      | 0.009      | 0.006      | 0.001      | 0.001              | 0.002      | 0.002      | 0.001      |
| absolute expected statistical | 0.049      | 0.069      | 0.077      | 0.075      | 0.048              | 0.078      | 0.081      | 0.069      |

Table 8.18: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [11.0, 12.5]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | -0.2       | 0.3        | -0.0       | 0.2        | -0.2               | -0.2       | -0.1       | -0.2       |
| signal mass model             | 0.0        | -0.0       | -0.0       | 0.0        | 0.0                | -0.0       | 0.0        | -0.0       |
| angular background model      | 9.6        | -3.3       | 2.9        | -0.2       | 0.0                | -0.0       | -0.0       | 0.0        |
| acceptance order              | 0.7        | -0.2       | 0.2        | -0.0       | 0.0                | -0.0       | 0.0        | -0.0       |
| simulation correction         | 8.4        | 1.1        | 0.5        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| PID correction                | 0.7        | 1.6        | 0.2        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| simulation statistics         | 0.0        | -0.2       | 0.0        | 0.0        | 0.0                | -0.0       | 0.0        | -0.0       |
| $q^2$ dependency              | -0.5       | -0.2       | 0.0        | -0.0       | -0.0               | 0.0        | 0.0        | 0.0        |
| S-wave                        | -0.2       | 4.3        | 7.0        | 0.2        | -0.2               | -0.0       | -0.1       | -0.2       |
| lifetime mismodelling         | 17.5       | -13.6      | 2.5        | 0.0        | 0.0                | -0.0       | -0.0       | 0.0        |
| time integration              | 17.1       | 14.6       | 7.3        | 1.0        | 1.1                | 0.0        | 1.2        | 1.1        |
| shared backgrounds            | 0.0        | 0.2        | -0.0       | 0.0        | -0.7               | 0.5        | 0.0        | 0.2        |
| remaining Bias                | -1.4       | -0.9       | 6.2        | -1.8       | 0.2                | 0.2        | -2.8       | -0.3       |
| quadratic sum                 | 27.6       | 20.8       | 12.5       | 2.0        | 1.3                | 0.5        | 3.0        | 1.1        |
| quadratic sum, absolute value | 0.012      | 0.013      | 0.008      | 0.001      | 0.001              | 0.000      | 0.002      | 0.001      |
| absolute expected statistical | 0.043      | 0.065      | 0.063      | 0.062      | 0.045              | 0.067      | 0.069      | 0.065      |

Table 8.19: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [1.1, 6.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | -0.3       | -0.5       | 0.2        | -0.2       | -0.3               | -0.2       | -0.0       | 0.0        |
| signal mass model             | 6.2        | 1.4        | -3.0       | -0.0       | 0.3                | 0.4        | 0.0        | -0.2       |
| angular background model      | 5.6        | 0.9        | -3.2       | 0.0        | 0.3                | 0.2        | 0.0        | -0.2       |
| acceptance order              | -0.3       | 0.2        | -0.0       | -0.0       | 0.0                | 0.0        | -0.0       | 0.0        |
| simulation correction         | 9.6        | 1.2        | 0.7        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| PID correction                | 1.1        | 0.9        | 0.0        | 0.0        | 0.0                | 0.0        | 0.0        | 0.0        |
| simulation statistics         | -0.3       | -0.2       | 0.0        | -0.0       | 0.0                | 0.0        | 0.0        | -0.0       |
| S-wave                        | -17.7      | 0.5        | 3.7        | -0.2       | -0.0               | 0.7        | 0.2        | 0.5        |
| lifetime mismodelling         | 25.0       | -10.7      | 2.4        | -0.0       | 0.0                | 0.0        | -0.0       | 0.0        |
| time integration              | 17.1       | 9.0        | 1.7        | 0.6        | 0.3                | 0.2        | 0.7        | 0.7        |
| shared backgrounds            | 0.0        | -0.0       | -0.0       | -0.0       | -0.0               | -0.2       | -0.0       | -0.2       |
| remaining Bias                | -3.7       | -2.1       | 3.4        | -0.9       | -3.5               | -3.3       | -1.5       | 1.6        |
| quadratic sum                 | 37.5       | 14.3       | 7.3        | 1.1        | 3.5                | 3.4        | 1.7        | 1.9        |
| quadratic sum, absolute value | 0.013      | 0.006      | 0.004      | 0.001      | 0.001              | 0.002      | 0.001      | 0.001      |
| absolute expected statistical | 0.036      | 0.043      | 0.059      | 0.054      | 0.034              | 0.055      | 0.060      | 0.043      |

Table 8.20: Systematic uncertainties relative to the statistical uncertainty in % in the region  $q^2 \in [15.0, 18.9]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|-------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| peaking backgrounds           | -4.2       | -0.0       | 0.2        | 0.2        | -0.4               | -0.2       | -0.0       | -0.3       |
| signal mass model             | -16.7      | 1.3        | -0.2       | -0.0       | 0.0                | -0.0       | -0.0       | -0.0       |
| angular background model      | 15.3       | -3.6       | 3.1        | 0.0        | 0.0                | -0.0       | -0.0       | 0.0        |
| acceptance order              | -31.7      | -11.5      | 0.6        | 1.1        | -2.7               | 0.8        | 2.2        | -1.3       |
| simulation correction         | 43.6       | 4.4        | 0.4        | 1.1        | 1.2                | 2.7        | 1.0        | 0.8        |
| PID correction                | 6.6        | 5.2        | 1.0        | 0.6        | 1.5                | 0.4        | 0.6        | 0.3        |
| simulation statistics         | -19.9      | -3.9       | 0.2        | 0.2        | -1.2               | -0.2       | -0.0       | 0.0        |
| S-wave                        | 0.3        | 14.1       | 10.2       | 0.2        | -0.0               | -0.2       | -0.0       | -0.5       |
| lifetime mismodelling         | 19.9       | -21.6      | 1.9        | 0.0        | 0.0                | 0.2        | -0.2       | -0.0       |
| time integration              | 12.2       | 18.8       | 6.4        | 0.8        | 1.9                | 0.0        | 1.8        | 2.3        |
| shared backgrounds            | -0.0       | -0.0       | 0.0        | -0.0       | 0.8                | -0.2       | -0.0       | -0.0       |
| remaining Bias                | 9.8        | 3.9        | 0.6        | -1.5       | -1.2               | -2.3       | -1.2       | -0.3       |
| quadratic sum                 | 67.2       | 35.2       | 12.7       | 2.4        | 4.3                | 3.7        | 3.4        | 2.9        |
| quadratic sum, absolute value | 0.019      | 0.014      | 0.006      | 0.001      | 0.001              | 0.002      | 0.002      | 0.001      |
| absolute expected statistical | 0.029      | 0.038      | 0.048      | 0.048      | 0.026              | 0.048      | 0.049      | 0.038      |

Table 8.21: Absolute values for the quadratic sum of the systematic uncertainties.

| $q^2$ region                    | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{FB}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
|---------------------------------|-------|-------|-------|-------|---------------|-------|-------|-------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | 0.017 | 0.014 | 0.005 | 0.001 | 0.009         | 0.001 | 0.004 | 0.001 |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | 0.015 | 0.004 | 0.002 | 0.002 | 0.001         | 0.001 | 0.001 | 0.001 |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | 0.016 | 0.012 | 0.010 | 0.003 | 0.002         | 0.003 | 0.002 | 0.001 |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | 0.012 | 0.009 | 0.006 | 0.001 | 0.001         | 0.002 | 0.002 | 0.001 |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 0.012 | 0.013 | 0.008 | 0.001 | 0.001         | 0.000 | 0.002 | 0.001 |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 0.013 | 0.006 | 0.004 | 0.001 | 0.001         | 0.002 | 0.001 | 0.001 |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 0.019 | 0.014 | 0.006 | 0.001 | 0.001         | 0.002 | 0.002 | 0.001 |

Table 8.22: Values for the quadratic sum of the systematic uncertainties relative to the expected statistical uncertainty.

| $q^2$ bin                       | $F_L$ in % | $S_3$ in % | $S_4$ in % | $A_5$ in % | $A_{FB}^{CP}$ in % | $S_7$ in % | $A_8$ in % | $A_9$ in % |
|---------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| [0.1, 0.98] $\text{GeV}^2/c^4$  | 34.8       | 20.5       | 6.6        | 0.8        | 14.5               | 1.3        | 5.0        | 2.0        |
| [1.1, 4.0] $\text{GeV}^2/c^4$   | 32.1       | 8.1        | 2.3        | 3.0        | 1.2                | 1.2        | 1.5        | 1.8        |
| [4.0, 6.0] $\text{GeV}^2/c^4$   | 32.2       | 18.4       | 12.8       | 4.2        | 4.4                | 3.5        | 1.9        | 2.2        |
| [6.0, 8.0] $\text{GeV}^2/c^4$   | 25.3       | 13.7       | 7.5        | 1.8        | 2.4                | 2.5        | 2.3        | 1.1        |
| [11.0, 12.5] $\text{GeV}^2/c^4$ | 27.6       | 20.8       | 12.5       | 2.0        | 1.3                | 0.5        | 3.0        | 1.1        |
| [1.1, 6.0] $\text{GeV}^2/c^4$   | 37.5       | 14.3       | 7.3        | 1.1        | 3.5                | 3.4        | 1.7        | 1.9        |
| [15.0, 18.9] $\text{GeV}^2/c^4$ | 67.2       | 35.2       | 12.7       | 2.4        | 4.3                | 3.7        | 3.4        | 2.9        |



# Chapter 9

## Results for the angular observables

In this Chapter, the results of the measurement of the angular observables ( $F_L$ ,  $S_{3,4,7}$ ,  $A_{\text{FB}}^{CP}$  and  $A_{5,8,9}$ ) introduced in Sec. 2.3.4 for the rare mode  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  are given. The angular observables are extracted in several  $q^2$  regions using an unbinned maximum likelihood fit to LHCb data taken during Run 1 and Run 2 of the LHC as described in Chap. 7. The statistical uncertainties on the angular observables are corrected for over- and undercoverage as described in Sec. 7.3. The systematic uncertainties have been studied and described in Chap. 8.

The results for the angular observables are given numerically and graphically in Sec. 9.1 and the correlations are given in Sec. 9.2. Finally, the compatibility with the SM is tested by varying the Wilson coefficient  $\mathcal{R}e(\mathcal{C}_9^{\text{NP}})$  in Sec. 9.3.

### 9.1 Results for angular observables

The four-dimensional unbinned maximum likelihood fit to the three decay angles and the reconstructed mass of the  $B_s^0$  candidates has been performed according to the fit model introduced in Sec. 7.2. The fit describes the data well as can be seen in the fit projections given in Figs. F.1-F.21 in App. F and shows no sign of systematic deviations. The numerical values for the  $CP$ -averages ( $F_L$  and  $S_{3,4,7}$ ) and  $CP$ -asymmetries ( $A_{\text{FB}}^{CP}$  and  $A_{5,8,9}$ ) are given in Tab. 9.1, where the first uncertainty is always of statistical nature and the second is the total systematic uncertainty. The results are statistically limited for all  $q^2$  regions and observables. A visual representation of the results for the angular observables is given in Fig. 9.1 for each of the  $q^2$  region except for the region of  $[1.1, 6.0] \text{ GeV}^2/c^4$ . In this figure the results of this measurement [46] are given as black crosses and are overlaid with the result from the previously published LHCb measurement [15] as grey crosses and the SM predictions from Refs. [109, 243–245], obtained from the **flavio** software package [109], as blue boxes, where available. The light grey bands indicate the regions of the charmonium resonances and the  $B_s^0 \rightarrow (\phi \rightarrow K^+ K^-)(\phi \rightarrow \mu^+ \mu^-)$  region, which are removed from this analysis.

The  $CP$ -averages are generally consistent with the SM predictions provided by the **flavio** software package. The largest local deviation are of  $\mathcal{O}(2.1 \sigma)$  of local significance. The largest deviation is found in  $S_7$  in the  $q^2$  region of  $[0.1, 0.98] \text{ GeV}^2/c^4$  with a local significance of  $-2.1 \sigma$ . While  $S_{3,4,7}$  do not show strong signs of a systematic shift from the SM prediction, the measured values of the angular observable  $F_L$  consistently lie below the prediction in the three narrow  $q^2$  regions in the range of  $[0.1, 6.0] \text{ GeV}^2/c^4$ . The global agreement with the SM is determined in Sec. 9.3. The  $CP$ -asymmetries are compatible with the SM prediction of zero in all  $q^2$  regions,

with the largest deviation being a shift of  $-2.1\sigma$  found in the observable  $A_8$  in the  $q^2$  region of  $[6.0, 8.0] \text{ GeV}^2/c^4$ . This measurement constitutes the most precise measurement of the angular observables in  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decays to date.

Table 9.1:  $CP$ -averages  $F_L$  and  $S_{3,4,7}$  and  $CP$ -asymmetries  $A_{\text{FB}}^{CP}$  and  $A_{5,8,9}$  obtained from the maximum likelihood fit. The first uncertainty is statistical and the second is the total systematic uncertainty, which was described in Chap. 8. This table also appeared in Ref. [46].

| $q^2 [\text{GeV}^2/c^4]$ | $F_L$                       | $S_3$                        | $S_4$                        | $S_7$                        |
|--------------------------|-----------------------------|------------------------------|------------------------------|------------------------------|
| [0.1, 0.98]              | $0.254 \pm 0.045 \pm 0.017$ | $-0.004 \pm 0.068 \pm 0.014$ | $0.213 \pm 0.082 \pm 0.005$  | $-0.178 \pm 0.072 \pm 0.001$ |
| [1.1, 4.0]               | $0.723 \pm 0.053 \pm 0.015$ | $-0.030 \pm 0.057 \pm 0.004$ | $-0.110 \pm 0.079 \pm 0.002$ | $-0.101 \pm 0.075 \pm 0.001$ |
| [4.0, 6.0]               | $0.701 \pm 0.050 \pm 0.016$ | $-0.162 \pm 0.067 \pm 0.012$ | $-0.222 \pm 0.092 \pm 0.010$ | $0.175 \pm 0.089 \pm 0.003$  |
| [6.0, 8.0]               | $0.624 \pm 0.051 \pm 0.012$ | $0.013 \pm 0.080 \pm 0.009$  | $-0.176 \pm 0.078 \pm 0.006$ | $0.033 \pm 0.081 \pm 0.002$  |
| [11.0, 12.5]             | $0.353 \pm 0.044 \pm 0.012$ | $-0.138 \pm 0.071 \pm 0.013$ | $-0.319 \pm 0.061 \pm 0.008$ | $-0.170 \pm 0.069 \pm 0.000$ |
| [1.1, 6.0]               | $0.715 \pm 0.036 \pm 0.013$ | $-0.083 \pm 0.047 \pm 0.006$ | $-0.155 \pm 0.058 \pm 0.004$ | $0.020 \pm 0.059 \pm 0.001$  |
| [15.0, 18.9]             | $0.359 \pm 0.031 \pm 0.019$ | $-0.247 \pm 0.042 \pm 0.014$ | $-0.208 \pm 0.047 \pm 0.006$ | $0.003 \pm 0.046 \pm 0.002$  |

| $q^2 [\text{GeV}^2/c^4]$ | $A_5$                        | $A_{\text{FB}}^{CP}$         | $A_8$                        | $A_9$                        |
|--------------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| [0.1, 0.98]              | $0.043 \pm 0.067 \pm 0.001$  | $0.068 \pm 0.064 \pm 0.009$  | $-0.007 \pm 0.073 \pm 0.004$ | $-0.030 \pm 0.079 \pm 0.001$ |
| [1.1, 4.0]               | $0.026 \pm 0.067 \pm 0.002$  | $0.023 \pm 0.054 \pm 0.001$  | $0.038 \pm 0.082 \pm 0.001$  | $0.020 \pm 0.068 \pm 0.001$  |
| [4.0, 6.0]               | $-0.084 \pm 0.084 \pm 0.003$ | $-0.030 \pm 0.051 \pm 0.002$ | $0.012 \pm 0.090 \pm 0.002$  | $-0.008 \pm 0.061 \pm 0.001$ |
| [6.0, 8.0]               | $-0.022 \pm 0.082 \pm 0.001$ | $0.032 \pm 0.049 \pm 0.001$  | $-0.170 \pm 0.080 \pm 0.002$ | $-0.012 \pm 0.090 \pm 0.001$ |
| [11.0, 12.5]             | $0.035 \pm 0.063 \pm 0.001$  | $0.034 \pm 0.048 \pm 0.001$  | $0.046 \pm 0.070 \pm 0.002$  | $0.017 \pm 0.071 \pm 0.001$  |
| [1.1, 6.0]               | $-0.007 \pm 0.051 \pm 0.001$ | $0.006 \pm 0.036 \pm 0.001$  | $0.016 \pm 0.062 \pm 0.001$  | $0.009 \pm 0.046 \pm 0.001$  |
| [15.0, 18.9]             | $-0.025 \pm 0.043 \pm 0.001$ | $-0.011 \pm 0.033 \pm 0.001$ | $0.072 \pm 0.051 \pm 0.002$  | $0.021 \pm 0.042 \pm 0.001$  |

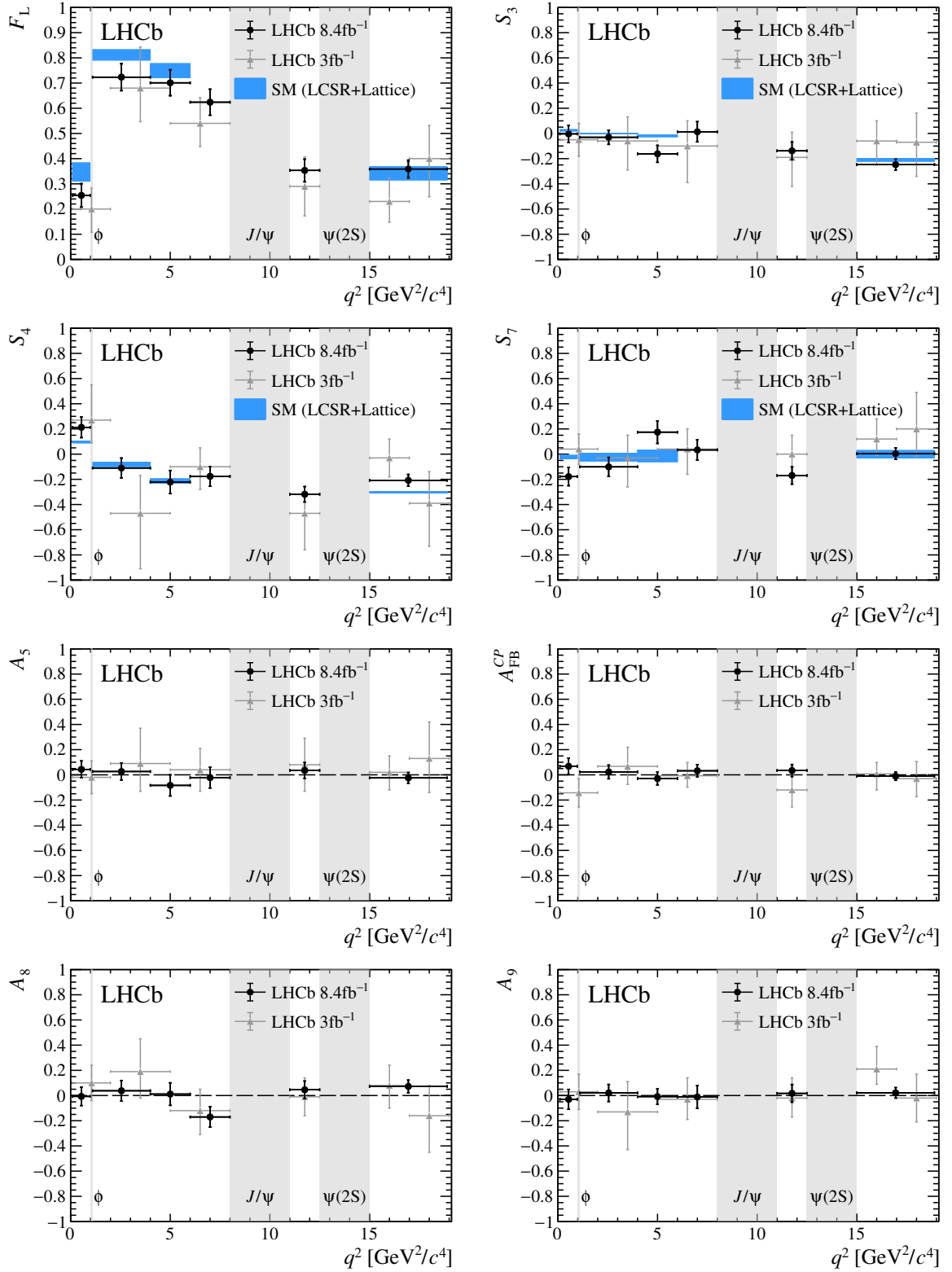


Figure 9.1:  $CP$ -averaged angular observables  $F_L$  and  $S_{3,4,7}$  and  $CP$ -asymmetries  $A_{\text{FB}}^{CP}$  and  $A_{5,8,9}$  shown by black crosses, overlaid with the SM prediction [109, 243–245] as blue boxes, where available. The grey crosses indicate the results from Ref. [15]. The light grey bands indicate the regions of the charmonium resonances and the  $B_s^0 \rightarrow \phi\phi$  region. This figure also appeared in Ref. [46].

## 9.2 Correlation matrices

In Tabs. 9.2-9.4 the correlation matrices for the angular observables in each of the  $q^2$  regions are given. The angular observables generally show little correlation with each other, typically at the level of 10% or lower. The largest correlation coefficient is found to be 29% in the  $q^2$  region of  $0.1 < q^2 < 0.98 \text{ GeV}^2/c^4$  between  $S_7$  and  $A_8$ . The correlations are an important input when combining angular observables in a global fit. These global fits are often used to determine the shift of Wilson coefficients, such as the shift in the real part of the muonic vector coupling ( $\Delta \mathcal{R}e(\mathcal{C}_9^{\text{NP}})$ ).

Table 9.2: Correlation matrix for the  $q^2$  regions  $[0.1, 0.98] \text{ GeV}^2/c^4$  and  $[1.1, 4.0] \text{ GeV}^2/c^4$ . This table also appeared in Ref. [46].

| Correlation matrix for $0.1 < q^2 < 0.98 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|-------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                             | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                       | 1.00  | -0.03 | 0.06  | 0.10  | 0.02                 | -0.20 | -0.05 | -0.04 |
| $S_3$                                                       |       | 1.00  | 0.11  | 0.00  | -0.06                | 0.07  | -0.05 | 0.02  |
| $S_4$                                                       |       |       | 1.00  | 0.03  | -0.06                | 0.08  | -0.05 | -0.01 |
| $A_5$                                                       |       |       |       | 1.00  | 0.12                 | -0.07 | 0.06  | -0.09 |
| $A_{\text{FB}}^{CP}$                                        |       |       |       |       | 1.00                 | 0.02  | -0.04 | 0.07  |
| $S_7$                                                       |       |       |       |       |                      | 1.00  | 0.29  | -0.03 |
| $A_8$                                                       |       |       |       |       |                      |       | 1.00  | 0.06  |
| $A_9$                                                       |       |       |       |       |                      |       |       | 1.00  |

| Correlation matrix for $1.1 < q^2 < 4.0 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                            | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                      | 1.00  | -0.12 | 0.03  | -0.03 | -0.07                | -0.12 | 0.03  | -0.01 |
| $S_3$                                                      |       | 1.00  | -0.07 | -0.00 | 0.03                 | 0.04  | -0.04 | 0.05  |
| $S_4$                                                      |       |       | 1.00  | 0.02  | 0.09                 | 0.01  | -0.02 | -0.02 |
| $A_5$                                                      |       |       |       | 1.00  | -0.08                | 0.02  | 0.01  | 0.06  |
| $A_{\text{FB}}^{CP}$                                       |       |       |       |       | 1.00                 | 0.02  | -0.02 | 0.03  |
| $S_7$                                                      |       |       |       |       |                      | 1.00  | -0.06 | 0.06  |
| $A_8$                                                      |       |       |       |       |                      |       | 1.00  | -0.04 |
| $A_9$                                                      |       |       |       |       |                      |       |       | 1.00  |

Table 9.3: Correlation matrix for the  $q^2$  regions  $[4.0, 6.0] \text{ GeV}^2/c^4$ ,  $[6.0, 8.0] \text{ GeV}^2/c^4$  and  $[11.0, 12.5] \text{ GeV}^2/c^4$ .

| Correlation matrix for $4.0 < q^2 < 6.0 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                            | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                      | 1.00  | 0.15  | 0.06  | -0.08 | 0.01                 | -0.01 | -0.04 | -0.08 |
| $S_3$                                                      |       | 1.00  | -0.04 | -0.02 | 0.05                 | -0.06 | -0.04 | 0.22  |
| $S_4$                                                      |       |       | 1.00  | -0.12 | -0.02                | 0.05  | -0.04 | -0.07 |
| $A_5$                                                      |       |       |       | 1.00  | -0.11                | -0.06 | 0.04  | 0.05  |
| $A_{\text{FB}}^{CP}$                                       |       |       |       |       | 1.00                 | -0.01 | 0.10  | 0.01  |
| $S_7$                                                      |       |       |       |       |                      | 1.00  | -0.13 | 0.02  |
| $A_8$                                                      |       |       |       |       |                      |       | 1.00  | -0.05 |
| $A_9$                                                      |       |       |       |       |                      |       |       | 1.00  |

| Correlation matrix for $6.0 < q^2 < 8.0 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                            | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                      | 1.00  | 0.03  | 0.07  | -0.04 | -0.10                | 0.01  | -0.03 | -0.04 |
| $S_3$                                                      |       | 1.00  | -0.08 | -0.02 | -0.02                | -0.01 | 0.04  | -0.05 |
| $S_4$                                                      |       |       | 1.00  | -0.09 | 0.02                 | -0.05 | -0.08 | -0.06 |
| $A_5$                                                      |       |       |       | 1.00  | -0.12                | -0.05 | -0.05 | 0.05  |
| $A_{\text{FB}}^{CP}$                                       |       |       |       |       | 1.00                 | -0.04 | -0.04 | -0.01 |
| $S_7$                                                      |       |       |       |       |                      | 1.00  | -0.14 | 0.03  |
| $A_8$                                                      |       |       |       |       |                      |       | 1.00  | -0.05 |
| $A_9$                                                      |       |       |       |       |                      |       |       | 1.00  |

| Correlation matrix for $11.0 < q^2 < 12.5 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|--------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                              | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                        | 1.00  | 0.09  | 0.02  | -0.05 | -0.10                | -0.03 | -0.04 | -0.07 |
| $S_3$                                                        |       | 1.00  | -0.14 | -0.01 | -0.02                | 0.11  | -0.06 | -0.06 |
| $S_4$                                                        |       |       | 1.00  | -0.04 | 0.18                 | -0.01 | -0.07 | -0.05 |
| $A_5$                                                        |       |       |       | 1.00  | -0.23                | -0.11 | 0.01  | -0.02 |
| $A_{\text{FB}}^{CP}$                                         |       |       |       |       | 1.00                 | 0.04  | -0.06 | 0.02  |
| $S_7$                                                        |       |       |       |       |                      | 1.00  | 0.06  | -0.05 |
| $A_8$                                                        |       |       |       |       |                      |       | 1.00  | -0.11 |
| $A_9$                                                        |       |       |       |       |                      |       |       | 1.00  |

Table 9.4: Correlation matrix for the  $q^2$  region  $[1.1, 6.0] \text{ GeV}^2/c^4$  and  $[15.0, 18.9] \text{ GeV}^2/c^4$ . This table also appeared in Ref. [46].

| Correlation matrix for $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                            | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                      | 1.00  | -0.03 | 0.05  | -0.02 | -0.07                | -0.08 | -0.04 | -0.03 |
| $S_3$                                                      |       | 1.00  | -0.07 | 0.01  | 0.03                 | -0.07 | -0.07 | 0.10  |
| $S_4$                                                      |       |       | 1.00  | -0.00 | -0.00                | -0.06 | -0.02 | -0.01 |
| $A_5$                                                      |       |       |       | 1.00  | -0.07                | -0.00 | 0.05  | 0.08  |
| $A_{\text{FB}}^{CP}$                                       |       |       |       |       | 1.00                 | 0.01  | 0.05  | 0.07  |
| $S_7$                                                      |       |       |       |       |                      | 1.00  | -0.08 | 0.01  |
| $A_8$                                                      |       |       |       |       |                      |       | 1.00  | -0.03 |
| $A_9$                                                      |       |       |       |       |                      |       |       | 1.00  |

| Correlation matrix for $15.0 < q^2 < 18.9 \text{ GeV}^2/c^4$ |       |       |       |       |                      |       |       |       |
|--------------------------------------------------------------|-------|-------|-------|-------|----------------------|-------|-------|-------|
|                                                              | $F_L$ | $S_3$ | $S_4$ | $A_5$ | $A_{\text{FB}}^{CP}$ | $S_7$ | $A_8$ | $A_9$ |
| $F_L$                                                        | 1.00  | 0.20  | -0.04 | -0.03 | 0.03                 | 0.00  | -0.03 | -0.10 |
| $S_3$                                                        |       | 1.00  | -0.06 | 0.03  | -0.11                | -0.08 | 0.00  | 0.13  |
| $S_4$                                                        |       |       | 1.00  | -0.13 | -0.03                | -0.04 | 0.13  | 0.06  |
| $A_5$                                                        |       |       |       | 1.00  | -0.11                | 0.05  | -0.08 | 0.05  |
| $A_{\text{FB}}^{CP}$                                         |       |       |       |       | 1.00                 | 0.10  | -0.00 | -0.03 |
| $S_7$                                                        |       |       |       |       |                      | 1.00  | 0.03  | 0.01  |
| $A_8$                                                        |       |       |       |       |                      |       | 1.00  | -0.07 |
| $A_9$                                                        |       |       |       |       |                      |       |       | 1.00  |

### 9.3 Compatibility of the $CP$ -averages with the SM

The  $CP$ -averages measured in Sec. 9.1 are generally consistent with the SM predictions at the level of  $2\sigma$  local significance. To determine the global agreement of the measured angular observables with the SM prediction, a global fit to the  $CP$ -averaged angular observables is performed using the  $q^2$  regions  $[0.1, 0.98] \text{ GeV}^2/c^4$ ,  $[1.1, 4.0] \text{ GeV}^2/c^4$ ,  $[4.0, 6.0] \text{ GeV}^2/c^4$  and  $[15.0, 18.9] \text{ GeV}^2/c^4$ . The fit is performed with the `flavio` software package [109] and is used to determine a shift in the real part of the Wilson coefficient  $\mathcal{C}_9$  ( $\Delta\mathcal{R}e(\mathcal{C}_9)$ ) which describes the vector coupling to the muons. The  $CP$ -asymmetries are omitted from the fit given that the SM prediction is not implemented in the `flavio` software package. The  $q^2$  regions  $[6.0, 8.0] \text{ GeV}^2/c^4$  and  $[11.0, 12.5] \text{ GeV}^2/c^4$  are not taken into account in this global fit, as they are sensitive to long-distance effects from charmonium resonances which are not fully understood and currently under discussion [103–107]. The correlations between the  $CP$ -averages in a  $q^2$  region given in Sec. 9.2 are taken into account in the likelihood used to the global fit. The logarithmic likelihood is defined as [109]

$$-2 \log \mathcal{L} = \sum_i \left( \vec{O}_i^{\text{exp}} - \vec{O}_i^{\text{th}}(\vec{\mathcal{C}}, \vec{\theta}_0) \right)^T \left[ C_{\text{exp}} + C_{\text{th}}(\vec{\theta}_0) \right]^{-1} \left( \vec{O}_i^{\text{exp}} - \vec{O}_i^{\text{th}}(\vec{\mathcal{C}}, \vec{\theta}_0) \right),$$

where  $\vec{O}_i^{\text{exp(th)}}$  indicates the experimental measurement (theoretical prediction) for the observables,  $\vec{\theta}_0$  indicates the set of central values of the theory parameters,  $C_{\text{exp(th)}}$  are the covariance matrices for the experimental measurements (theoretical prediction) and  $\vec{\mathcal{C}}$  indicates the Wilson coefficients. The resulting logarithmic likelihood profile ( $-\Delta \log \mathcal{L}$ ) is shown in Fig. 9.2 as an orange line. The best fit value is found at the minimum of the negative logarithmic likelihood to be  $\Delta\mathcal{R}e(\mathcal{C}_9) = -1.3^{+0.7}_{-0.6}$  which is consistent with the SM value of  $\Delta\mathcal{R}e(\mathcal{C}_9) = 0$  at the level of  $1.9\sigma$ , assuming the uncertainties follow that of a Gaussian distribution. The given uncertainties are taken as the difference of the best fit value to the values at  $-2\Delta \log \mathcal{L} = 1$ . Consistent shifts in  $\Delta\mathcal{R}e(\mathcal{C}_9)$  have been observed in other angular analyses of  $b \rightarrow s\ell^+\ell^-$  transitions performed at LHCb, namely the angular analyses of  $B^+ \rightarrow K^{*+}\mu^+\mu^-$  [21] and  $B^0 \rightarrow K^{*0}\mu^+\mu^-$  [22]. This is illustrated in Fig. 9.3. While it is intriguing that the shift has consistently been measured for different spectator quarks, it does not constitute a strong sign of NP, yet. This is due to the fact that it is not clear if this is a genuine NP effect or induced by long-distance effects from the charmonium resonances, which can not theoretically be calculated from first principles in the SM [56]. The effect of the long-distance effects from the charmonium resonances is a topic of discussion [103–107].

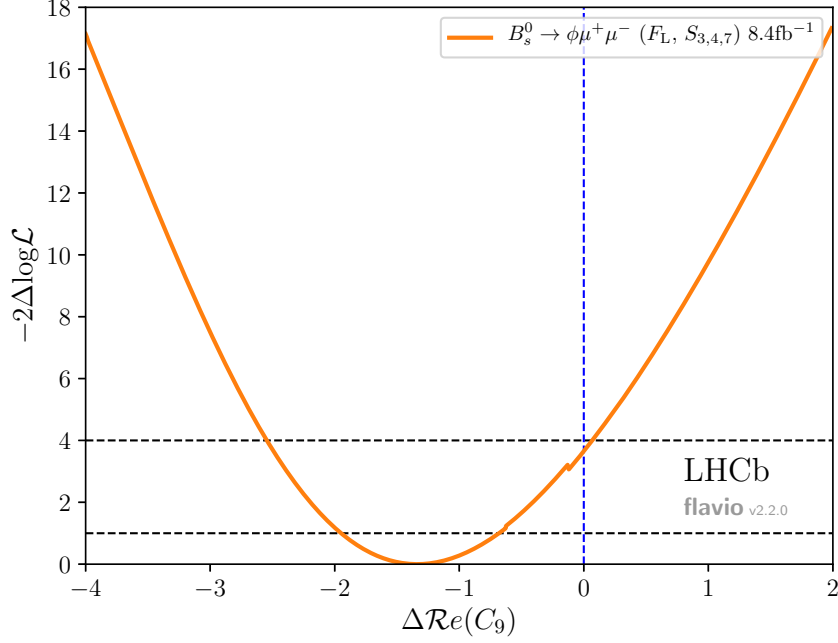


Figure 9.2: Negative logarithmic likelihood for the scan over  $\Delta\mathcal{R}e(C_9)$ . The fit uses the  $CP$ -averaged angular observables in the  $q^2$  regions  $[0.1, 0.98] \text{ GeV}^2/c^4$ ,  $[1.1, 4.0] \text{ GeV}^2/c^4$ ,  $[4.0, 6.0] \text{ GeV}^2/c^4$  and  $[15.0, 18.9] \text{ GeV}^2/c^4$ . This figure also appeared in Ref. [46].

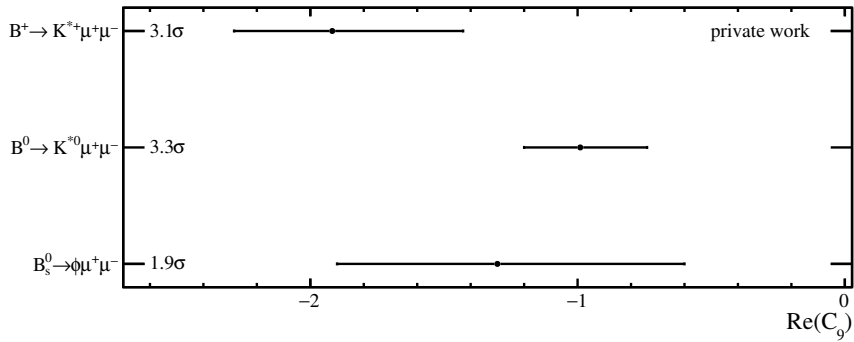


Figure 9.3: Comparisons of global fit results to angular analyses of  $b \rightarrow s\ell^+\ell^-$  transitions at LHCb [21, 22].



# Chapter 10

## Summary and outlook

This thesis presents an angular analysis of the rare decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  using proton-proton collision data corresponding to an integrated luminosity of  $8.4 \text{ fb}^{-1}$  collected with the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV during the LHC Run 1 and Run 2. The  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay mode is an especially interesting probe of the SM due to the narrow  $\phi$  resonance. The narrow  $\phi$  resonance allows for a particularly clean selection with very little background contributions and small contribution from the  $S$ -wave configuration of the  $K^+ K^- \mu^+ \mu^-$  final compared to  $B \rightarrow K^* \mu^+ \mu^-$  decays. The angular observables are measured in regions of the invariant dimuon mass squared ( $q^2$ ) and correspond to either  $CP$ -averaged ( $F_L$ ,  $S_{3,4,7}$ ) or  $CP$ -asymmetric ( $A_{\text{FB}}^{CP}$ ,  $A_{5,8,9}$ ) observables. The measured values for these observables, along with their associated statistical and systematic uncertainty are given in Chap. 9.1, where they are also displayed graphically alongside the SM predictions. Each of the  $CP$ -averages is consistent with the considered SM prediction [109, 243–245], typically at the level of  $2\sigma$  and each of the  $CP$ -asymmetries is consistent with the SM prediction of zero, typically at the level of  $1\sigma$ . The largest local deviations for the  $CP$ -averages and  $CP$ -asymmetries are found in  $S_7$  in the  $q^2$  region of  $[0.1, 0.98] \text{ GeV}^2/c^4$  with  $-2.1\sigma$  and in  $A_8$  in the  $q^2$  region of  $[6.0, 8.0] \text{ GeV}^2/c^4$  with  $-2.1\sigma$ , respectively.

The global compatibility of the  $CP$ -averages with the SM across the  $q^2$  regions has been evaluated with a fit of the shift in the muonic vector coupling ( $\Delta \mathcal{R}e(\mathcal{C}_9)$ ) using the `flavio` software package. The shift was found to be  $\Delta \mathcal{R}e(\mathcal{C}_9) = -1.3_{-0.6}^{+0.7}$  which corresponds to a global agreement at the level of  $1.9\sigma$ , assuming the uncertainty follows that of a Gaussian distribution. While the results agree with the SM, they show the same trend of a downwards shift in  $\Delta \mathcal{R}e(\mathcal{C}_9)$  as seen in other angular analyses of muonic final states [21, 22]. As such additional data taken in future runs of the LHC will be crucial to be able to draw conclusions regarding NP scenarios.

This measurement of the angular observables represents the most precise measurement to date with the finest partition of the low  $q^2$  region yet, and, therefore, supersedes the previous measurement performed at LHCb [15]. The increased statistical power and the ability to use the finer  $q^2$  partition are driven by an improved selection alongside the addition of Run 2 data which adds nearly a four-fold increase in  $B_s^0$  mesons compared to Run 1 data due to the additional luminosity and the higher center-of-mass energy.

## Outlook

A future update with data taken in Runs 3 and 4 of the LHC with the upgraded LHCb detector [246,247] is, therefore, eagerly expected. By 2030 an integrated luminosity of  $50 \text{ fb}^{-1}$  is expected which would enlarge the sample size by more than a factor of 6 assuming the same efficiencies for the new data set. The enlarged data set could be used to perform the most precise measurement of the angular observables. With the increase in statistics, an untagged, time-dependent angular analysis could be performed to measure the effective lifetime of the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  for the first time. This is a crucial step towards a full time-dependent angular analysis [248] which additionally requires flavour-tagging of the  $B_s^0$  and  $\bar{B}_s^0$  mesons. Given the increased size of the Run 3 data set and assuming an effective tagging power of 5%, the full time-dependent angular analysis is expected to be able to be performed. The full time-dependent angular analysis would allow for the measurement of the  $CP$ -averages  $S_5$ ,  $S_6^{s,c}$ ,  $S_8$  and  $S_9$ , the  $CP$ -asymmetries  $A_1^{s,c}$ ,  $A_3$ ,  $A_4$  and  $A_7$  as well as the  $h_i$  and  $s_i$  observables introduced in Ref. [57]. Considering that branching fraction of the  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  decay has only been measured at D0 [249], CDF [250] and LHCb [11,15,24] and LHCb is the only experiment to perform angular analyses of this decay [11,15,46], both, the improved statistical power and the additional avenues of probing for New Physics are interesting prospects.



# Appendix A

## Definitions for additional time-dependent angular observables

In this Appendix, the angular observables related to the time-dependent angular distribution  $h_i$  and  $s_i$  are defined following Ref. [57]. The new angular distributions depend on the transversity amplitudes introduced in Sec. 2.3.3. Here, the scalar amplitude  $A_S$  is also taken into account when defining the  $h_i$  and  $s_i$  terms. Furthermore, the notation  $\tilde{A}_X$  denotes the amplitude of the  $\bar{B}_s^0$  meson decaying into the final state without applying the  $CP$ -conjugation ( $\tilde{A} = \eta_X \bar{A}_X$ , where  $\bar{A}$  denotes the amplitude of the  $\bar{B}_s^0$  meson and can be obtained by changing the signs of all weak phases and  $\eta_X$  is the  $CP$ -parity of the amplitude).

The  $h_i$  are connected to the real part  $\mathcal{Re}$  of combinations of the transversity amplitudes and given by

$$\begin{aligned}
h_{1s} &= \frac{2 + \beta_\ell^2}{2} \mathcal{Re}[e^{i\phi} \{\tilde{A}_\perp^L A_\perp^{L*} + \tilde{A}_\parallel^L A_\parallel^{L*} + \tilde{A}_\perp^R A_\perp^{R*} + \tilde{A}_\parallel^R A_\parallel^{R*}\}] \\
&\quad + \frac{4m_\ell^2}{q^2} \mathcal{Re}[e^{i\phi} \{\tilde{A}_\perp^L A_\perp^{R*} + \tilde{A}_\parallel^L A_\parallel^{R*}\} + e^{-i\phi} \{A_\perp^L \tilde{A}_\perp^{R*} + A_\parallel^L \tilde{A}_\parallel^{R*}\}] \\
h_{1c} &= 2\mathcal{Re}[e^{i\phi} \{\tilde{A}_0^L A_0^{L*} + \tilde{A}_0^R A_0^{R*}\}] \\
&\quad + \frac{8m_\ell^2}{q^2} [\mathcal{Re}[e^{i\phi} \tilde{A}_t A_t^*] + \mathcal{Re}\{e^{i\phi} \tilde{A}_0^L A_0^{R*} + e^{-i\phi} A_0^L \tilde{A}_0^{R*}\}] + 2\beta_\ell^2 \mathcal{Re}[e^{i\phi} \{\tilde{A}_S A_S^*\}] \\
h_{2s} &= \frac{\beta_\ell^2}{2} \mathcal{Re}[e^{i\phi} \{\tilde{A}_\perp^L A_\perp^{L*} + \tilde{A}_\parallel^L A_\parallel^{L*} + \tilde{A}_\perp^R A_\perp^{R*} + \tilde{A}_\parallel^R A_\parallel^{R*}\}] \\
h_{2c} &= -2\beta_\ell^2 \mathcal{Re}[e^{i\phi} \{\tilde{A}_0^L A_0^{L*} + \tilde{A}_0^R A_0^{R*}\}] \\
h_3 &= \beta_\ell^2 \mathcal{Re}[e^{i\phi} \{\tilde{A}_\perp^L A_\perp^{L*} - \tilde{A}_\parallel^L A_\parallel^{L*} + \tilde{A}_\perp^R A_\perp^{R*} - \tilde{A}_\parallel^R A_\parallel^{R*}\}] \\
h_4 &= \frac{1}{\sqrt{2}} \beta_\ell^2 \mathcal{Re}[e^{i\phi} \{\tilde{A}_0^L A_\parallel^{L*} + \tilde{A}_0^R A_\parallel^{R*}\} + e^{-i\phi} \{A_0^L \tilde{A}_\parallel^{L*} + A_0^R \tilde{A}_\parallel^{R*}\}] \\
h_5 &= \sqrt{2} \beta_\ell \left[ \mathcal{Re}[e^{i\phi} \{\tilde{A}_0^L A_\perp^{L*} - \tilde{A}_0^R A_\perp^{R*}\} + e^{-i\phi} \{A_0^L \tilde{A}_\perp^{L*} - A_0^R \tilde{A}_\perp^{R*}\}] \right. \\
&\quad \left. - \frac{m_\ell}{\sqrt{q^2}} \mathcal{Re}[e^{i\phi} \{\tilde{A}_\parallel^L A_S^* + \tilde{A}_\parallel^R A_S^*\} + e^{-i\phi} \{A_\parallel^L \tilde{A}_S^* + A_\parallel^R \tilde{A}_S^*\}] \right] \\
h_{6s} &= 2\beta_\ell \mathcal{Re}[e^{i\phi} \{\tilde{A}_\parallel^L A_\perp^{L*} - \tilde{A}_\parallel^R A_\perp^{R*}\} + e^{-i\phi} \{A_\parallel^L \tilde{A}_\perp^{L*} - A_\parallel^R \tilde{A}_\perp^{R*}\}]
\end{aligned}$$

$$\begin{aligned}
h_{6c} &= 4\beta_\ell \frac{m_\ell}{\sqrt{q^2}} \mathcal{R}e[e^{i\phi}\{\tilde{A}_0^L A_S^* + \tilde{A}_0^R A_S^*\} + e^{-i\phi}\{A_0^L \tilde{A}_S^* + A_0^R \tilde{A}_S^*\}] \\
h_7 &= \sqrt{2}\beta_\ell \left[ \mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_{||}^{L*} - \tilde{A}_0^R A_{||}^{R*}\} + e^{-i\phi}\{A_0^L \tilde{A}_{||}^{L*} - A_0^R \tilde{A}_{||}^{R*}\}] \right. \\
&\quad \left. + \frac{m_\ell}{\sqrt{q^2}} \mathcal{I}m[e^{i\phi}\{\tilde{A}_\perp^L A_S^* + \tilde{A}_\perp^R A_S^*\} + e^{-i\phi}\{A_\perp^L \tilde{A}_S^* + A_\perp^R \tilde{A}_S^*\}] \right] \\
h_8 &= \frac{1}{\sqrt{2}}\beta_\ell^2 \mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_\perp^{L*} + \tilde{A}_0^R A_\perp^{R*}\} + e^{-i\phi}\{A_0^L \tilde{A}_\perp^{L*} + A_0^R \tilde{A}_\perp^{R*}\}] \\
h_9 &= -\beta_\ell^2 \mathcal{I}m[e^{i\phi}\{\tilde{A}_{||}^L A_\perp^{L*} + \tilde{A}_{||}^R A_\perp^{R*}\} + e^{-i\phi}\{A_{||}^L \tilde{A}_\perp^{L*} + A_{||}^R \tilde{A}_\perp^{R*}\}].
\end{aligned}$$

The  $s_i$  are connected to the imaginary part  $\mathcal{I}m$  of combinations of the transversity amplitudes and given by

$$\begin{aligned}
s_{1s} &= \frac{2 + \beta_\ell^2}{2} \mathcal{I}m[e^{i\phi}\{\tilde{A}_\perp^L A_\perp^{L*} + \tilde{A}_{||}^L A_{||}^{L*} + \tilde{A}_\perp^R A_\perp^{R*} + \tilde{A}_{||}^R A_{||}^{R*}\}] \\
&\quad + \frac{4m_\ell^2}{q^2} \mathcal{I}m[e^{i\phi}\{\tilde{A}_\perp^L A_\perp^{R*} + \tilde{A}_{||}^L A_{||}^{R*}\} - e^{-i\phi}\{A_\perp^L \tilde{A}_\perp^{R*} + A_{||}^L \tilde{A}_{||}^{R*}\}] \\
s_{1c} &= 2\mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_0^{L*} + \tilde{A}_0^R A_0^{R*}\}] \\
&\quad + \frac{8m_\ell^2}{q^2} \left[ \mathcal{I}m[e^{i\phi}\{\tilde{A}_t A_t^*\}] + \mathcal{I}m[e^{i\phi}\tilde{A}_0^L A_0^{R*} - e^{-i\phi}A_0^L \tilde{A}_0^{R*}] \right] + 2\beta_\ell^2 \mathcal{I}m[e^{i\phi}\tilde{A}_S A_S^*] \\
s_{2s} &= \frac{\beta_\ell^2}{2} \mathcal{I}m[e^{i\phi}\{\tilde{A}_\perp^L A_\perp^{L*} + \tilde{A}_{||}^L A_{||}^{L*} + \tilde{A}_\perp^R A_\perp^{R*} + \tilde{A}_{||}^R A_{||}^{R*}\}] \\
s_{2c} &= -2\beta_\ell^2 \mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_0^{L*} + \tilde{A}_0^R A_0^{R*}\}] \\
s_3 &= \beta_\ell^2 \mathcal{I}m[e^{i\phi}\{\tilde{A}_\perp^L A_\perp^{L*} - \tilde{A}_{||}^L A_{||}^{L*} + \tilde{A}_\perp^R A_\perp^{R*} - \tilde{A}_{||}^R A_{||}^{R*}\}] \\
s_4 &= \frac{1}{\sqrt{2}}\beta_\ell^2 \mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_{||}^{L*} + \tilde{A}_0^R A_{||}^{R*}\} - e^{-i\phi}\{A_0^L \tilde{A}_{||}^{L*} + A_0^R \tilde{A}_{||}^{R*}\}] \\
s_5 &= \sqrt{2}\beta_\ell \left[ \mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_\perp^{L*} - \tilde{A}_0^R A_\perp^{R*}\} - e^{-i\phi}\{A_0^L \tilde{A}_\perp^{L*} - A_0^R \tilde{A}_\perp^{R*}\}] \right. \\
&\quad \left. - \frac{m_\ell}{\sqrt{q^2}} \mathcal{I}m[e^{i\phi}\{\tilde{A}_{||}^L A_S^* + \tilde{A}_{||}^R A_S^*\} - e^{-i\phi}\{A_{||}^L \tilde{A}_S^* + A_{||}^R \tilde{A}_S^*\}] \right] \\
s_{6s} &= 2\beta_\ell \mathcal{I}m[e^{i\phi}\{\tilde{A}_{||}^L A_\perp^{L*} - \tilde{A}_{||}^R A_\perp^{R*}\} - e^{-i\phi}\{A_{||}^L \tilde{A}_\perp^{L*} - A_{||}^R \tilde{A}_\perp^{R*}\}] \\
s_{6c} &= 4\beta_\ell \frac{m_\ell}{\sqrt{q^2}} \mathcal{I}m[e^{i\phi}\{\tilde{A}_0^L A_S^* + \tilde{A}_0^R A_S^*\} - e^{-i\phi}\{A_0^L \tilde{A}_S^* + A_0^R \tilde{A}_S^*\}] \\
s_7 &= -\sqrt{2}\beta_\ell \left[ \mathcal{R}e[e^{i\phi}\{\tilde{A}_0^L A_{||}^{L*} - \tilde{A}_0^R A_{||}^{R*}\} - e^{-i\phi}\{A_0^L \tilde{A}_{||}^{L*} - A_0^R \tilde{A}_{||}^{R*}\}] \right. \\
&\quad \left. + \frac{m_\ell}{\sqrt{q^2}} \mathcal{R}e[e^{i\phi}\{\tilde{A}_\perp^L A_S^* + \tilde{A}_\perp^R A_S^*\} - e^{-i\phi}\{A_\perp^L \tilde{A}_S^* + A_\perp^R \tilde{A}_S^*\}] \right] \\
s_8 &= -\frac{1}{\sqrt{2}}\beta_\ell^2 \mathcal{R}e[e^{i\phi}\{\tilde{A}_0^L A_\perp^{L*} + \tilde{A}_0^R A_\perp^{R*}\} - e^{-i\phi}\{A_0^L \tilde{A}_\perp^{L*} + A_0^R \tilde{A}_\perp^{R*}\}] \\
s_9 &= \beta_\ell^2 \mathcal{R}e[e^{i\phi}\{\tilde{A}_{||}^L A_\perp^{L*} + \tilde{A}_{||}^R A_\perp^{R*}\} - e^{-i\phi}\{A_{||}^L \tilde{A}_\perp^{L*} + A_{||}^R \tilde{A}_\perp^{R*}\}].
\end{aligned}$$

# Appendix B

## Comparison of control mode data and simulation for all years

This Appendix shows the complete comparison for relevant distributions between simulation and data for the control mode  $B_s^0 \rightarrow J/\psi \phi$ . All distributions are shown for the used simulation in this thesis, namely under data taking conditions as they were in 2012, 2016 and 2017 data. The corrected and non-corrected simulation is given in red and green, respectively, and the data is given as black points. Full details of the corrections are discussed in Sec. 5.5.5.

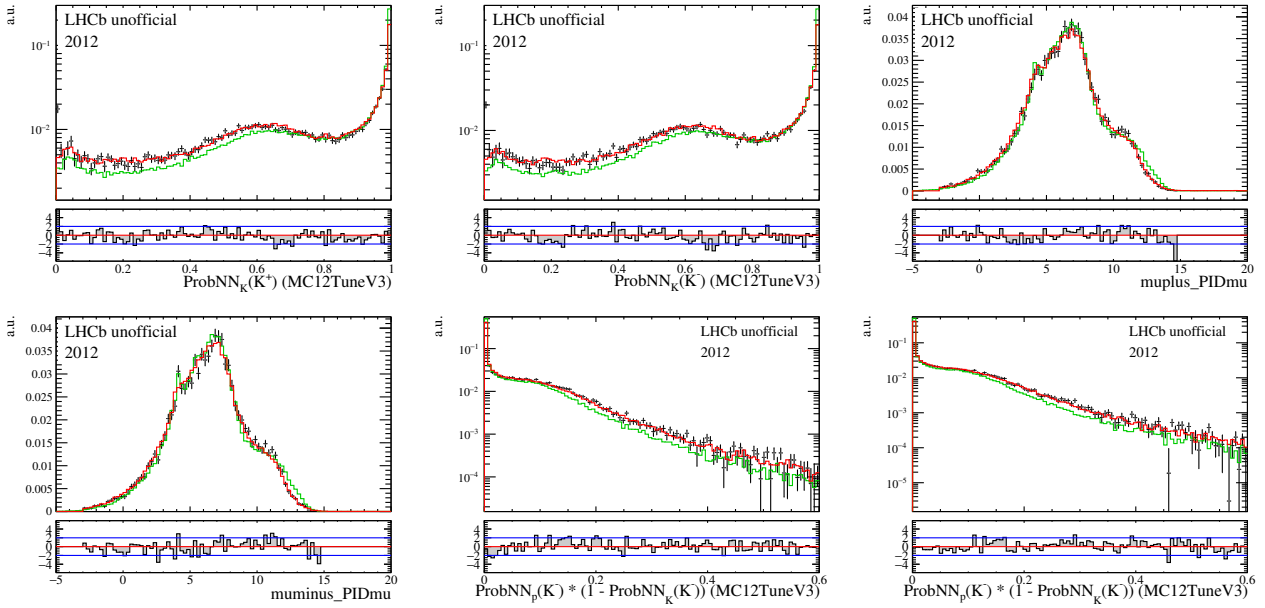


Figure B.1: Data and simulation comparison of the PID variables that are used in the preselection for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2012 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

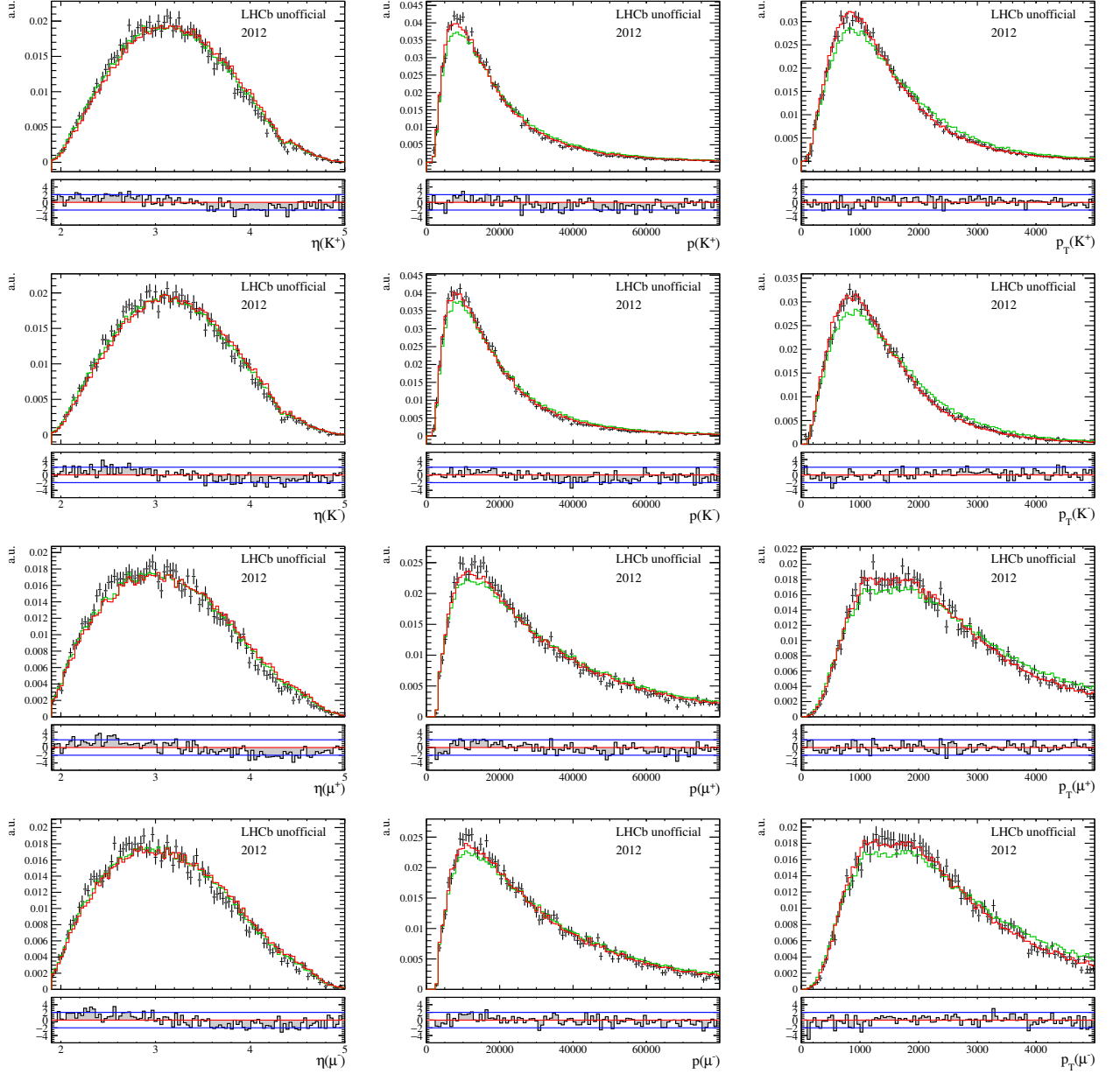


Figure B.2: Data and simulation comparison of the momenta  $p$  and  $p_T$  as well as  $\eta$  of the daughter particles for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2012 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

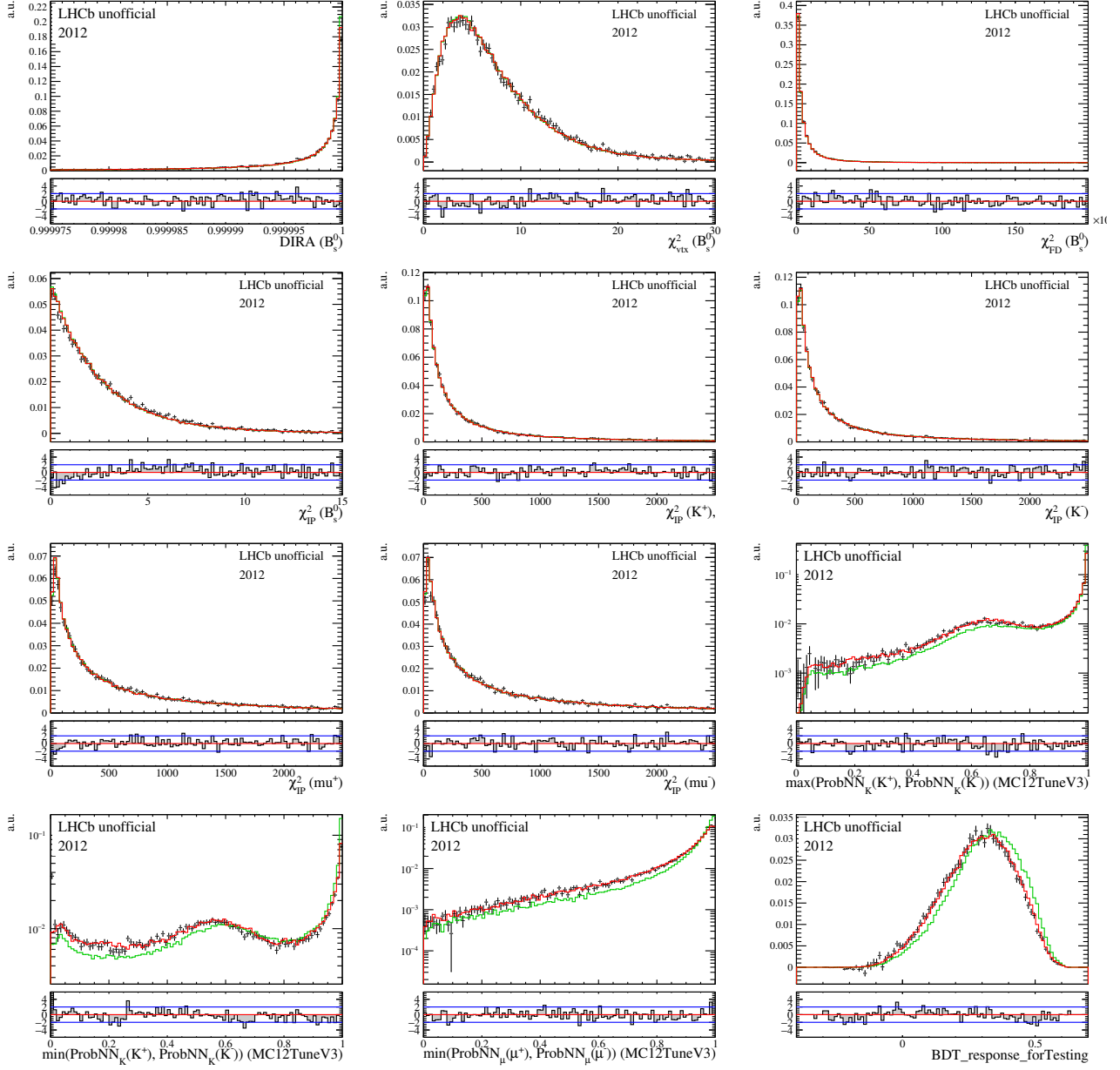


Figure B.3: Data and simulation comparison of the BDT input variables and the BDT response (bottom right) for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2012 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.



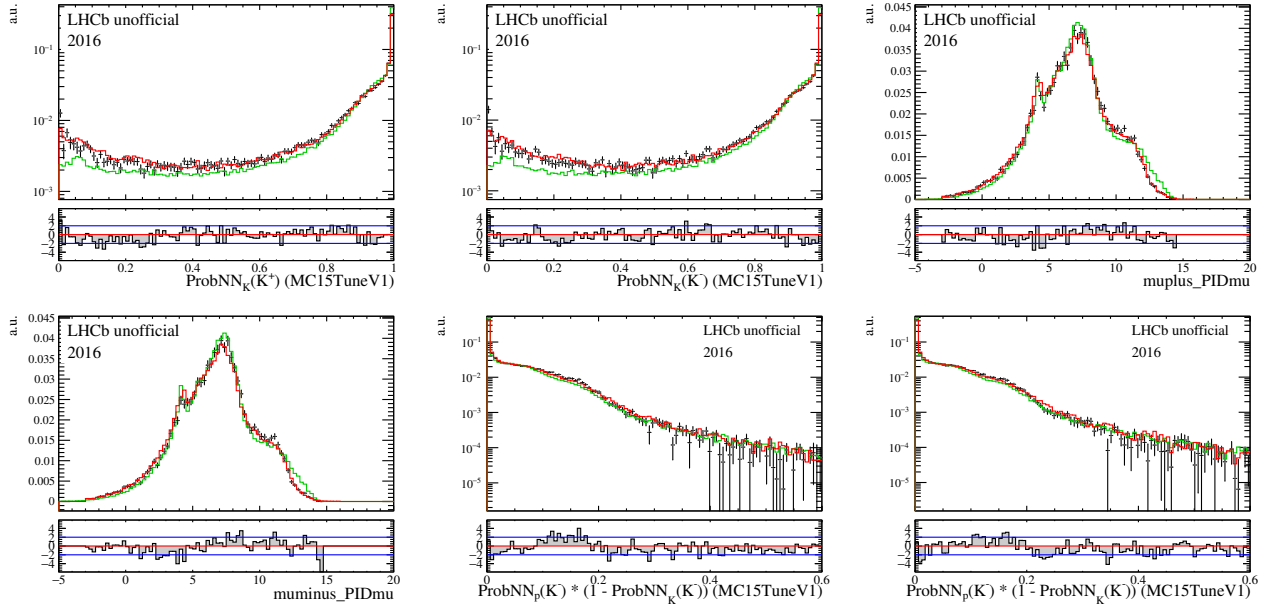


Figure B.4: Data and simulation comparison of the PID variables that are used in the preselection for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2016 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

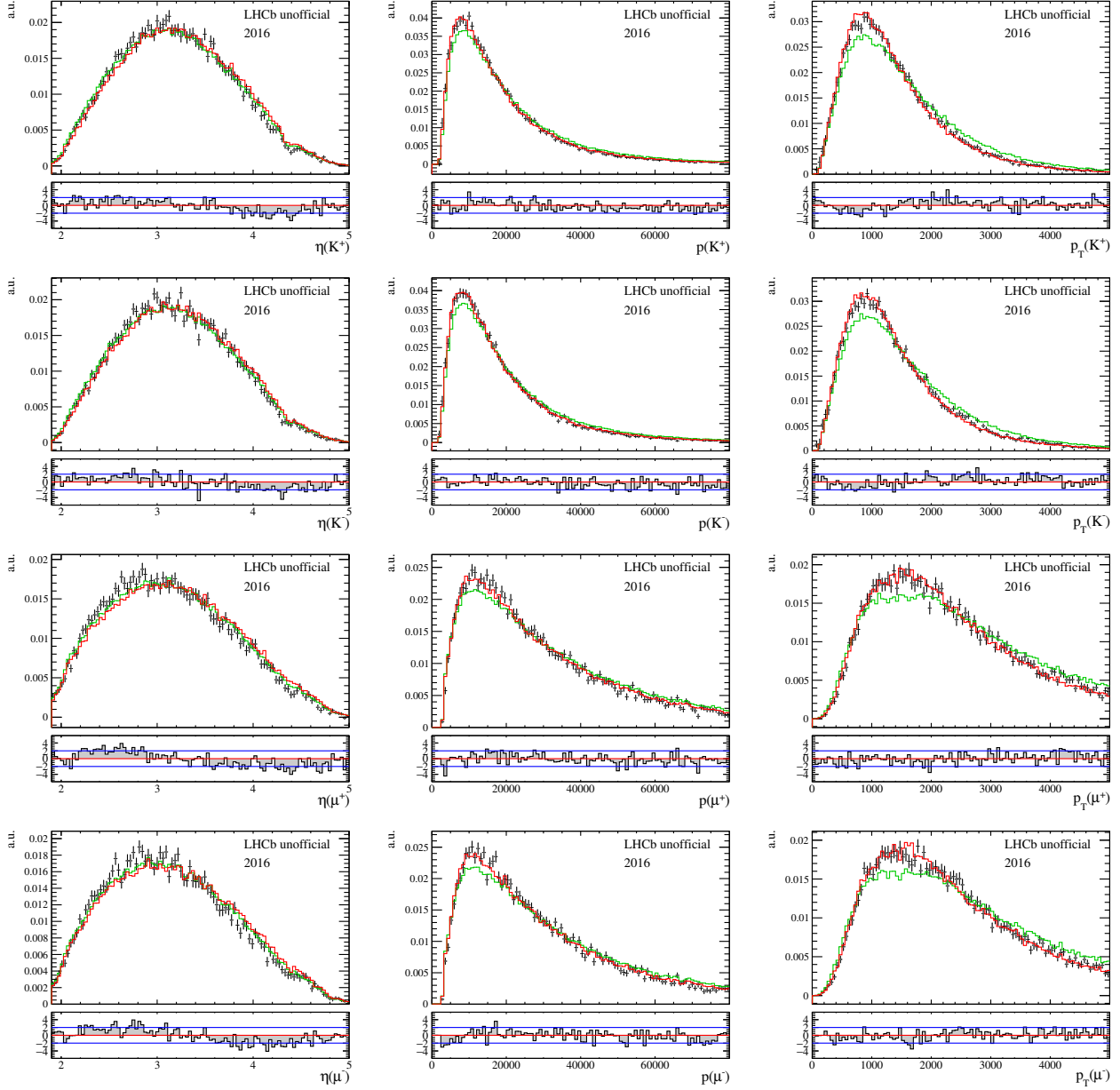


Figure B.5: Data and simulation comparison of the momenta  $p$  and  $p_T$  as well as  $\eta$  of the daughter particles for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2016 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

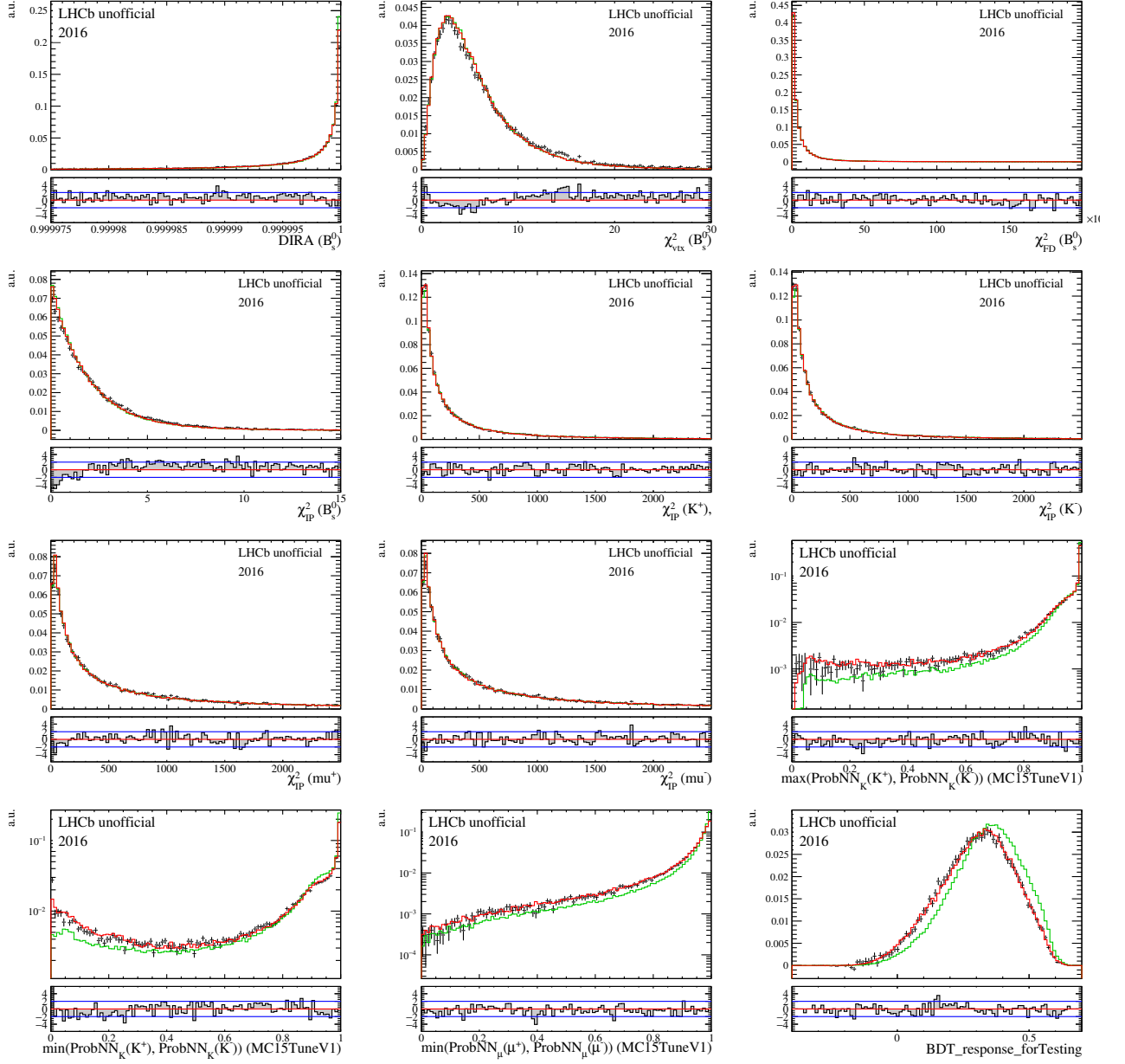


Figure B.6: Data and simulation comparison of the BDT input variables and the BDT response (bottom right) for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2016 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

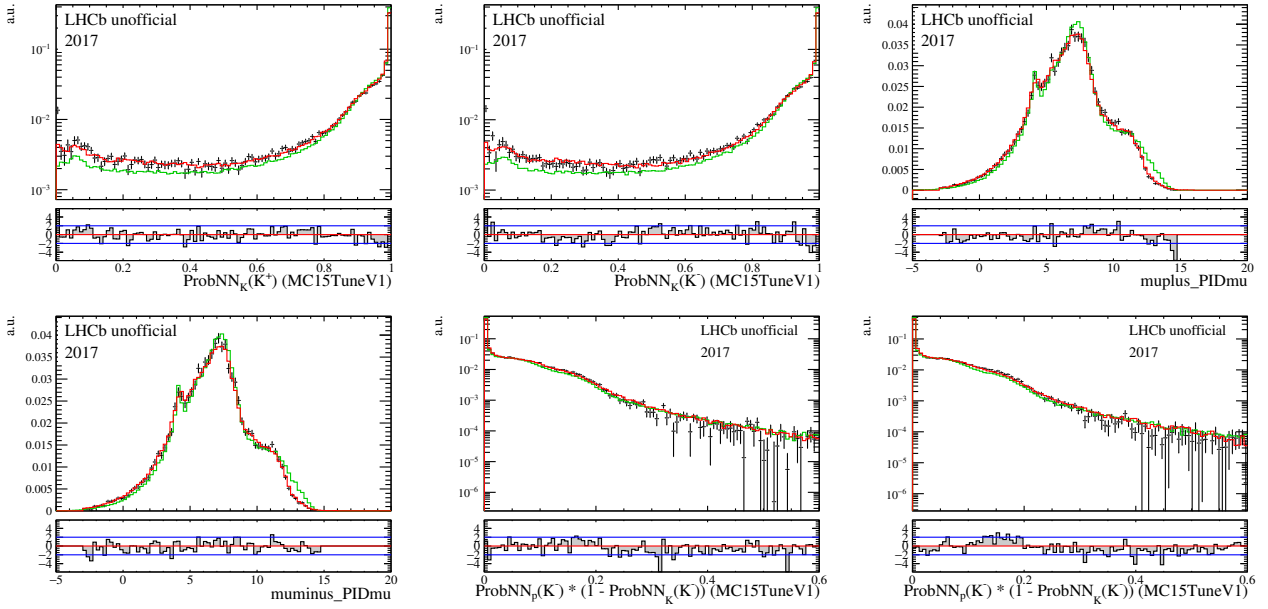


Figure B.7: Data and simulation comparison of the PID variables that are used in the preselection for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2017 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

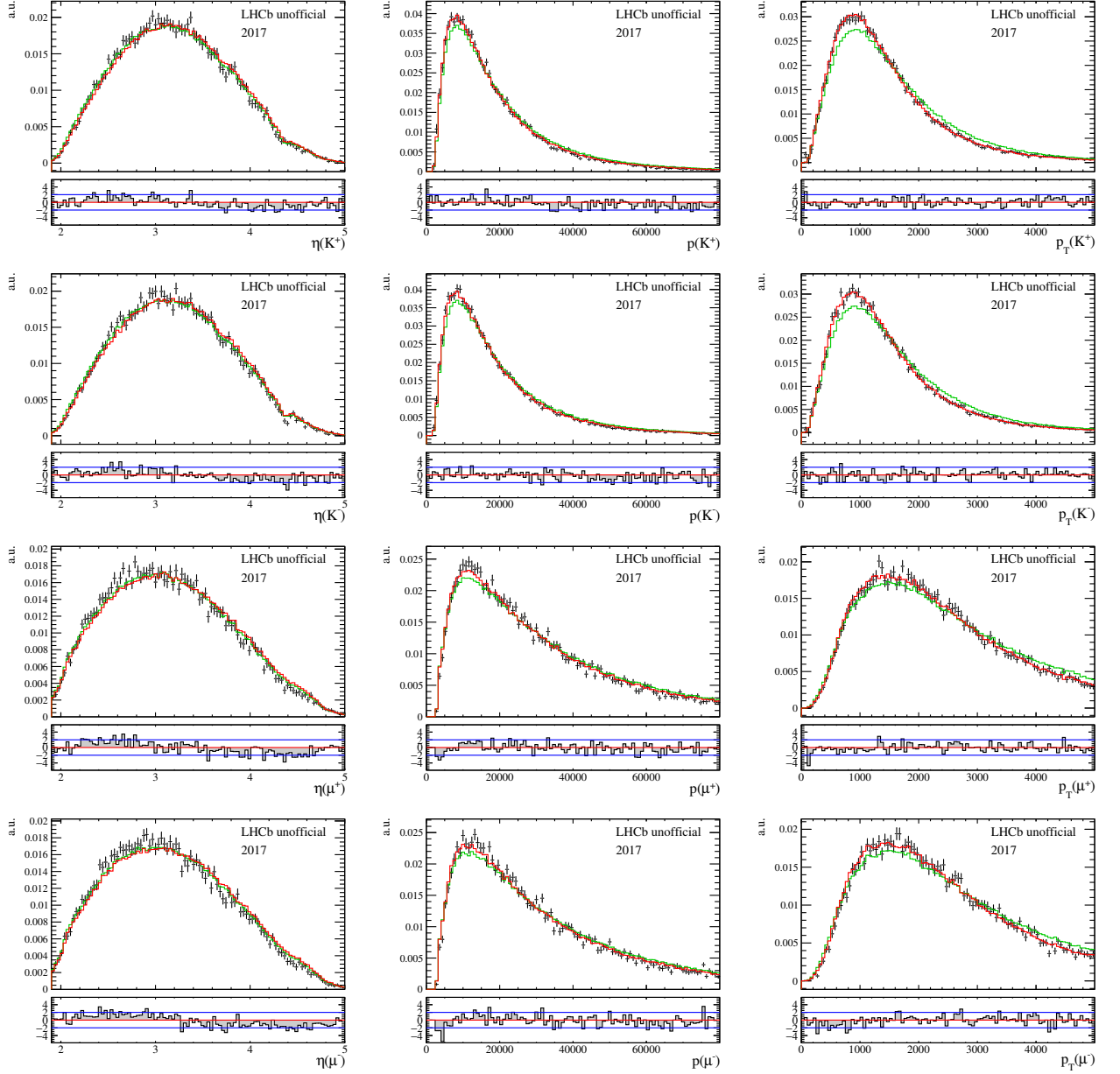


Figure B.8: Data and simulation comparison of the momenta  $p$  and  $p_T$  as well as  $\eta$  of the daughter particles for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2017 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

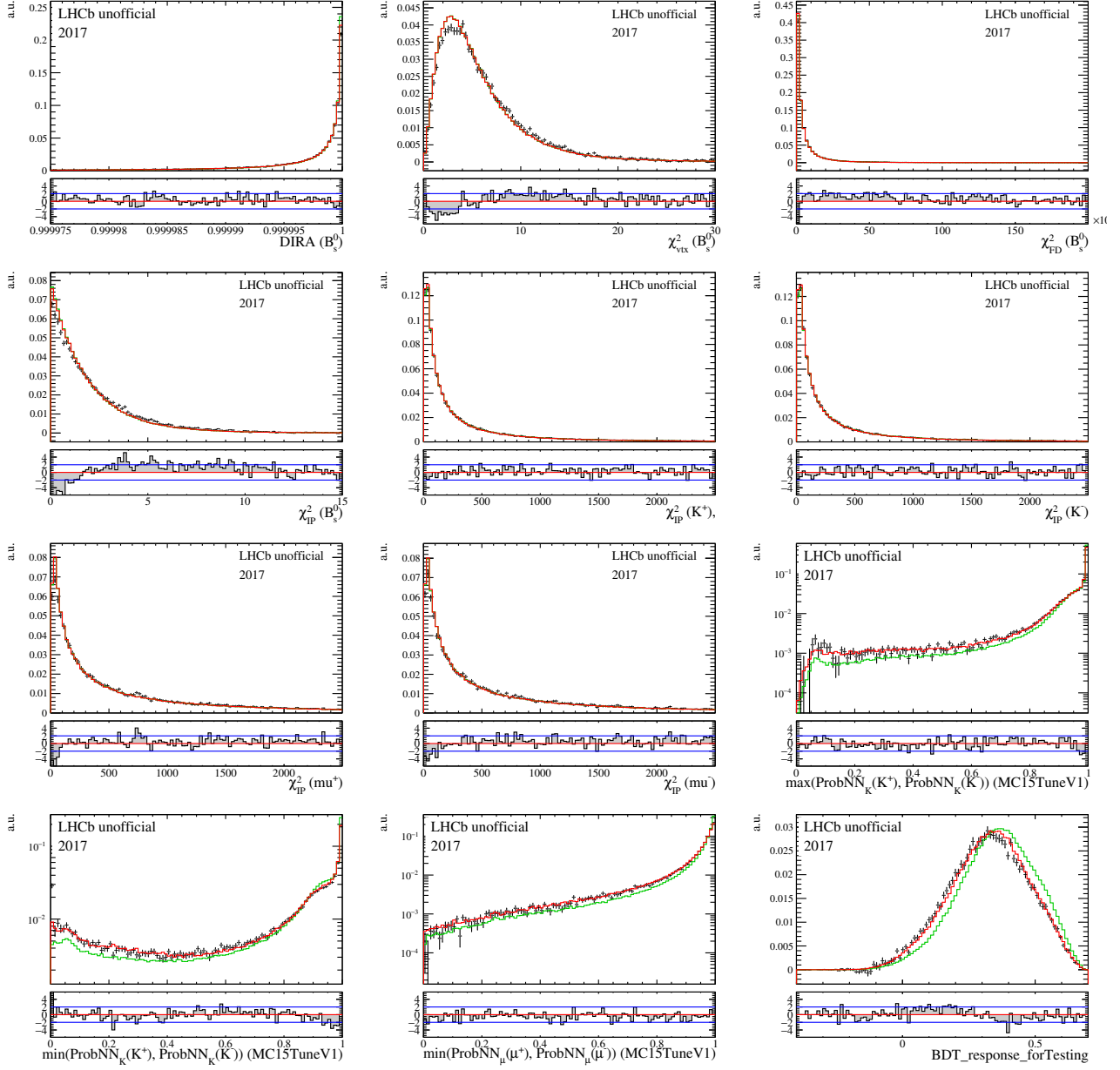


Figure B.9: Data and simulation comparison of the BDT input variables and the BDT response (bottom right) for the  $B_s^0 \rightarrow J/\psi \phi$  control mode under 2017 data taking conditions. The data (black points), as well as the raw (green) and corrected (red) simulation are shown.

# Appendix C

## Determining amplitudes from angular observables

In this Appendix the conversion of angular observables into amplitudes is derived taking the time dependence into account. This is an important step when comparing the fits to control mode data discussed in Sec. 7.3.2 to the dedicated  $B_s^0 \rightarrow J/\psi \phi$  angular analyses performed at LHCb in Refs. [239] and [240]. In Refs. [239, 240], the amplitudes of the  $\phi$  meson ( $|A_0|^2$  and  $|A_\perp|^2$ ) are measured at  $t = 0$ . Therefore, a comparison with these results, requires the  $|A_0(0)|^2$  and  $|A_\perp(0)|^2$  to be expressed in terms of the time-integrated angular observables extracted using the maximum-likelihood approach employed in this analysis. In Sec. 7.3.2.2 the relationship between the amplitudes and the angular observables is given in Eqs. 7.3-7.4 as

$$|A_0(0)|^2 + |\bar{A}_0(0)|^2 = F_L \left( 1 + \frac{(2S_3 - F_L + 1)}{F_+ (1 - 2S_3 + F_L) + F_- (1 + 2S_3 - F_L)} (F_+ - F_-) \right)$$

$$|A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 = \frac{F_- (2S_3 - F_L + 1)}{F_+ (1 - 2S_3 + F_L) + F_- (1 + 2S_3 - F_L)},$$

where  $F_-$  and  $F_+$  are constants which account for the time-dependence of  $(|A_{(0,\parallel)}(t)|^2 + |\bar{A}_{(0,\parallel)}(t)|^2)$  and  $(|A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2)$  [241], respectively and are defined by

$$F_\mp = \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \left( \cosh \left( \frac{\Delta\Gamma_s}{2} t \right) \mp \cos(\phi_s) \sinh \left( \frac{\Delta\Gamma_s}{2} t \right) \right) dt.$$

The SM values used to calculate  $F_\mp$  are  $\Gamma_s = 65.6895 \times 10^{10} s^{-1}$  and  $\Delta\Gamma_s/\Gamma_s = 0.166$  [65]. The  $\epsilon(t)$  term is the time-dependent efficiency and is determined on  $B_s^0 \rightarrow J/\psi \phi$  simulation.

The derivation of Eqs. 7.3-7.4 starts with the definition of  $F_L$  given in Eq. 2.7 as

$$F_L = \frac{\int_0^\infty \epsilon(t) (|A_0(t)|^2 + |\bar{A}_0(t)|^2) dt}{\int_0^\infty \epsilon(t) (|A_0(t)|^2 + |\bar{A}_0(t)|^2 + |A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2 + |A_\parallel(t)|^2 + |\bar{A}_\parallel(t)|^2) dt}.$$

The denominator can be rewritten by defining the time-dependent normalisation ( $\mathcal{N}(t)$ ) as  $\mathcal{N}(t) = \sum_{i=0,\perp,\parallel} (|A_i(t)|^2 + |\bar{A}_i(t)|^2)$ . At  $t = 0$ , the normalisation ( $\mathcal{N}(0)$ ) is unity, explicitly  $\mathcal{N}(0) = \sum_{i=0,\perp,\parallel} (|A_i(0)|^2 + |\bar{A}_i(0)|^2) = 1$ . With this definition, the denominator in the

definition of  $F_L$  can be rewritten as

$$\begin{aligned}
\int_0^\infty \epsilon(t) \mathcal{N}(t) &= \int_0^\infty \epsilon(t) \sum_{i=0,\perp,\parallel} \left( |A_i(t)|^2 + |\bar{A}_i(t)|^2 \right) dt \\
&= \sum_{i=0,\parallel} \left( |A_i(0)|^2 + |\bar{A}_i(0)|^2 \right) F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) F_+ \\
&= \left[ 1 - \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) \right] F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) F_+ \\
&= F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ - F_-).
\end{aligned}$$

With this relationship,  $F_L$  can be given in terms of the transversity amplitudes at  $t = 0$  as

$$(C.1) \quad F_L = \frac{\left( |A_0(0)|^2 + |\bar{A}_0(0)|^2 \right) F_-}{F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ - F_-)}.$$

In Eq. 2.7,  $S_3$  is similarly defined as

$$S_3 = \frac{1}{2} \frac{\int_0^\infty \epsilon(t) (|A_\perp(t)|^2 + |\bar{A}_\perp(t)|^2 - |A_\parallel(t)|^2 - |\bar{A}_\parallel(t)|^2) dt}{\int_0^\infty \epsilon(t) \mathcal{N}(t)}.$$

$S_3$  can analogously be given in terms of the amplitudes at  $t = 0$  by rearranging:

$$\begin{aligned}
2S_3 &= \frac{\left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) F_+ - \left( |A_\parallel(0)|^2 + |\bar{A}_\parallel(0)|^2 \right) F_-}{F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ - F_-)} \\
&= \frac{\left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ + F_-) - \left( 1 - \left( |A_0(0)|^2 + |\bar{A}_0(0)|^2 \right) \right) F_-}{F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ - F_-)} \\
&= \frac{\left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ + F_-) - F_-}{F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ - F_-)} + F_L.
\end{aligned}$$

With this dependency, the transversity amplitude  $\left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right)$  can be solved for by rearranging and is found to be

$$(C.2) \quad \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) = \frac{F_- (2S_3 - F_L + 1)}{F_+ (1 - 2S_3 + F_L) + F_- (1 + 2S_3 - F_L)}.$$

Similarly, rearranging Eq C.1 for  $\left( |A_0(0)|^2 + |\bar{A}_0(0)|^2 \right)$  and substituting Eq. C.2 to write the transversity amplitude in terms of angular observables gives

$$\begin{aligned}
\left( |A_0(0)|^2 + |\bar{A}_0(0)|^2 \right) &= F_L \frac{F_- + \left( |A_\perp(0)|^2 + |\bar{A}_\perp(0)|^2 \right) (F_+ - F_-)}{F_-} \\
&= F_L \left( 1 + \frac{(2S_3 - F_L + 1)}{F_+ (1 - 2S_3 + F_L) + F_- (1 + 2S_3 - F_L)} (F_+ - F_-) \right).
\end{aligned}$$



# Appendix D

## Distributions of the pulls from the fit validation

In this Appendix the pulls for the each of the angular observables from the pseudoexperiments for the fit validation are shown in Figs. D.1-D.7 for the angular observables from blind fits to data.

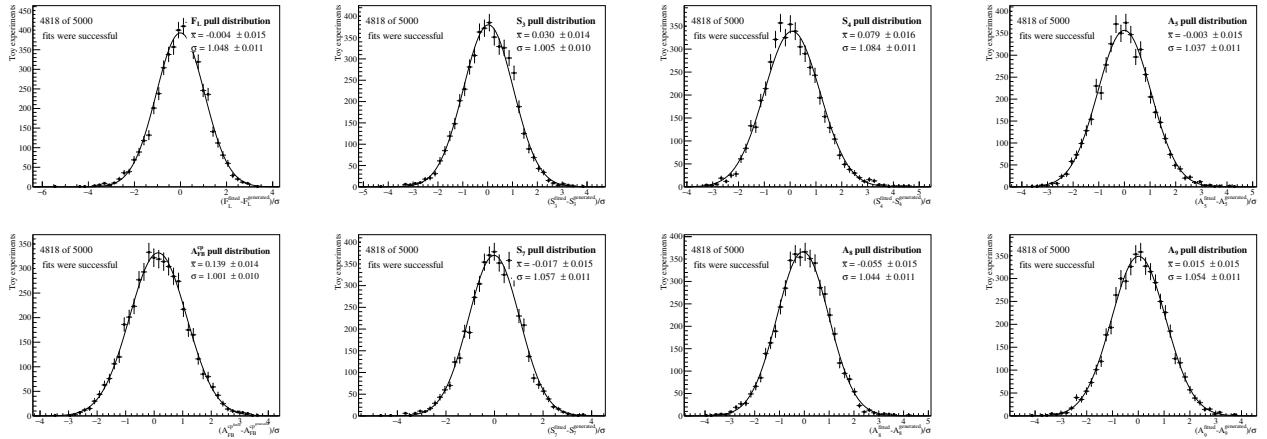


Figure D.1: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[0.1 \text{ GeV}^2/c^4, 0.98 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

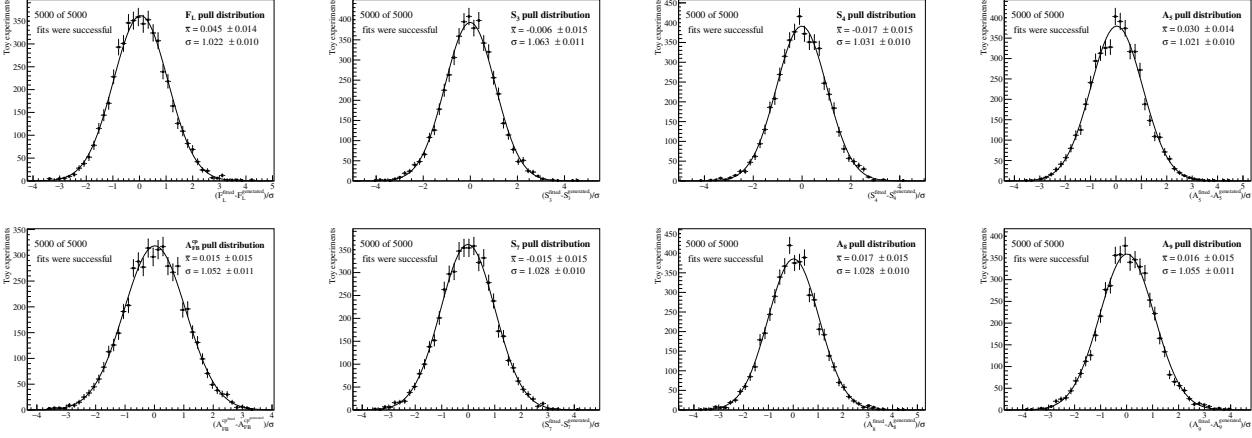


Figure D.2: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[1.1 \text{ GeV}^2/c^4, 4.0 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

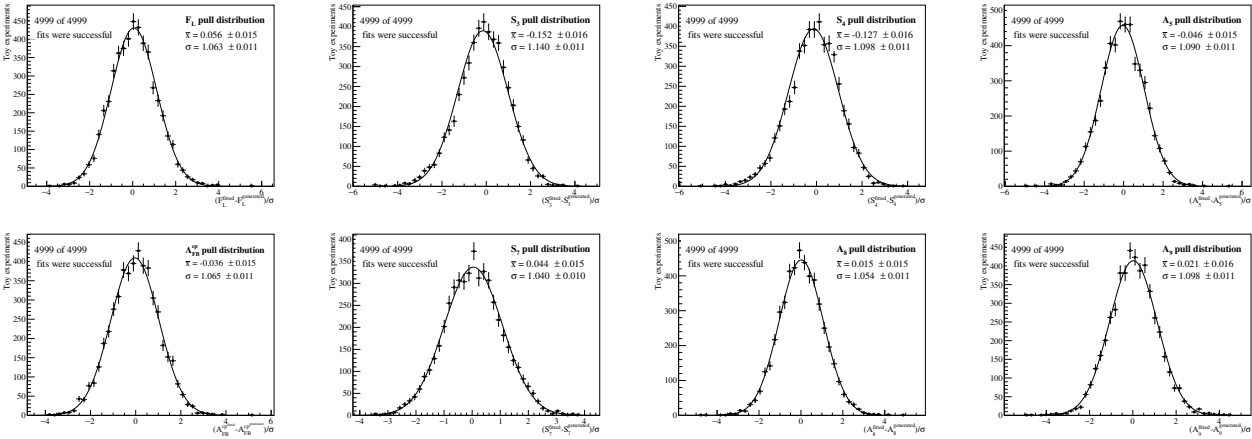


Figure D.3: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[4.0 \text{ GeV}^2/c^4, 6.0 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

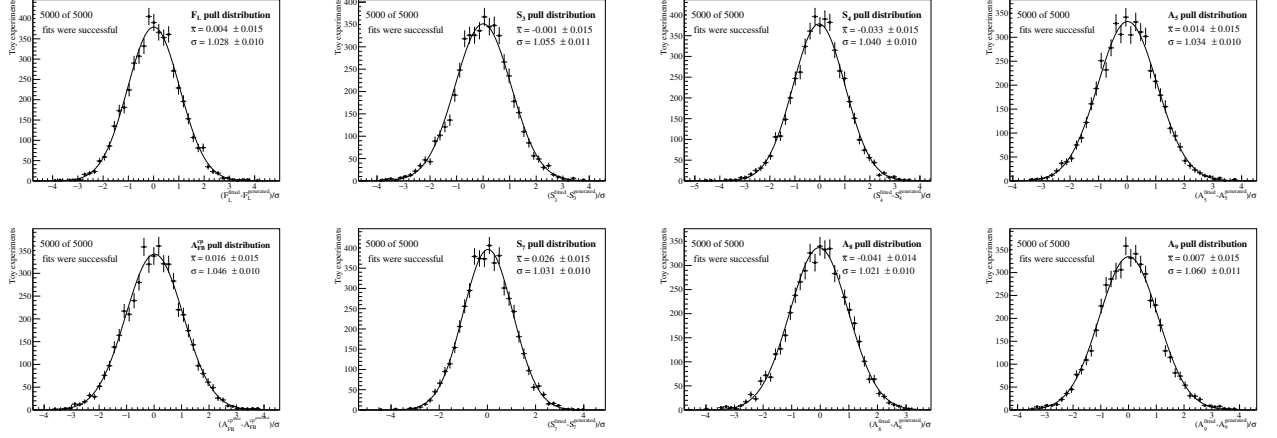


Figure D.4: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[6.0 \text{ GeV}^2/c^4, 8.0 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

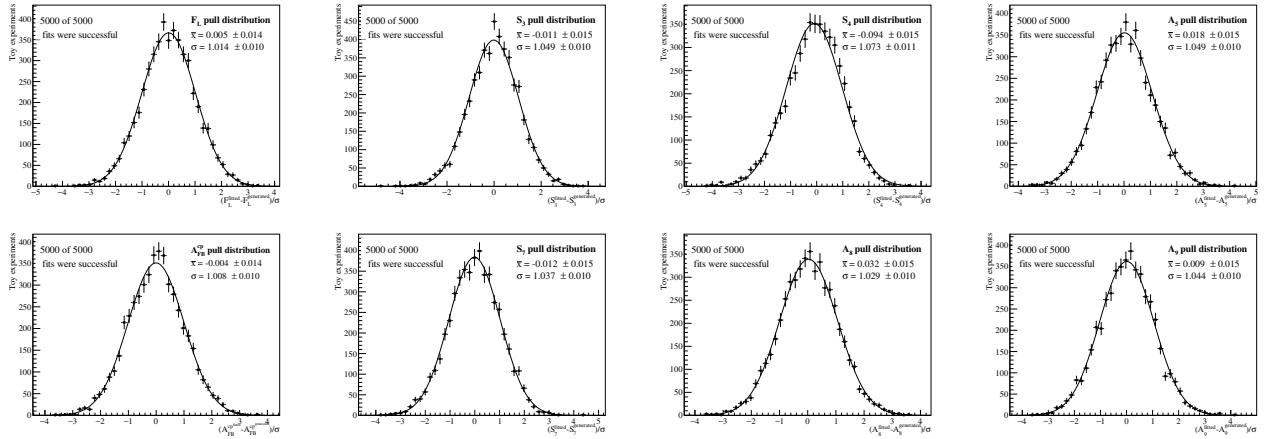


Figure D.5: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[11.0 \text{ GeV}^2/c^4, 12.5 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

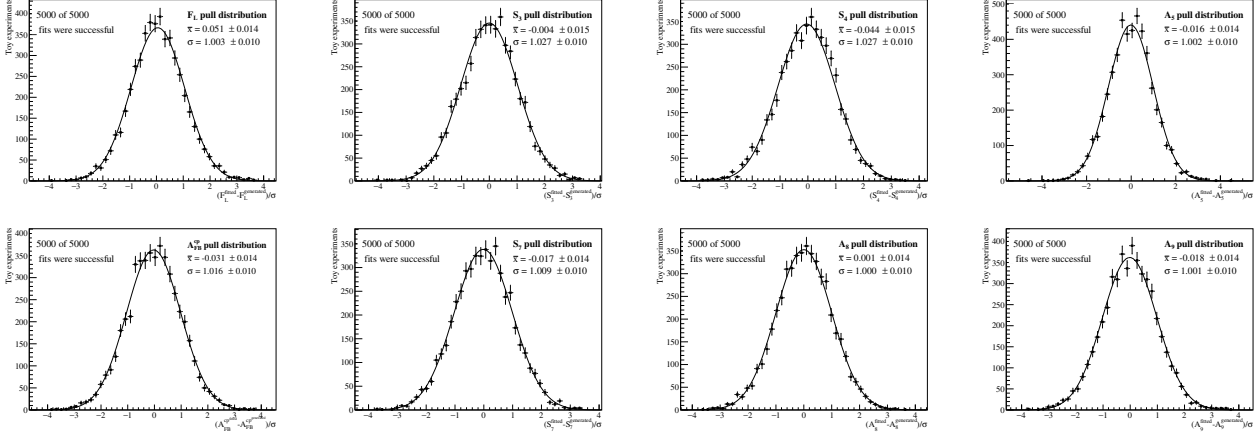


Figure D.6: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[1.1 \text{ GeV}^2/c^4, 6.0 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

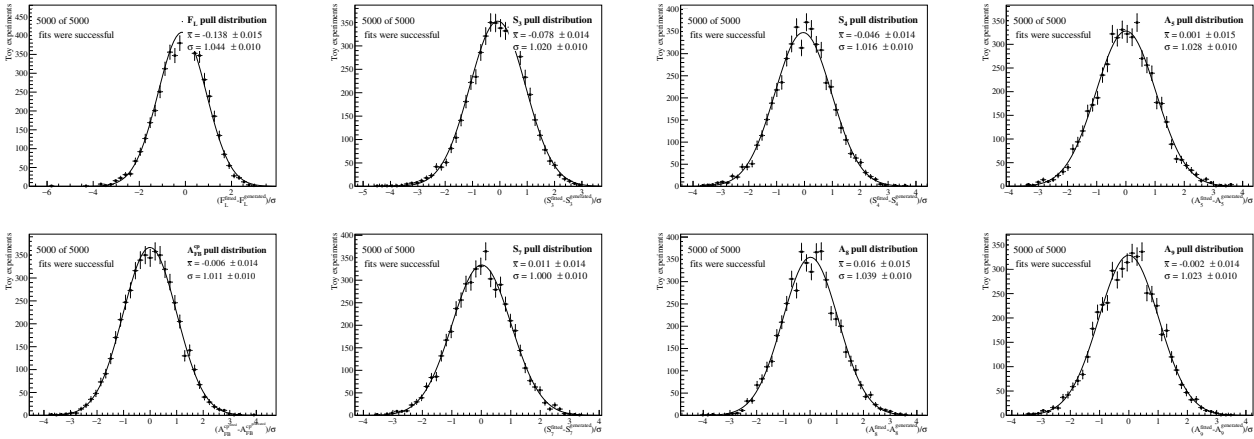


Figure D.7: The pulls from the pseudoexperiments for the fit validation for each angular observable in the  $q^2$  bin of  $[15.0 \text{ GeV}^2/c^4, 18.9 \text{ GeV}^2/c^4]$ . The angular observables are taken from blind fits to data.

# Appendix E

## Alternative tables for the systematic uncertainties

In this Appendix the absolute values for each systematic uncertainty and the quadratic sum are given for each  $q^2$  bin independently. In the last line, the expected statistical uncertainty is also given. The summary table with the systematic uncertainty relative to the expected statistical uncertainty is also given. Furthermore, in Tab. E.8 the absolute values for the remaining bias using the fit validation pseudoexperiments with angular observables from the SM are given for completeness.

Table E.1: Absolute systematic uncertainties in the bin  $q^2 \in [0.1, 0.98]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| peaking backgrounds           | 0.000  | 0.000  | 0.000  | 0.000  | 0.000                | -0.000 | 0.000  | -0.000 |
| signal mass model             | 0.000  | 0.000  | 0.000  | 0.000  | 0.000                | -0.000 | -0.000 | -0.000 |
| angular background model      | -0.000 | -0.001 | 0.001  | -0.000 | 0.000                | 0.000  | 0.000  | 0.000  |
| acceptance order              | -0.001 | -0.002 | -0.000 | -0.000 | -0.000               | -0.000 | -0.000 | 0.000  |
| simulation correction         | 0.005  | 0.001  | 0.000  | 0.000  | 0.000                | 0.000  | 0.000  | 0.000  |
| PID correction                | 0.001  | 0.001  | 0.000  | 0.000  | 0.000                | 0.000  | 0.000  | 0.000  |
| simulation statistics         | 0.000  | 0.000  | 0.000  | 0.000  | -0.000               | 0.000  | -0.000 | -0.000 |
| $q^2$ dependency              | -0.003 | 0.001  | -0.000 | -0.000 | -0.000               | 0.000  | -0.000 | -0.000 |
| S-wave                        | -0.000 | -0.001 | -0.002 | 0.000  | 0.000                | 0.001  | 0.000  | -0.000 |
| lifetime mismodelling         | 0.011  | -0.010 | -0.000 | -0.000 | 0.000                | -0.000 | -0.000 | -0.000 |
| time integration              | 0.009  | 0.009  | 0.003  | 0.000  | 0.000                | 0.001  | 0.000  | 0.001  |
| shared backgrounds            | 0.006  | -0.001 | 0.000  | -0.000 | -0.001               | -0.000 | 0.000  | -0.000 |
| remaining Bias                | -0.002 | -0.002 | -0.004 | 0.000  | -0.009               | -0.000 | 0.004  | -0.001 |
| quadratic sum                 | 0.017  | 0.014  | 0.005  | 0.001  | 0.009                | 0.001  | 0.004  | 0.001  |
| absolute expected statistical | 0.048  | 0.068  | 0.080  | 0.068  | 0.063                | 0.067  | 0.081  | 0.068  |

Table E.2: Absolute systematic uncertainties in the bin  $q^2 \in [1.1, 4.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|---------------|--------|--------|--------|
| peaking backgrounds           | -0.000 | 0.000  | -0.000 | -0.000 | -0.000        | -0.000 | 0.000  | 0.000  |
| signal mass model             | 0.000  | 0.000  | -0.000 | 0.000  | -0.000        | -0.000 | 0.000  | -0.000 |
| angular background model      | -0.006 | 0.000  | -0.001 | 0.000  | -0.000        | 0.000  | 0.000  | 0.000  |
| acceptance order              | -0.001 | 0.000  | 0.000  | 0.000  | -0.000        | -0.000 | 0.000  | -0.000 |
| simulation correction         | 0.003  | 0.000  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| PID correction                | 0.000  | 0.000  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| simulation statistics         | -0.000 | -0.000 | -0.000 | 0.000  | -0.000        | -0.000 | -0.000 | 0.000  |
| $q^2$ dependency              | 0.002  | -0.000 | 0.000  | -0.000 | 0.000         | 0.000  | 0.000  | 0.000  |
| S-wave                        | -0.009 | 0.000  | 0.001  | -0.000 | -0.000        | 0.001  | 0.000  | 0.000  |
| lifetime mismodelling         | 0.007  | -0.003 | 0.001  | -0.000 | -0.000        | 0.000  | -0.000 | 0.000  |
| time integration              | 0.006  | 0.004  | 0.001  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| shared backgrounds            | -0.002 | 0.000  | -0.000 | -0.000 | 0.000         | -0.000 | -0.000 | -0.000 |
| remaining Bias                | -0.002 | 0.000  | 0.000  | -0.002 | -0.001        | 0.001  | -0.001 | -0.001 |
| quadratic sum                 | 0.015  | 0.004  | 0.002  | 0.002  | 0.001         | 0.001  | 0.001  | 0.001  |
| absolute expected statistical | 0.047  | 0.054  | 0.079  | 0.071  | 0.045         | 0.071  | 0.080  | 0.054  |

Table E.3: Absolute systematic uncertainties in the bin  $q^2 \in [4.0, 6.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|---------------|--------|--------|--------|
| peaking backgrounds           | -0.000 | -0.000 | 0.000  | -0.000 | 0.000         | -0.000 | -0.000 | 0.000  |
| signal mass model             | -0.000 | 0.000  | 0.000  | -0.000 | -0.000        | 0.000  | -0.000 | 0.000  |
| angular background model      | 0.011  | 0.001  | -0.003 | 0.000  | 0.001         | 0.000  | 0.000  | -0.000 |
| acceptance order              | 0.001  | 0.000  | 0.000  | -0.000 | 0.000         | 0.000  | 0.000  | -0.000 |
| simulation correction         | 0.000  | 0.001  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| PID correction                | 0.001  | 0.001  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| simulation statistics         | 0.000  | -0.000 | 0.000  | -0.000 | -0.000        | -0.000 | 0.000  | -0.000 |
| $q^2$ dependency              | -0.001 | -0.000 | -0.000 | -0.000 | 0.000         | -0.000 | 0.000  | 0.000  |
| S-wave                        | -0.006 | 0.001  | 0.004  | 0.000  | 0.000         | 0.000  | -0.000 | 0.000  |
| lifetime mismodelling         | 0.008  | -0.004 | 0.002  | -0.000 | 0.000         | -0.000 | 0.000  | -0.000 |
| time integration              | 0.006  | 0.004  | 0.002  | 0.001  | 0.000         | 0.000  | 0.001  | 0.000  |
| shared backgrounds            | 0.000  | 0.000  | 0.000  | 0.000  | -0.000        | -0.000 | -0.000 | -0.000 |
| remaining Bias                | -0.002 | 0.010  | 0.009  | 0.003  | 0.002         | -0.003 | -0.001 | -0.001 |
| quadratic sum                 | 0.016  | 0.012  | 0.010  | 0.003  | 0.002         | 0.003  | 0.002  | 0.001  |
| absolute expected statistical | 0.050  | 0.066  | 0.080  | 0.078  | 0.047         | 0.080  | 0.083  | 0.066  |

Table E.4: Absolute systematic uncertainties in the bin  $q^2 \in [6.0, 8.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|---------------|--------|--------|--------|
| peaking backgrounds           | 0.000  | -0.000 | 0.000  | -0.000 | 0.000         | 0.000  | -0.000 | 0.000  |
| signal mass model             | -0.000 | 0.000  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| angular background model      | -0.003 | -0.002 | -0.001 | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| acceptance order              | 0.001  | 0.000  | 0.000  | 0.000  | -0.000        | 0.000  | -0.000 | 0.000  |
| simulation correction         | 0.002  | 0.000  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| PID correction                | 0.001  | 0.001  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| simulation statistics         | 0.000  | -0.000 | 0.000  | -0.000 | -0.000        | -0.000 | -0.000 | -0.000 |
| $q^2$ dependency              | -0.000 | -0.000 | -0.000 | 0.000  | -0.000        | -0.000 | -0.000 | -0.000 |
| S-wave                        | -0.004 | 0.001  | 0.004  | -0.000 | 0.000         | 0.000  | -0.000 | 0.000  |
| lifetime mismodelling         | 0.009  | -0.006 | 0.002  | -0.000 | 0.000         | -0.000 | -0.000 | -0.000 |
| time integration              | 0.008  | 0.006  | 0.003  | 0.001  | 0.000         | 0.000  | 0.001  | 0.001  |
| shared backgrounds            | 0.000  | 0.000  | 0.000  | 0.000  | 0.000         | 0.000  | -0.000 | -0.000 |
| remaining Bias                | 0.000  | 0.000  | 0.001  | -0.001 | -0.001        | -0.002 | 0.002  | -0.001 |
| quadratic sum                 | 0.012  | 0.009  | 0.006  | 0.001  | 0.001         | 0.002  | 0.002  | 0.001  |
| absolute expected statistical | 0.049  | 0.069  | 0.077  | 0.075  | 0.048         | 0.078  | 0.081  | 0.069  |

Table E.5: Absolute systematic uncertainties in the bin  $q^2 \in [11.0, 12.5]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{FB}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|---------------|--------|--------|--------|
| peaking backgrounds           | -0.000 | 0.000  | -0.000 | 0.000  | -0.000        | -0.000 | -0.000 | -0.000 |
| signal mass model             | 0.000  | -0.000 | -0.000 | 0.000  | 0.000         | -0.000 | 0.000  | -0.000 |
| angular background model      | 0.004  | -0.002 | 0.002  | -0.000 | 0.000         | -0.000 | -0.000 | 0.000  |
| acceptance order              | 0.000  | -0.000 | 0.000  | -0.000 | 0.000         | -0.000 | 0.000  | -0.000 |
| simulation correction         | 0.004  | 0.001  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| PID correction                | 0.000  | 0.001  | 0.000  | 0.000  | 0.000         | 0.000  | 0.000  | 0.000  |
| simulation statistics         | 0.000  | -0.000 | 0.000  | 0.000  | 0.000         | -0.000 | 0.000  | -0.000 |
| $q^2$ dependency              | -0.000 | -0.000 | 0.000  | -0.000 | -0.000        | 0.000  | 0.000  | 0.000  |
| S-wave                        | -0.000 | 0.003  | 0.004  | 0.000  | -0.000        | -0.000 | -0.000 | -0.000 |
| lifetime mismodelling         | 0.007  | -0.009 | 0.002  | 0.000  | 0.000         | -0.000 | -0.000 | 0.000  |
| time integration              | 0.007  | 0.009  | 0.005  | 0.001  | 0.001         | 0.000  | 0.001  | 0.001  |
| shared backgrounds            | 0.000  | 0.000  | -0.000 | 0.000  | -0.000        | 0.000  | 0.000  | 0.000  |
| remaining Bias                | -0.001 | -0.001 | 0.004  | -0.001 | 0.000         | 0.000  | -0.002 | -0.000 |
| quadratic sum                 | 0.012  | 0.013  | 0.008  | 0.001  | 0.001         | 0.000  | 0.002  | 0.001  |
| absolute expected statistical | 0.043  | 0.065  | 0.063  | 0.062  | 0.045         | 0.067  | 0.069  | 0.065  |

Table E.6: Absolute systematic uncertainties in the bin  $q^2 \in [1.1, 6.0]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| peaking backgrounds           | -0.000 | -0.000 | 0.000  | -0.000 | -0.000               | -0.000 | -0.000 | 0.000  |
| signal mass model             | 0.002  | 0.001  | -0.002 | -0.000 | 0.000                | 0.000  | 0.000  | -0.000 |
| angular background model      | 0.002  | 0.000  | -0.002 | 0.000  | 0.000                | 0.000  | 0.000  | -0.000 |
| acceptance order              | -0.000 | 0.000  | -0.000 | -0.000 | 0.000                | 0.000  | -0.000 | 0.000  |
| simulation correction         | 0.003  | 0.001  | 0.000  | 0.000  | 0.000                | 0.000  | 0.000  | 0.000  |
| PID correction                | 0.000  | 0.000  | 0.000  | 0.000  | 0.000                | 0.000  | 0.000  | 0.000  |
| simulation statistics         | -0.000 | -0.000 | 0.000  | -0.000 | 0.000                | 0.000  | 0.000  | -0.000 |
| S-wave                        | -0.006 | 0.000  | 0.002  | -0.000 | -0.000               | 0.000  | 0.000  | 0.000  |
| lifetime mismodelling         | 0.009  | -0.005 | 0.001  | -0.000 | 0.000                | 0.000  | -0.000 | 0.000  |
| time integration              | 0.006  | 0.004  | 0.001  | 0.000  | 0.000                | 0.000  | 0.000  | 0.000  |
| shared backgrounds            | 0.000  | -0.000 | -0.000 | -0.000 | -0.000               | -0.000 | -0.000 | -0.000 |
| remaining Bias                | -0.001 | -0.001 | 0.002  | -0.001 | -0.001               | -0.002 | -0.001 | 0.001  |
| quadratic sum                 | 0.013  | 0.006  | 0.004  | 0.001  | 0.001                | 0.002  | 0.001  | 0.001  |
| absolute expected statistical | 0.036  | 0.043  | 0.059  | 0.054  | 0.034                | 0.055  | 0.060  | 0.043  |

Table E.7: Absolute systematic uncertainties in the bin  $q^2 \in [15.0, 18.9]$  and its quadratic sum in comparison with the statistical uncertainty.

| systematic uncertainty        | $F_L$  | $S_3$  | $S_4$  | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$  | $A_9$  |
|-------------------------------|--------|--------|--------|--------|----------------------|--------|--------|--------|
| peaking backgrounds           | -0.001 | -0.000 | 0.000  | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |
| signal mass model             | -0.005 | 0.001  | -0.000 | -0.000 | 0.000                | -0.000 | -0.000 | -0.000 |
| angular background model      | 0.004  | -0.001 | 0.002  | 0.000  | 0.000                | -0.000 | -0.000 | 0.000  |
| acceptance order              | -0.009 | -0.004 | 0.000  | 0.001  | -0.001               | 0.000  | 0.001  | -0.001 |
| simulation correction         | 0.013  | 0.002  | 0.000  | 0.001  | 0.000                | 0.001  | 0.001  | 0.000  |
| PID correction                | 0.002  | 0.002  | 0.001  | 0.000  | 0.000                | 0.000  | 0.000  | 0.000  |
| simulation statistics         | -0.006 | -0.002 | 0.000  | 0.000  | -0.000               | -0.000 | -0.000 | 0.000  |
| S-wave                        | 0.000  | 0.005  | 0.005  | 0.000  | -0.000               | -0.000 | -0.000 | -0.000 |
| lifetime mismodelling         | 0.006  | -0.008 | 0.001  | 0.000  | 0.000                | 0.000  | -0.000 | -0.000 |
| time integration              | 0.004  | 0.007  | 0.003  | 0.000  | 0.001                | 0.000  | 0.001  | 0.001  |
| shared backgrounds            | -0.000 | -0.000 | 0.000  | -0.000 | 0.000                | -0.000 | -0.000 | -0.000 |
| remaining Bias                | 0.003  | 0.002  | 0.000  | -0.001 | -0.000               | -0.001 | -0.001 | -0.000 |
| quadratic sum                 | 0.019  | 0.014  | 0.006  | 0.001  | 0.001                | 0.002  | 0.002  | 0.001  |
| absolute expected statistical | 0.029  | 0.038  | 0.048  | 0.048  | 0.026                | 0.048  | 0.049  | 0.038  |

Table E.8: The values for the systematic uncertainty originating from the remaining bias in the fit validation using the SM angular observables are shown.

| $q^2$ bin    | $F_L$  | $S_3$  | $S_4$ | $A_5$  | $A_{\text{FB}}^{CP}$ | $S_7$  | $A_8$ | $A_9$  |
|--------------|--------|--------|-------|--------|----------------------|--------|-------|--------|
| [0.1, 0.98]  | -0.001 | 0.000  | 0.000 | -0.001 | -0.001               | 0.001  | 0.000 | -0.001 |
| [1.1, 4.0]   | -0.003 | 0.002  | 0.002 | 0.001  | 0.001                | 0.002  | 0.001 | 0.002  |
| [4.0, 6.0]   | -0.002 | 0.001  | 0.003 | 0.001  | -0.001               | -0.000 | 0.002 | -0.000 |
| [6.0, 8.0]   | 0.000  | 0.001  | 0.002 | 0.001  | 0.001                | 0.000  | 0.003 | -0.001 |
| [11.0, 12.5] | 0.001  | -0.001 | 0.004 | 0.000  | 0.000                | -0.000 | 0.000 | 0.000  |
| [1.1, 6.0]   | -0.001 | 0.000  | 0.002 | -0.000 | -0.000               | -0.001 | 0.001 | -0.001 |
| [15.0, 18.9] | 0.003  | -0.001 | 0.001 | -0.000 | 0.000                | -0.001 | 0.001 | -0.000 |



# Appendix F

## Projections of fits to rare mode data

In this Appendix the fit projections of the four-dimensional unbinned maximum likelihood fit to the three decay angles and the mass of the  $B_s^0$  candidate to the Run 1 and Run 2 data sets are shown in Figs. F.1-F.21. The fit model describes the data well.

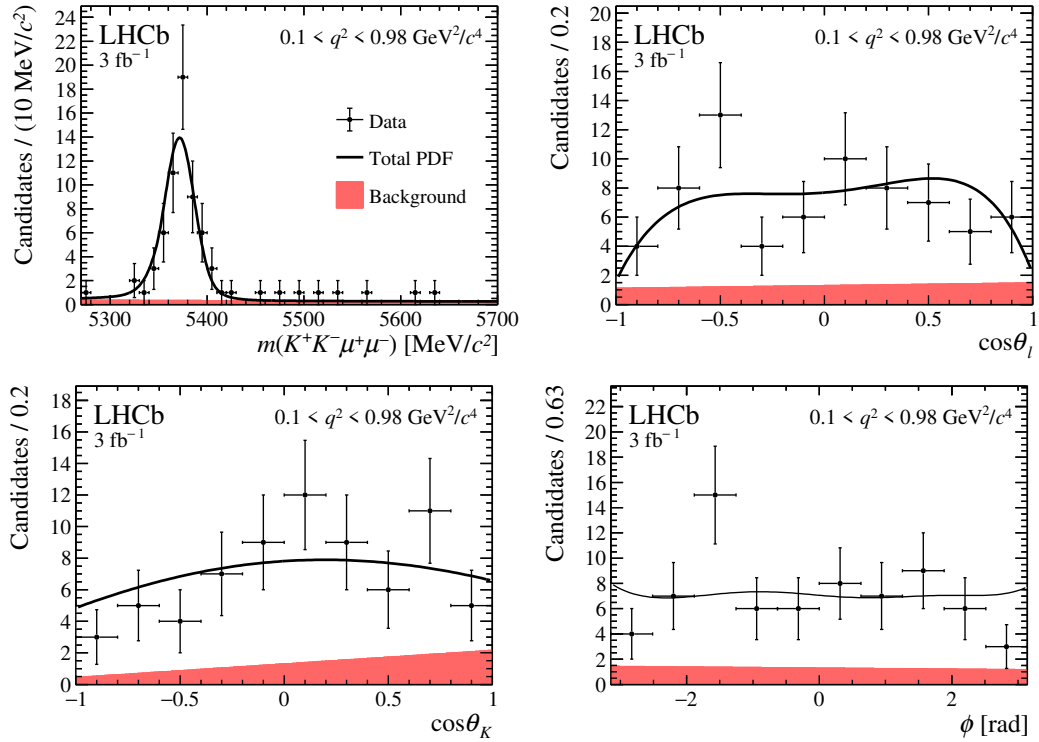


Figure F.1: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [0.1, 0.98] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

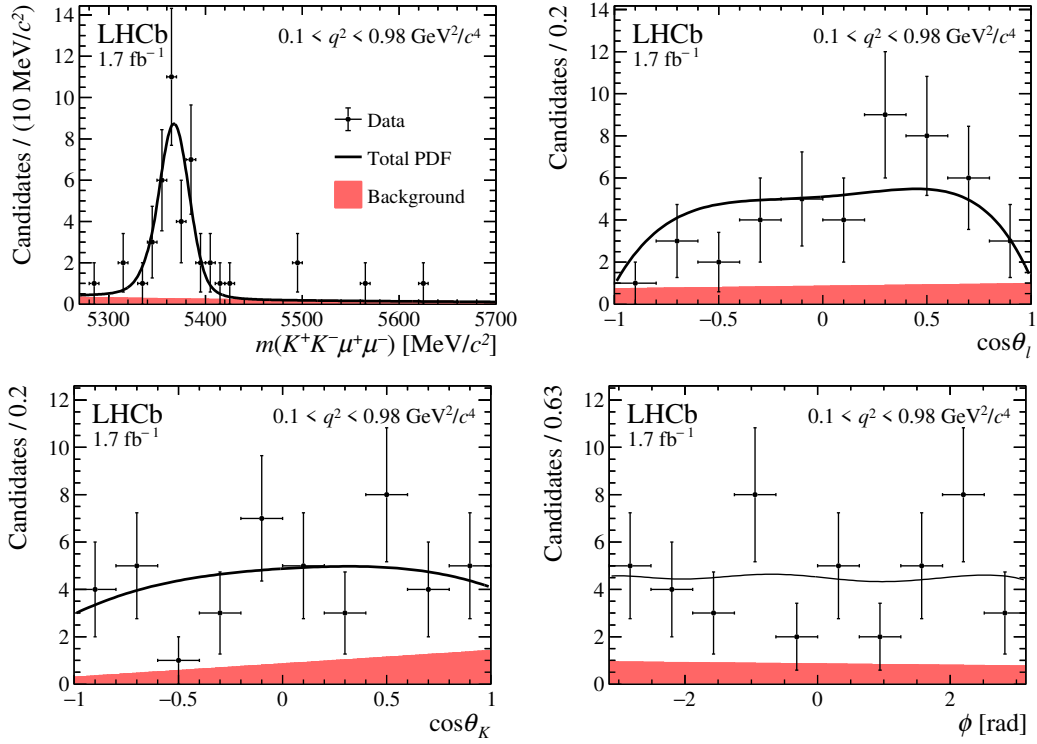


Figure F.2: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [0.1, 0.98] \text{ GeV}^2/c^4$  for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

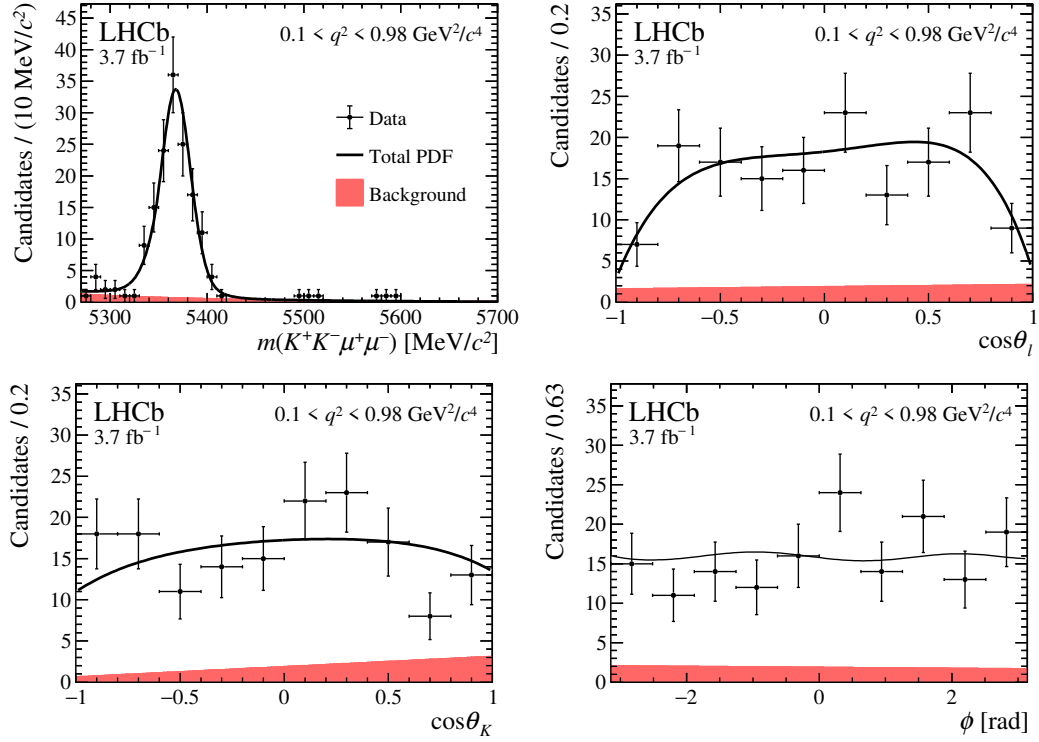


Figure F.3: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [0.1, 0.98] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

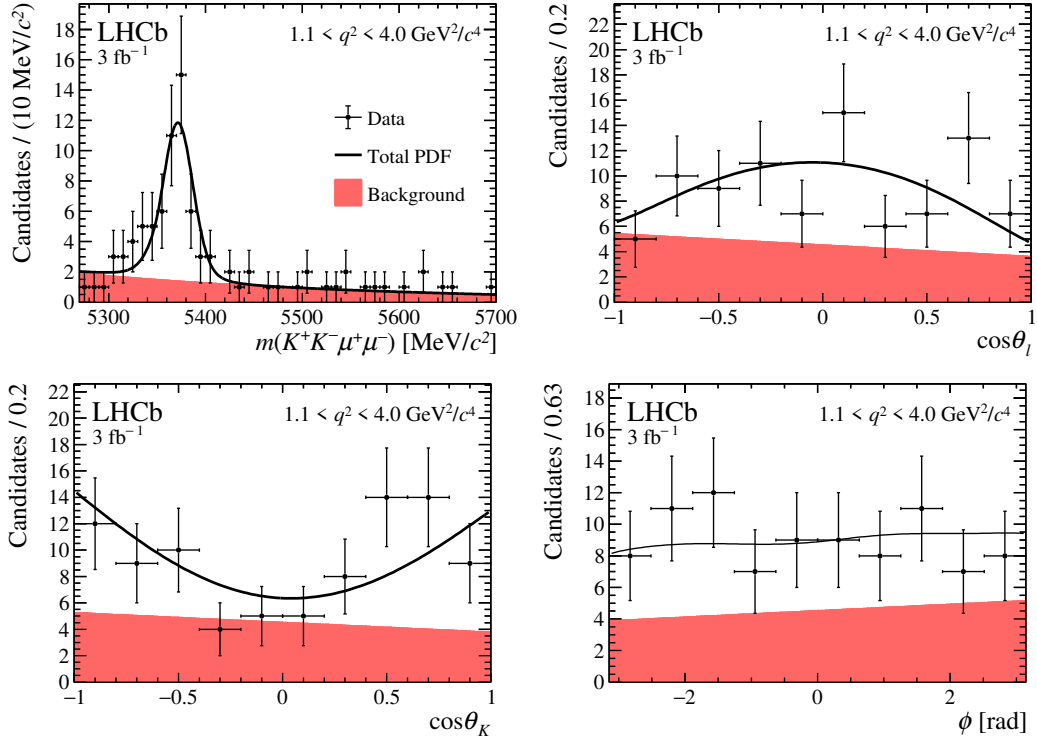


Figure F.4: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [1.1, 4.0] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

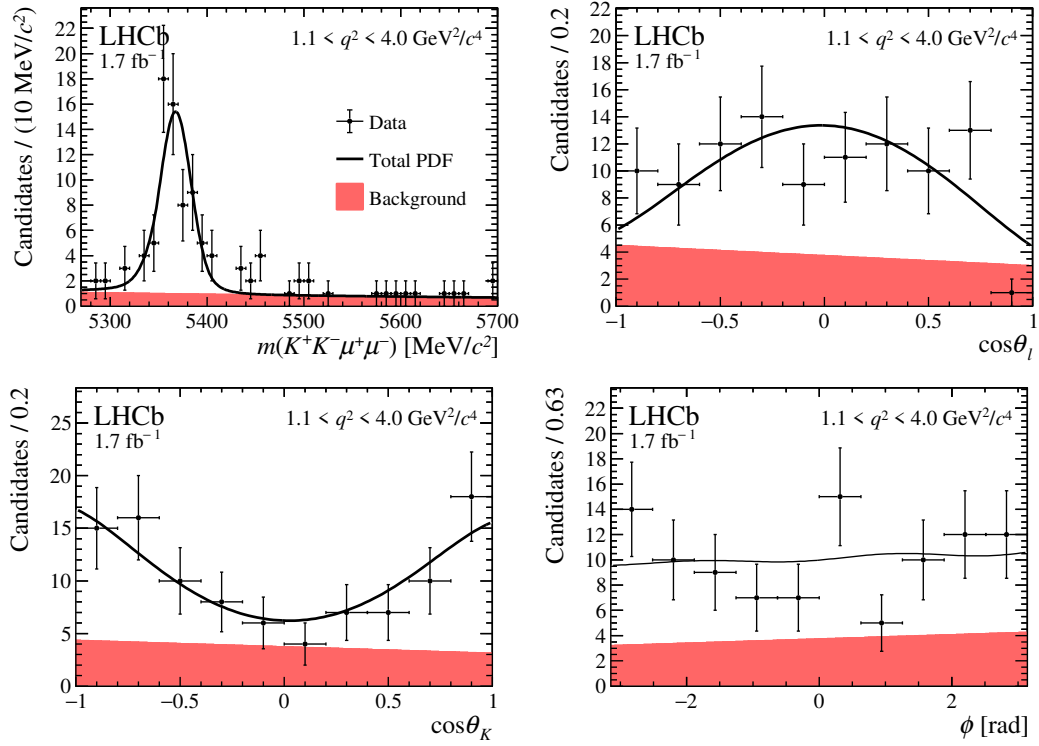


Figure F.5: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [1.1, 4.0] \text{ GeV}^2/c^4$  for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

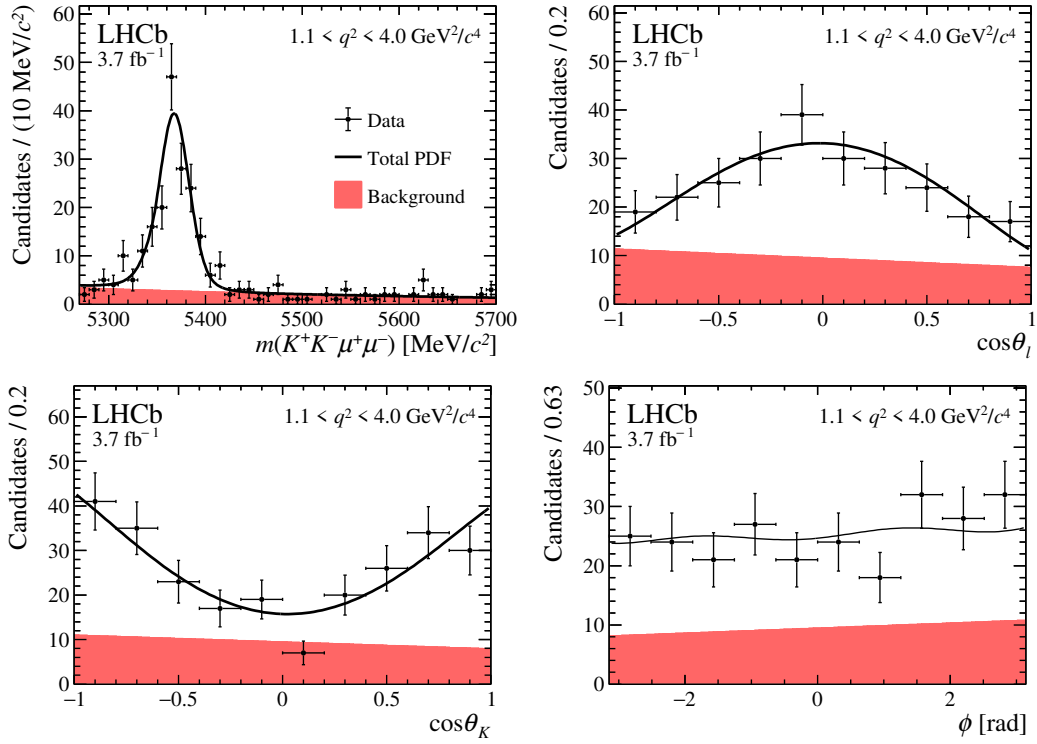


Figure F.6: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [1.1, 4.0] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

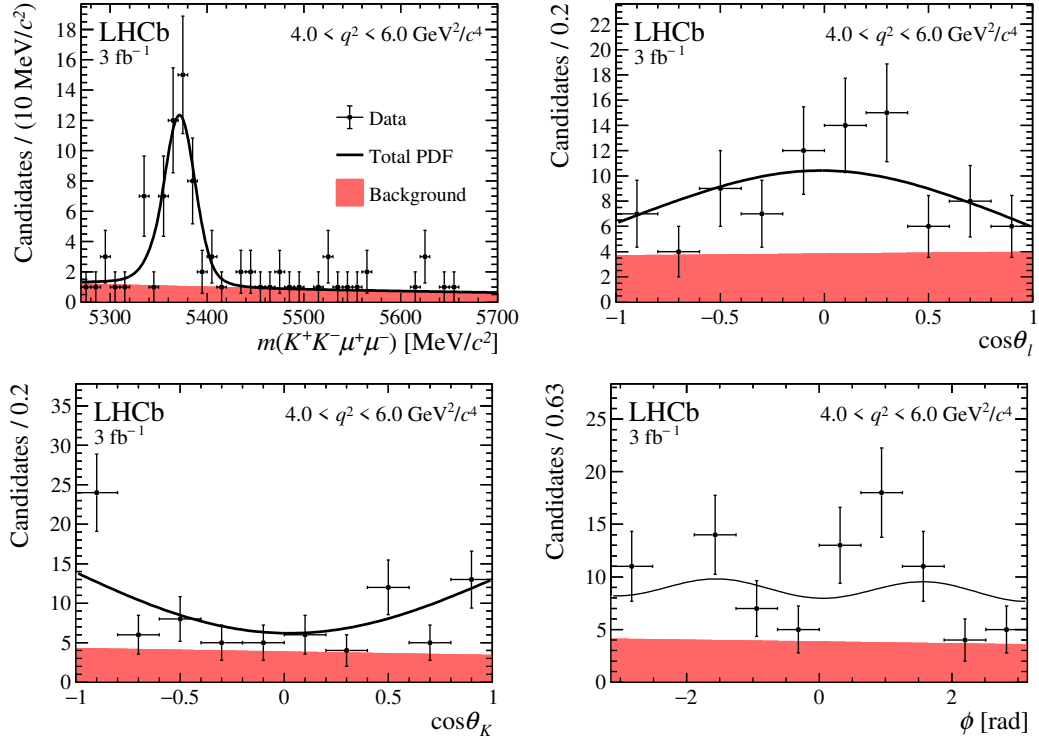


Figure F.7: Mass and angular distributions of  $B_s^0 \rightarrow \phi\mu^+\mu^-$  candidates in the region  $q^2 \in [4.0, 6.0] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

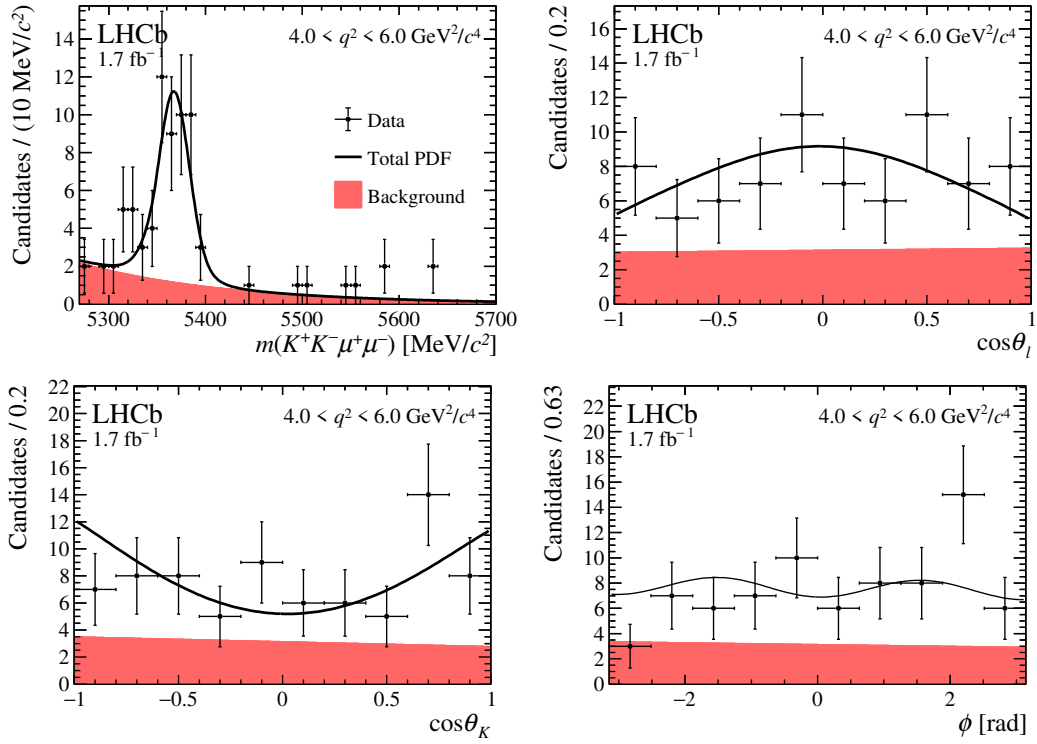


Figure F.8: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [4.0, 6.0] \text{ GeV}^2/c^4$  for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].



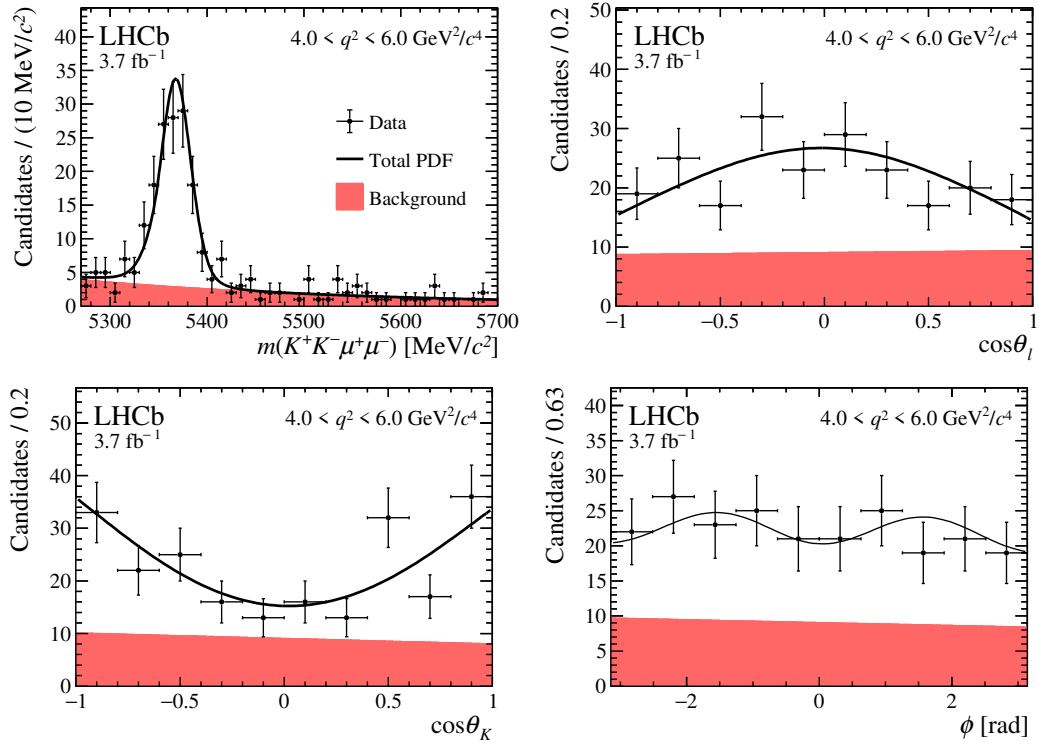


Figure F.9: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [4.0, 6.0] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

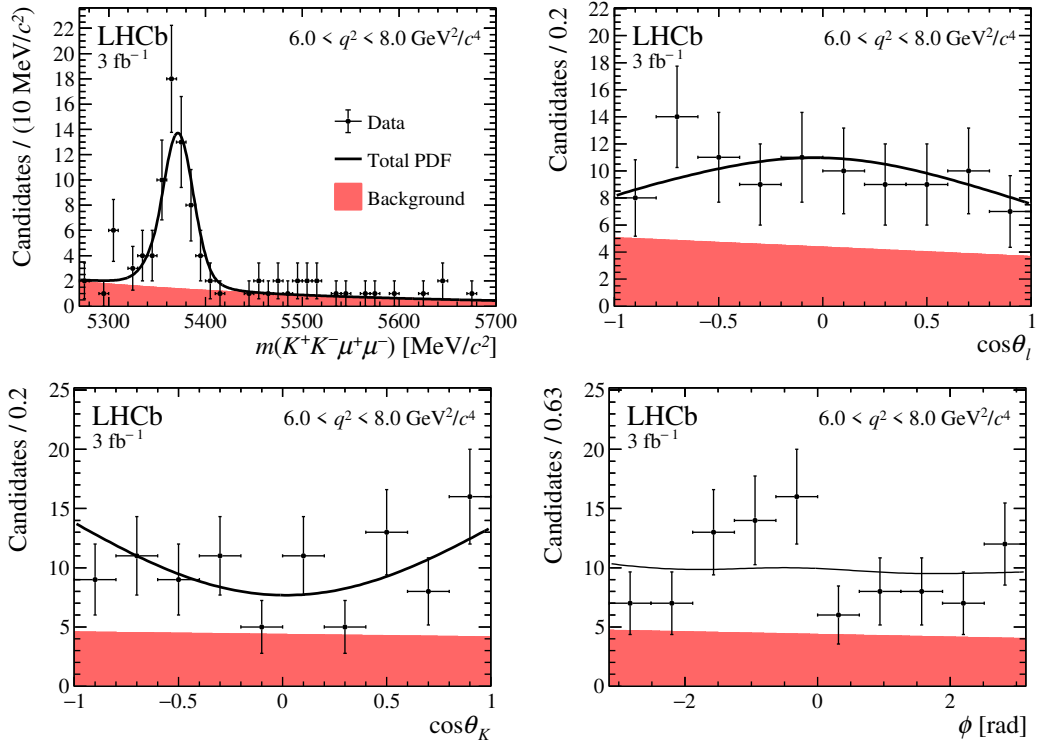


Figure F.10: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [6.0, 8.0] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

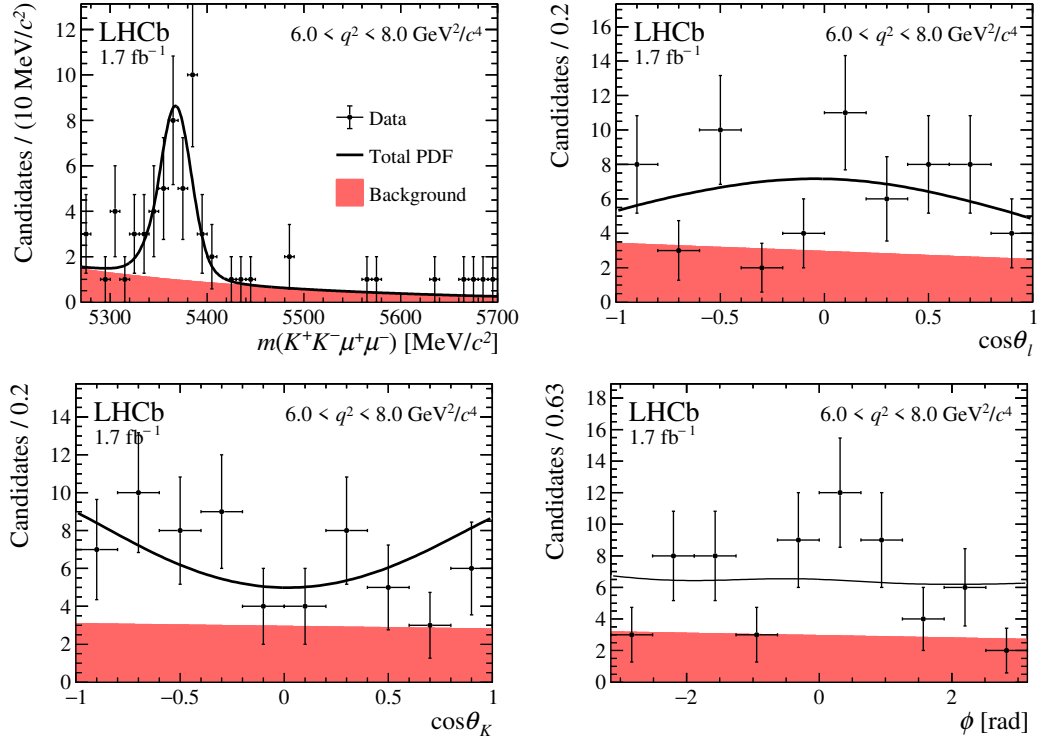


Figure F.11: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [6.0, 8.0]$  GeV<sup>2</sup>/c<sup>4</sup> for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

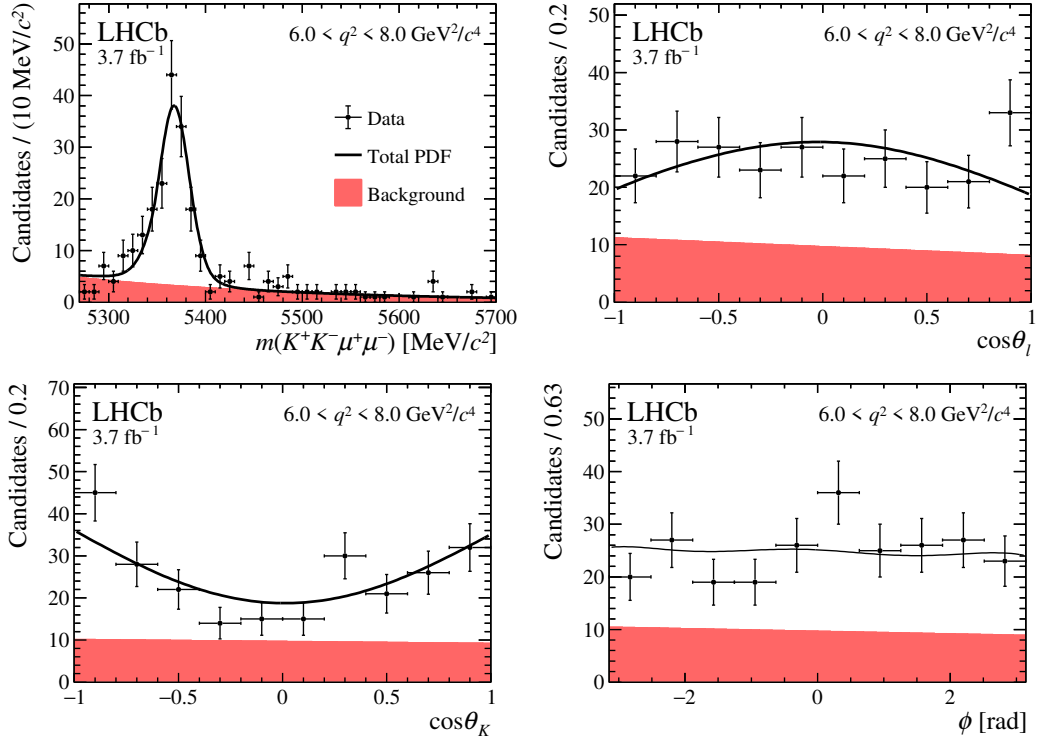


Figure F.12: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [6.0, 8.0] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

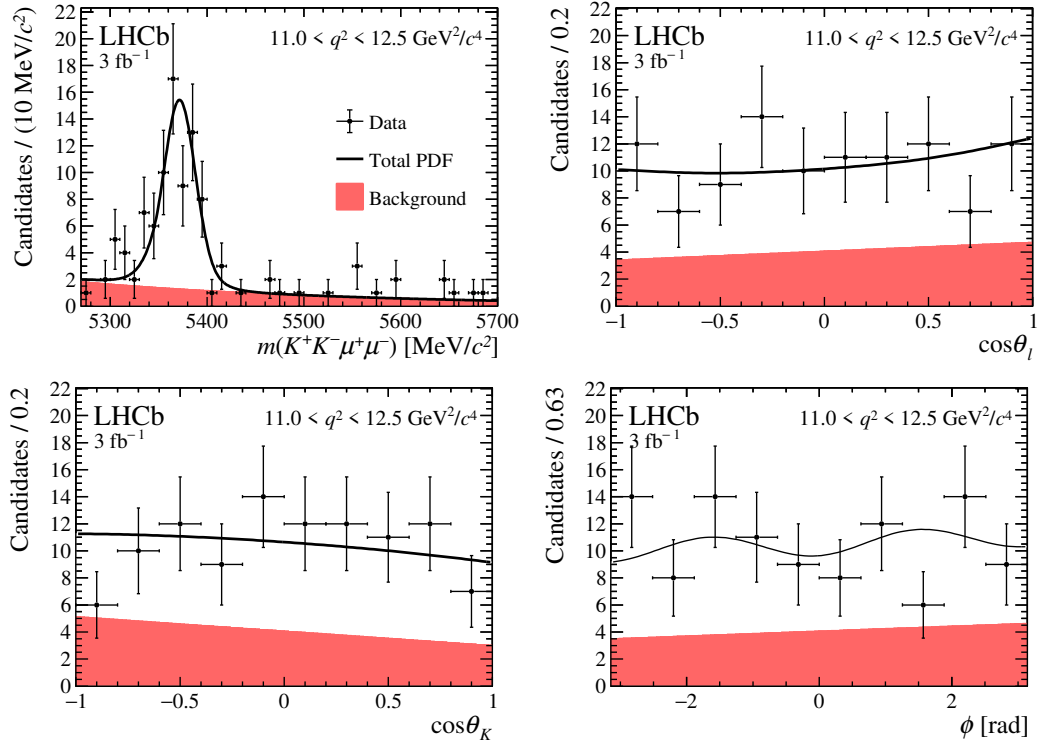


Figure F.13: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [11.0, 12.5] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

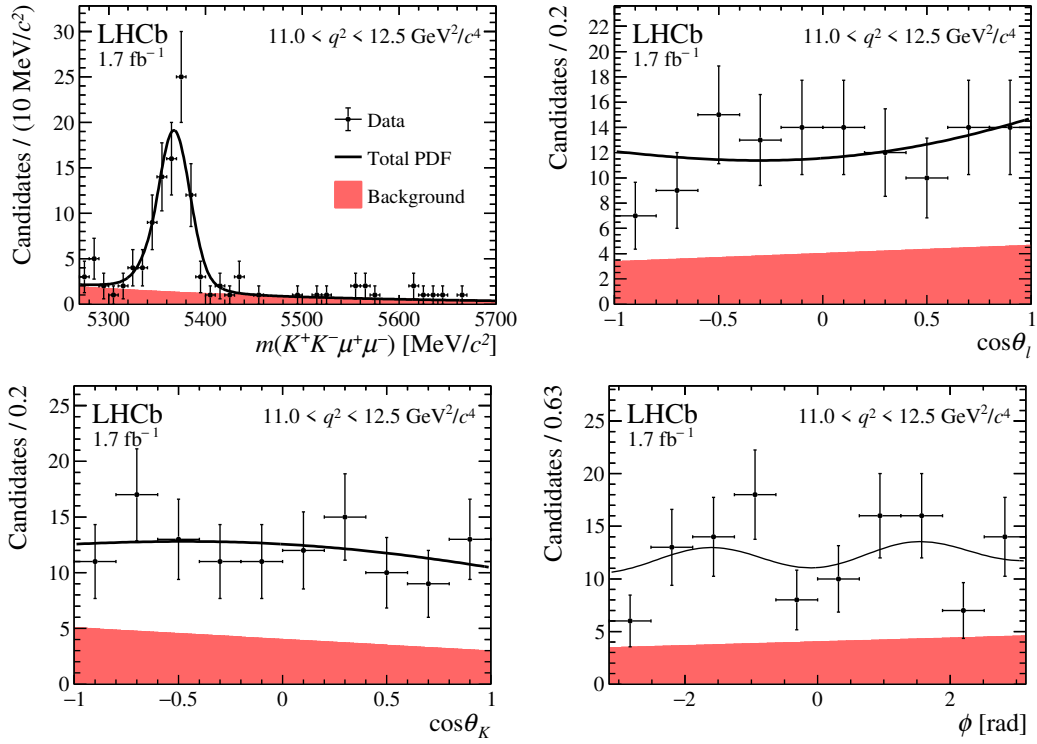


Figure F.14: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [11.0, 12.5]$  GeV<sup>2</sup>/c<sup>4</sup> for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

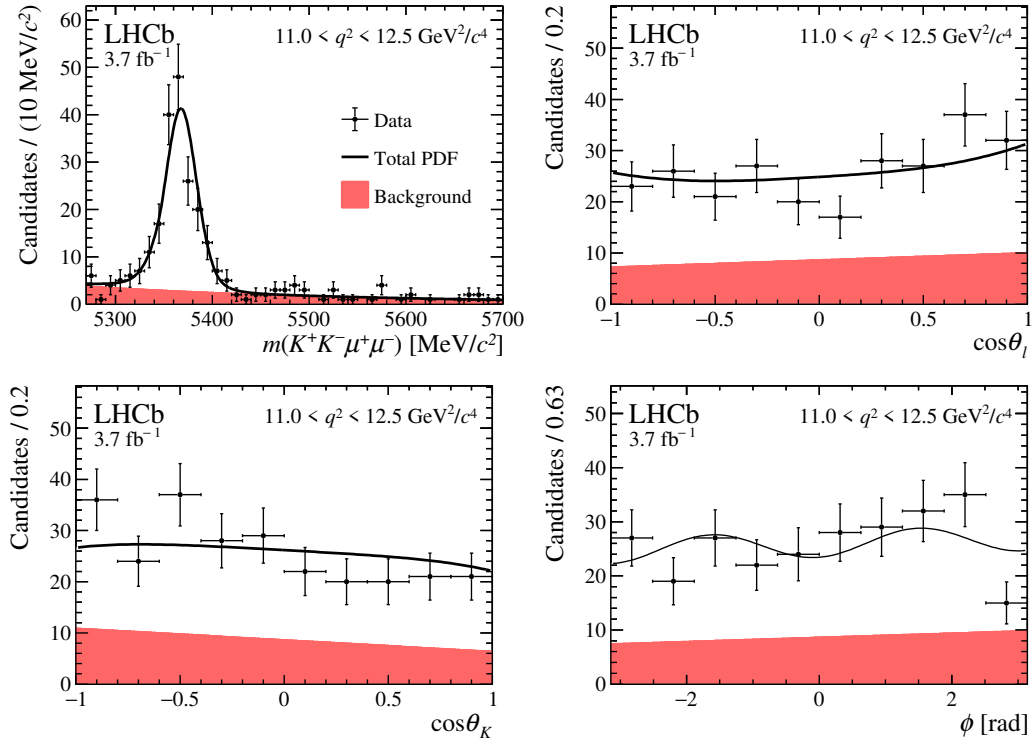


Figure F.15: Mass and angular distributions of  $B_s^0 \rightarrow \phi\mu^+\mu^-$  candidates in the region  $q^2 \in [11.0, 12.5] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

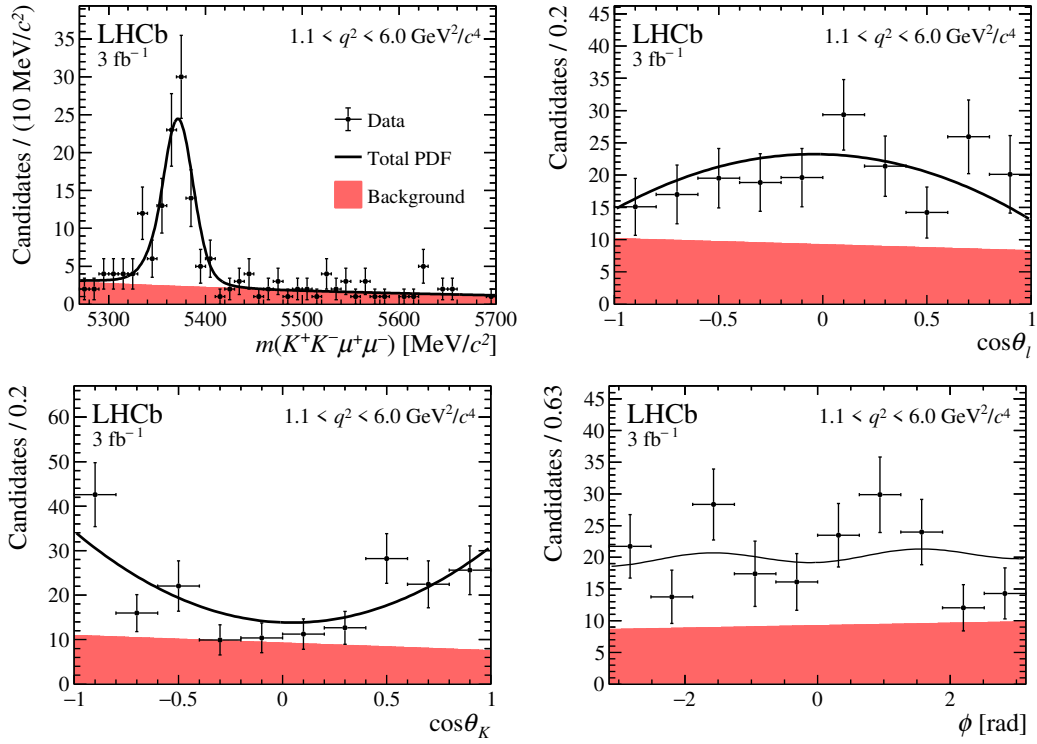


Figure F.16: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [1.1, 6.0] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].



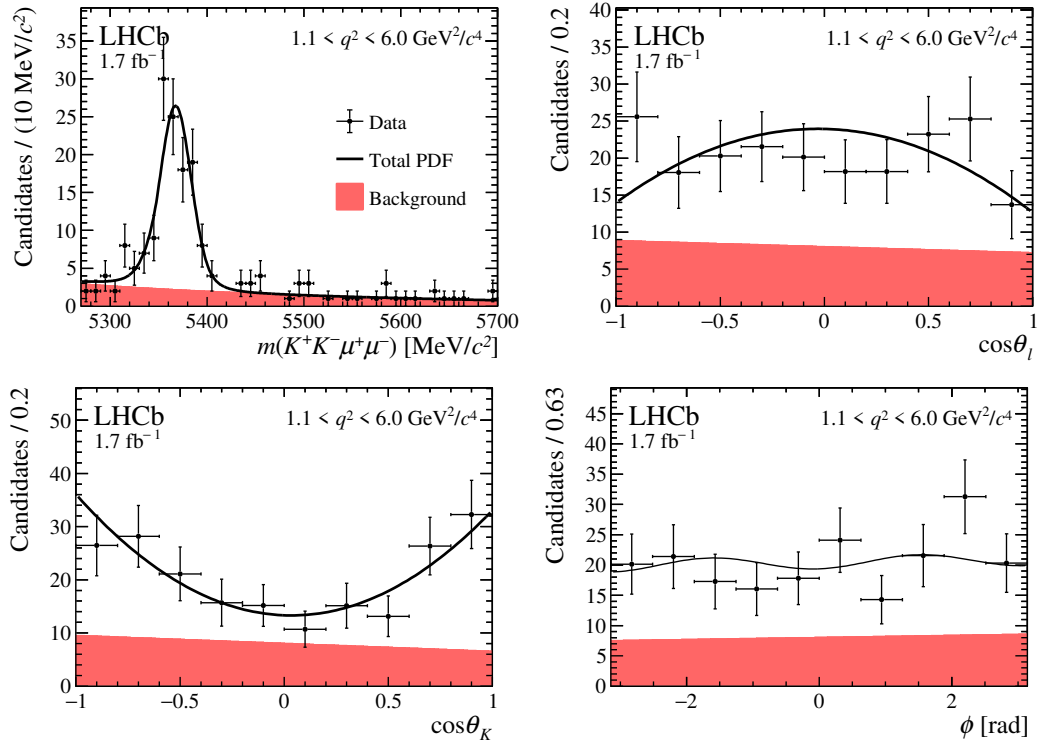


Figure F.17: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [1.1, 6.0] \text{ GeV}^2/c^4$  for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

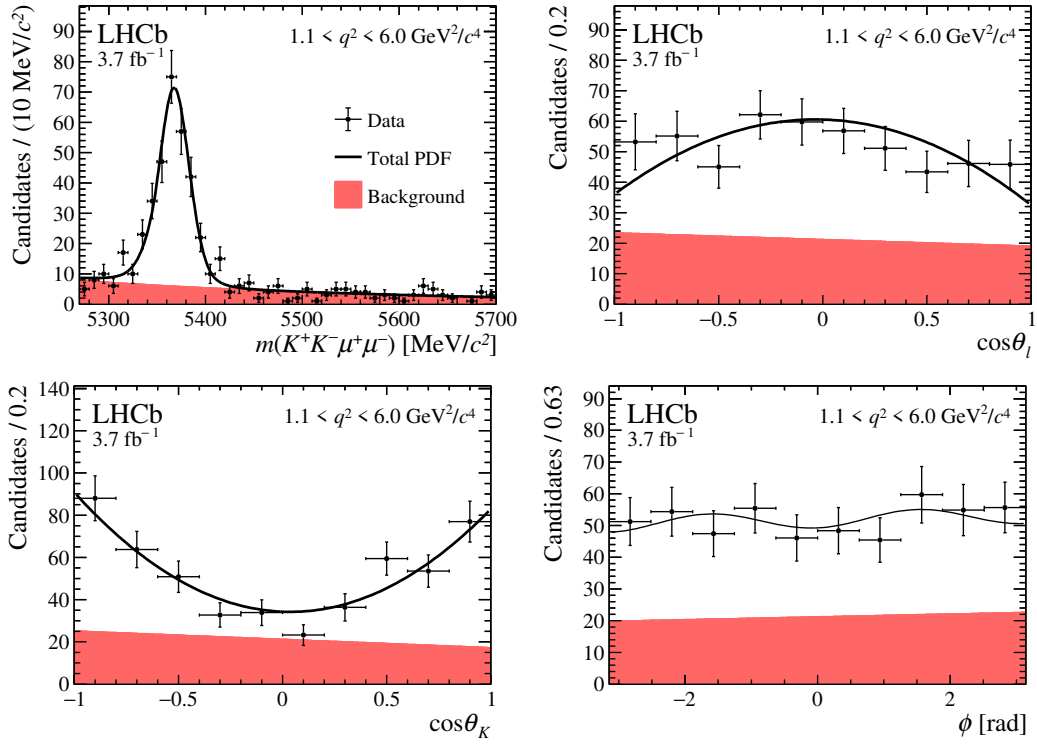


Figure F.18: Mass and angular distributions of  $B_s^0 \rightarrow \phi\mu^+\mu^-$  candidates in the region  $q^2 \in [1.1, 6.0] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

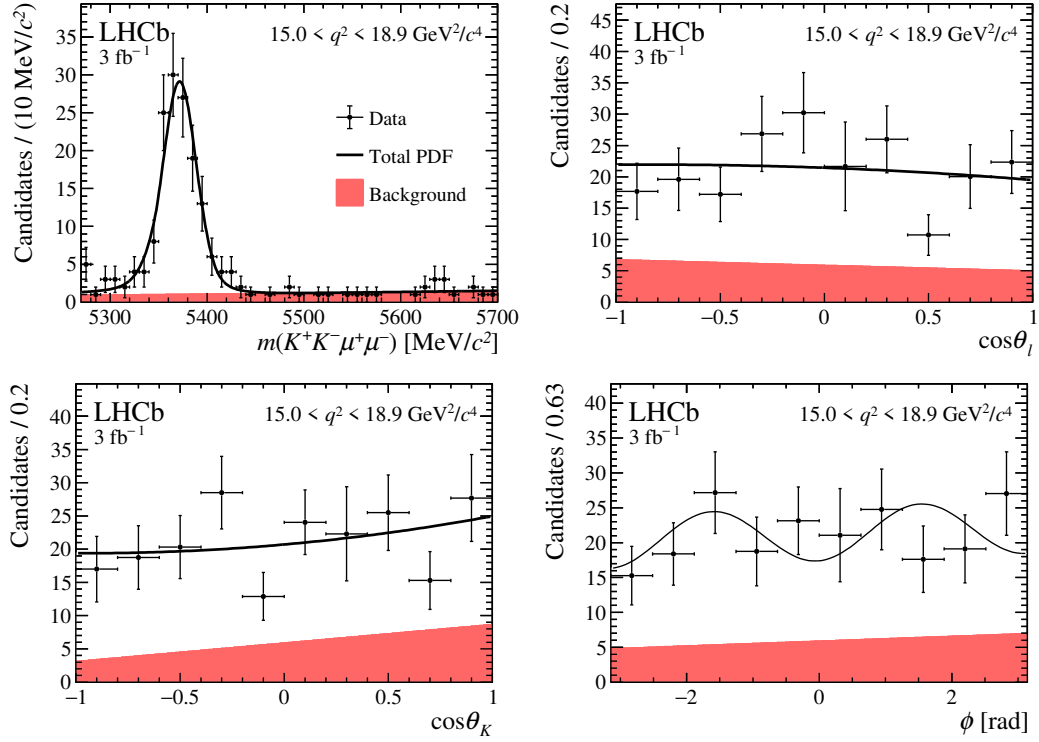


Figure F.19: Mass and angular distributions of  $B_s^0 \rightarrow \phi\mu^+\mu^-$  candidates in the region  $q^2 \in [15.0, 18.9] \text{ GeV}^2/c^4$  for data taken in 2011–2012. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

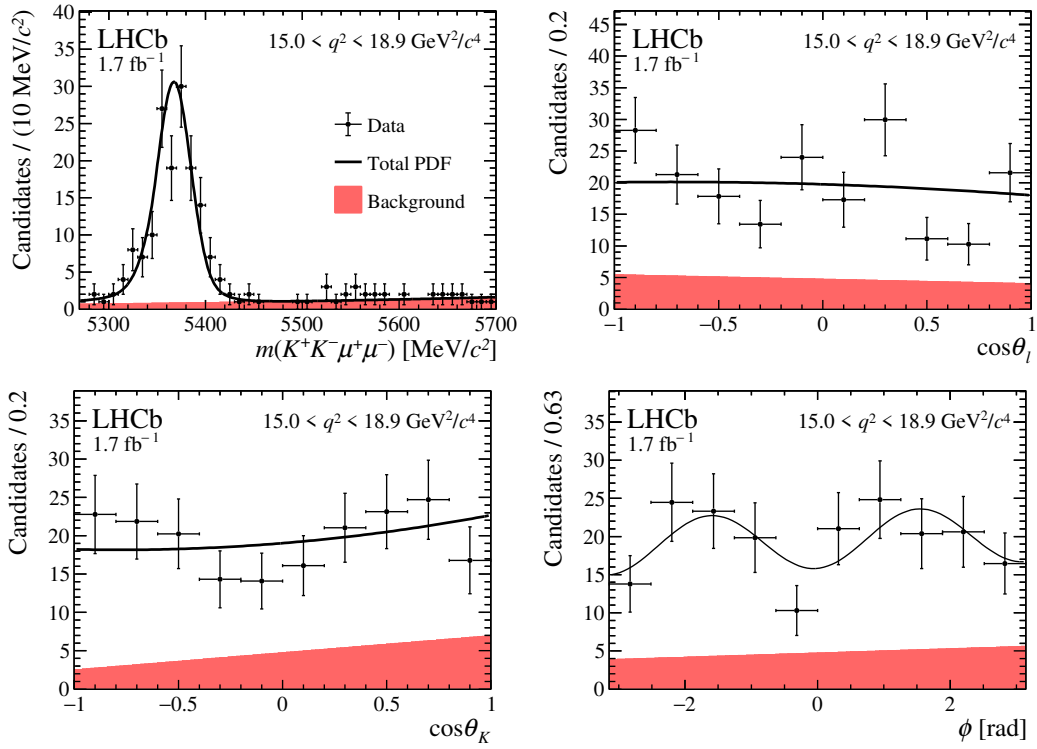


Figure F.20: Mass and angular distributions of  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  candidates in the region  $q^2 \in [15.0, 18.9] \text{ GeV}^2/c^4$  for data taken in 2016. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

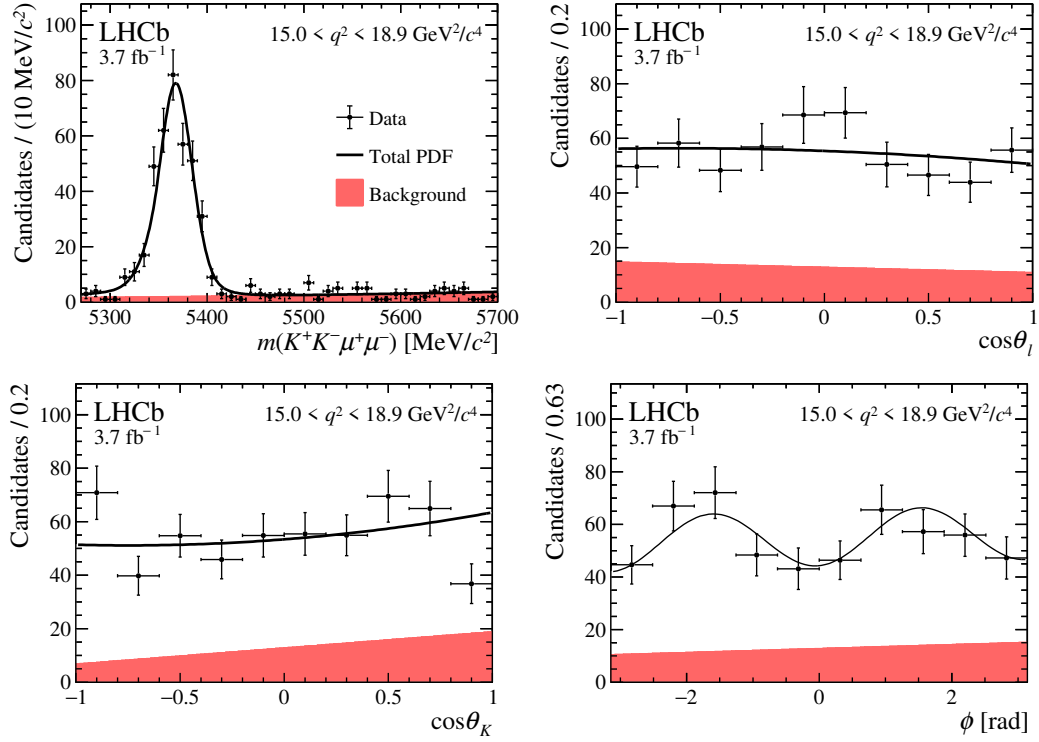


Figure F.21: Mass and angular distributions of  $B_s^0 \rightarrow \phi\mu^+\mu^-$  candidates in the region  $q^2 \in [15.0, 18.9] \text{ GeV}^2/c^4$  for data taken in 2017–2018. The data are overlaid with the projections of the fitted PDF. This figure also appeared in Ref. [46].

# Bibliography

- [1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.
- [2] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.
- [3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [4] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.
- [5] WMAP collaboration, C. L. Bennett *et al.*, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, Astrophys. J. Suppl. **208** (2013) 20, [arXiv:1212.5225](#).
- [6] WMAP collaboration, G. Hinshaw *et al.*, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*, Astrophys. J. Suppl. **208** (2013) 19, [arXiv:1212.5226](#).
- [7] Planck collaboration, N. Aghanim *et al.*, *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, Astron. Astrophys. **641** (2020) A1, [arXiv:1807.06205](#).
- [8] Planck collaboration, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, [arXiv:1807.06209](#), [Erratum: Astron.Astrophys. 652, C4 (2021)].
- [9] J. P. Ostriker and P. J. Steinhardt, *Cosmic concordance*, [arXiv:astro-ph/9505066](#).
- [10] N. A. Bahcall, J. P. Ostriker, S. Perlmutter, and P. J. Steinhardt, *The Cosmic triangle: Assessing the state of the universe*, Science **284** (1999) 1481, [arXiv:astro-ph/9906463](#).
- [11] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , JHEP **07** (2013) 084, [arXiv:1305.2168](#).
- [12] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using  $B^+ \rightarrow K^+ \ell^+ \ell^-$  decays*, Phys. Rev. Lett. **113** (2014) 151601, [arXiv:1406.6482](#).
- [13] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of  $B \rightarrow K^{(*)} \mu^+ \mu^-$  decays*, JHEP **06** (2014) 133, [arXiv:1403.8044](#).
- [14] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of  $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$  decays*, JHEP **06** (2015) 115, [arXiv:1503.07138](#), [Erratum: JHEP 09, 145 (2018)].

- [15] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , JHEP **09** (2015) 179, [arXiv:1506.08777](#).
- [16] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  decay using  $3 \text{ fb}^{-1}$  of integrated luminosity*, JHEP **02** (2016) 104, [arXiv:1512.04442](#).
- [17] LHCb collaboration, R. Aaij *et al.*, *Measurements of the S-wave fraction in  $B^0 \rightarrow K^+ \pi^- \mu^+ \mu^-$  decays and the  $B^0 \rightarrow K^*(892)^0 \mu^+ \mu^-$  differential branching fraction*, JHEP **11** (2016) 047, [arXiv:1606.04731](#), [Erratum: JHEP **04**, 142 (2017)].
- [18] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with  $B^0 \rightarrow K^{*0} \ell^+ \ell^-$  decays*, JHEP **08** (2017) 055, [arXiv:1705.05802](#).
- [19] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in  $B^+ \rightarrow K^+ \ell^+ \ell^-$  decays*, Phys. Rev. Lett. **122** (2019) 191801, [arXiv:1903.09252](#).
- [20] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with  $\Lambda_b^0 \rightarrow p K^- \ell^+ \ell^-$  decays*, JHEP **05** (2020) 040, [arXiv:1912.08139](#).
- [21] LHCb collaboration, R. Aaij *et al.*, *Angular Analysis of the  $B^+ \rightarrow K^{*+} \mu^+ \mu^-$  Decay*, Phys. Rev. Lett. **126** (2021) 161802, [arXiv:2012.13241](#).
- [22] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-Averaged Observables in the  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  Decay*, Phys. Rev. Lett. **125** (2020) 011802, [arXiv:2003.04831](#).
- [23] LHCb collaboration, R. Aaij *et al.*, *Tests of lepton universality using  $B^0 \rightarrow K_S^0 \ell^+ \ell^-$  and  $B^+ \rightarrow K^{*+} \ell^+ \ell^-$  decays*, [arXiv:2110.09501](#).
- [24] LHCb collaboration, R. Aaij *et al.*, *Branching Fraction Measurements of the Rare  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  and  $B_s^0 \rightarrow f_2'(1525) \mu^+ \mu^-$  Decays*, Phys. Rev. Lett. **127** (2021) 151801, [arXiv:2105.14007](#).
- [25] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, [arXiv:2103.11769](#).
- [26] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of  $B_d^0 \rightarrow K^* \mu^+ \mu^-$  decays in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$  with the ATLAS detector*, JHEP **10** (2018) 047, [arXiv:1805.04000](#).
- [27] CMS collaboration, V. Khachatryan *et al.*, *Angular analysis of the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  from pp collisions at  $\sqrt{s} = 8 \text{ TeV}$* , Phys. Lett. **B753** (2016) 424, [arXiv:1507.08126](#).
- [28] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  in proton-proton collisions at  $\sqrt{s} = 8 \text{ TeV}$* , Phys. Lett. **B781** (2018) 517, [arXiv:1710.02846](#).
- [29] Belle collaboration, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of  $B \rightarrow K^* \ell^+ \ell^-$* , Phys. Rev. Lett. **118** (2017) 111801, [arXiv:1612.05014](#).
- [30] BaBar collaboration, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays  $B \rightarrow K^{(*)} l^+ l^-$* , Phys. Rev. **D86** (2012) 032012, [arXiv:1204.3933](#).

- [31] Belle collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in  $B \rightarrow K\ell\ell$  decays*, JHEP **03** (2021) 105, [arXiv:1908.01848](#).
- [32] Belle collaboration, A. Abdesselam *et al.*, *Test of Lepton-Flavor Universality in  $B \rightarrow K^*\ell^+\ell^-$  Decays at Belle*, Phys. Rev. Lett. **126** (2021) 161801, [arXiv:1904.02440](#).
- [33] Belle collaboration, A. Abashian *et al.*, *The Belle Detector*, Nucl. Instrum. Meth. A **479** (2002) 117.
- [34] BaBar collaboration, B. Aubert *et al.*, *The BaBar detector*, Nucl. Instrum. Meth. A **479** (2002) 1, [arXiv:hep-ex/0105044](#).
- [35] BaBar collaboration, B. Aubert *et al.*, *The BABAR Detector: Upgrades, Operation and Performance*, Nucl. Instrum. Meth. A **729** (2013) 615, [arXiv:1305.3560](#).
- [36] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [37] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [38] LHCb collaboration, A. A. A. Jr, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [39] W. Altmannshofer and P. Stangl, *New physics in rare B decays after Moriond 2021*, Eur. Phys. J. C **81** (2021) 952, [arXiv:2103.13370](#).
- [40] T. Hurth, F. Mahmoudi, D. M. Santos, and S. Neshatpour, *More indications for lepton nonuniversality in  $b \rightarrow s\ell^+\ell^-$* , Phys. Lett. B **824** (2022) 136838, [arXiv:2104.10058](#).
- [41] M. Algueró *et al.*,  *$b \rightarrow s\ell\ell$  Global Fits after  $R_{K_S}$  and  $R_{K^{*+}}$* , [arXiv:2104.08921](#).
- [42] Muon g-2 collaboration, B. Abi *et al.*, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021) 141801, [arXiv:2104.03281](#).
- [43] Muon g-2 collaboration, T. Albahri *et al.*, *Measurement of the anomalous precession frequency of the muon in the Fermilab Muon g-2 Experiment*, Phys. Rev. D **103** (2021) 072002, [arXiv:2104.03247](#).
- [44] Muon g-2 collaboration, T. Albahri *et al.*, *Magnetic-field measurement and analysis for the Muon g-2 Experiment at Fermilab*, Phys. Rev. A **103** (2021) 042208, [arXiv:2104.03201](#).
- [45] Muon g-2 collaboration, T. Albahri *et al.*, *Beam dynamics corrections to the Run-1 measurement of the muon anomalous magnetic moment at Fermilab*, Phys. Rev. Accel. Beams **24** (2021) 044002, [arXiv:2104.03240](#).
- [46] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the rare decay  $B_s^0 \rightarrow \phi\mu^+\mu^-$* , JHEP **11** (2021) 043, [arXiv:2107.13428](#).
- [47] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.



- [48] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: a practical introduction; 3rd ed.*, Graduate student series in physics, IOP, Bristol, 2003.
- [49] D. Griffiths, *Introduction to elementary particles*, Wiley VCH, 2008.
- [50] G. ALTARELLI, *The Standard model of particle physics*, [arXiv:hep-ph/0510281](#).
- [51] U. Nierste, *Three Lectures on Meson Mixing and CKM phenomenology*, in *Helmholz International Summer School on Heavy Quark Physics*, 1–38, 2009, [arXiv:0904.1869](#).
- [52] Y. Grossman, *Introduction to flavor physics*, in *2009 European School of High-Energy Physics*, 111–144, 2010, [arXiv:1006.3534](#).
- [53] T. Blake, T. Gershon, and G. Hiller, *Rare  $b$  hadron decays at the LHC*, *Ann. Rev. Nucl. Part. Sci.* **65** (2015) 113, [arXiv:1501.03309](#).
- [54] S. Weinberg, *The Quantum Theory of Fields*, vol. 2, Cambridge University Press, 1996.
- [55] C. Bobeth, G. Hiller, and G. Piranishvili,  *$CP$  Asymmetries in  $\bar{B} \rightarrow \bar{K}^*(\rightarrow \bar{K}\pi)\bar{\ell}\ell$  and Untagged  $\bar{B}_s$ ,  $B_s \rightarrow \phi(\rightarrow K^+K^-)\bar{\ell}\ell$  Decays at NLO*, *JHEP* **0807** (2008) 106, [arXiv:0805.2525](#).
- [56] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of  $B \rightarrow K^*\mu^+\mu^-$  Decays in the Standard Model and Beyond*, *JHEP* **0901** (2009) 019, [arXiv:0811.1214](#).
- [57] S. Descotes-Genon and J. Virto, *Time dependence in  $B \rightarrow V\ell\ell$  decays*, *JHEP* **04** (2015) 045, [arXiv:1502.05509](#), [Erratum: *JHEP* **07**, 049 (2015)].
- [58] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [59] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [60] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [61] ATLAS, CMS collaboration, G. Aad *et al.*, *Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803, [arXiv:1503.07589](#).
- [62] S. Weinberg, *The Making of the standard model*, *Eur. Phys. J. C* **34** (2004) 5, [arXiv:hep-ph/0401010](#).
- [63] M. S. Lie, *Die Grundlagen für die Theorie der unendlichen kontinuierlichen Transformationsgruppen. I. Treatise*, *Leipz. Ber.* **3** (1891) 316.
- [64] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, *Phys. Rev.* **105** (1957) 1413.
- [65] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, *Prog. Theor. Exp. Phys.* **2020** (2020) 083C01.

- [66] B. R. Webber, *Hadronization*, in *Summer School on Hadronic Aspects of Collider Physics*, 49–77, 1994, [arXiv:hep-ph/9411384](#).
- [67] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV*, Phys. Lett. B **122** (1983) 103.
- [68] CDF collaboration, T. Affolder *et al.*, *Measurement of the  $W$  boson mass with the Collider Detector at Fermilab*, Phys. Rev. D **64** (2001) 052001, [arXiv:hep-ex/0007044](#).
- [69] ATLAS collaboration, M. Aaboud *et al.*, *Measurement of the  $W$ -boson mass in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 110, [arXiv:1701.07240](#), [Erratum: Eur.Phys.J.C 78, 898 (2018)].
- [70] LHCb collaboration, R. Aaij *et al.*, *Measurement of the  $W$  boson mass*, JHEP **01** (2022) 036, [arXiv:2109.01113](#).
- [71] UA1 collaboration, C. Albajar *et al.*, *Studies of Intermediate Vector Boson Production and Decay in UA1 at the CERN Proton - Antiproton Collider*, Z. Phys. C **44** (1989) 15.
- [72] ALEPH collaboration, R. Barate *et al.*, *Measurement of the  $Z$  resonance parameters at LEP*, Eur. Phys. J. C **14** (2000) 1.
- [73] OPAL collaboration, G. Abbiendi *et al.*, *Precise determination of the  $Z$  resonance parameters at LEP: 'Zedometry'*, Eur. Phys. J. C **19** (2001) 587, [arXiv:hep-ex/0012018](#).
- [74] H1 collaboration, V. Andreev *et al.*, *Determination of electroweak parameters in polarised deep-inelastic scattering at HERA*, Eur. Phys. J. C **78** (2018) 777, [arXiv:1806.01176](#).
- [75] K. G. Wilson, *Renormalization group and critical phenomena. ii. phase-space cell analysis of critical behavior*, Phys. Rev. B **4** (1971) 3184.
- [76] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.
- [77] M. Kobayashi and T. Maskawa,  *$CP$  Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [78] P. Ball *et al.*,  *$B$  Decays at the LHC*, in *Workshop on Standard Model Physics (and more) at the LHC (First Plenary Meeting)*, 2000, [arXiv:hep-ph/0003238](#).
- [79] *The CKM matrix and the unitarity triangle. Workshop, CERN, Geneva, Switzerland, 13-16 Feb 2002: Proceedings*, CERN Yellow Reports: Conference Proceedings, 2003. doi: 10.5170/CERN-2003-002-corr.
- [80] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.
- [81] A. De Roeck and R. S. Thorne, *Structure Functions*, Prog. Part. Nucl. Phys. **66** (2011) 727, [arXiv:1103.0555](#).
- [82] J. Blumlein, *The Theory of Deeply Inelastic Scattering*, Prog. Part. Nucl. Phys. **69** (2013) 28, [arXiv:1208.6087](#).

- [83] R. D. Ball *et al.*, *Parton Distribution Benchmarking with LHC Data*, JHEP **04** (2013) 125, [arXiv:1211.5142](#).
- [84] S. Forte and G. Watt, *Progress in the Determination of the Partonic Structure of the Proton*, Ann. Rev. Nucl. Part. Sci. **63** (2013) 291, [arXiv:1301.6754](#).
- [85] J. Gao, L. Harland-Lang, and J. Rojo, *The Structure of the Proton in the LHC Precision Era*, Phys. Rept. **742** (2018) 1, [arXiv:1709.04922](#).
- [86] C. Elsässer,  *$\bar{b}b$  production angle plots*, 2013. [Online; accessed 29 January 2021].
- [87] T. Sjöstrand, S. Mrenna, and P. Skands, *Pythia 6.4 physics and manual*, Journal of High Energy Physics **2006** (2006) 026–026.
- [88] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to pythia 8.1*, Computer Physics Communications **178** (2008) 852–867.
- [89] LHCb collaboration, R. Aaij *et al.*, *Measurement of the  $b$ -quark production cross-section in 7 and 13 TeV  $pp$  collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum *ibid.* **119** (2017) 169901, [arXiv:1612.05140](#).
- [90] M. Cacciari, M. L. Mangano, and P. Nason, *Gluon PDF constraints from the ratio of forward heavy-quark production at the LHC at  $\sqrt{S} = 7$  and 13 TeV*, 2015. doi: 10.1140/epjc/s10052-015-3814-x.
- [91] E. Fermi, *Versuch einer theorie der  $\beta$ -strahlen. i*, Zeitschrift für Physik **88** (1934) 161.
- [92] E. Fermi, *Tentativo di una teoria dei raggi  $\beta$* , Il Nuovo Cimento (1924-1942) **11** (1934) 1.
- [93] C. P. Burgess, *Introduction to Effective Field Theory*, Ann. Rev. Nucl. Part. Sci. **57** (2007) 329, [arXiv:hep-th/0701053](#).
- [94] K. G. Wilson and W. Zimmermann, *Operator product expansions and composite field operators in the general framework of quantum field theory*, Communications in Mathematical Physics **24** (1972) 87.
- [95] B. Grinstein, R. Springer, and M. B. Wise, *Strong-interaction effects in weak radiative  $b$ -meson decay*, Nuclear Physics B **339** (1990) 269.
- [96] K. G. Wilson, *Quantum field - theory models in less than 4 dimensions*, Phys. Rev. D **7** (1973) 2911.
- [97] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, [arXiv:hep-ph/9512380](#).
- [98] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive  $B \rightarrow V l^+ l^-$ ,  $V \gamma$  decays*, Nucl. Phys. B **612** (2001) 25, [arXiv:hep-ph/0106067](#).
- [99] V. M. Braun, *Light cone sum rules*, in *4th International Workshop on Progress in Heavy Quark Physics*, 105–118, 1997, [arXiv:hep-ph/9801222](#).

- [100] R. Gupta, *Introduction to lattice QCD: Course*, in *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*, 83–219, 1997, [arXiv:hep-lat/9807028](#).
- [101] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *Penguins with Charm and Quark-Hadron Duality*, *Eur. Phys. J. C* **61** (2009) 439, [arXiv:0902.4446](#).
- [102] M. Beylich, G. Buchalla, and T. Feldmann, *Theory of  $B \rightarrow K^{(*)}\ell^+\ell^-$  decays at high  $q^2$ : OPE and quark-hadron duality*, *Eur. Phys. J. C* **71** (2011) 1635, [arXiv:1101.5118](#).
- [103] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the  $B \rightarrow K^*\ell^+\ell^-$  decays in the large-recoil region: SM challenges and BSM opportunities*, *Phys. Rev. D* **93** (2016) 014028, [arXiv:1412.3183](#).
- [104] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in  $b \rightarrow s\ell^+\ell^-$ ?*, [arXiv:1406.0566](#).
- [105] M. Ciuchini *et al.*,  *$B \rightarrow K^*\ell^+\ell^-$  decays at large recoil in the Standard Model: a theoretical reappraisal*, *JHEP* **06** (2016) 116, [arXiv:1512.07157](#).
- [106] M. Ciuchini *et al.*, *On Flavourful Easter eggs for New Physics hunger and Lepton Flavour Universality violation*, *Eur. Phys. J. C* **77** (2017) 688, [arXiv:1704.05447](#).
- [107] N. Gubernari, D. van Dyk, and J. Virto, *Non-local matrix elements in  $B_{(s)} \rightarrow \{K^{(*)}, \phi\}\ell^+\ell^-$* , *JHEP* **02** (2021) 088, [arXiv:2011.09813](#).
- [108] C. Bobeth, G. Hiller, and D. van Dyk, *General analysis of  $\bar{B} \rightarrow \bar{K}^{(*)}\ell^+\ell^-$  decays at low recoil*, *Phys. Rev. D* **87** (2013) 034016, [arXiv:1212.2321](#).
- [109] D. M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, [arXiv:1810.08132](#).
- [110] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [111] M. A. Hone, *The duoplasmatron ion source for the new CERN Linac preinjector*, CERN, Geneva, 1979.
- [112] A. M. Lombardi, *The radio frequency quadrupole (RFQ)*, doi: 10.5170/CERN-2006-012.201.
- [113] D. J. Warner, *Project study for a new 50 MeV linear accelerator for the C. P. S.*, CERN-MPS-LINP-73-1, CERN, Geneva, 1973.
- [114] M. Benedikt *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004.
- [115] E. Mobs, *The CERN accelerator complex - August 2018*, 2018. [Online; accessed 28 December 2020].
- [116] R. Bailey and P. Collier, *Standard Filling Schemes for Various LHC Operation Modes*, CERN, Geneva, 2003.

- [117] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [118] W. Herr and B. Muratori, *Concept of luminosity*, doi: 10.5170/CERN-2006-002.361.
- [119] C. Pralavorio, *Record luminosity: well done LHC*, 2017. [Online; accessed 1 March 2020].
- [120] LHCb collaboration, H. Dijkstra, H. J. Hilke, T. Nakada, and T. Ypsilantis, *LHCb Letter of Intent*, LHCb-95-001, CERN, Geneva, 1995.
- [121] LHCb collaboration, R. Antunes-Nobrega *et al.*, *LHCb reoptimized detector design and performance: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003.
- [122] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, EPJ Tech. Instrum. **6** (2019) 1, arXiv:1803.00824.
- [123] LHCb collaboration, R. Aaij *et al.*, *Precision luminosity measurements at LHCb*, JINST **9** (2014) P12005, arXiv:1410.0149.
- [124] A. Iliencko, *Continuous counterparts of poisson and binomial distributions and their properties*, 2013.
- [125] B. Muratori and T. Pieloni, *Luminosity levelling techniques for the LHC*, doi: 10.5170/CERN-2014-004.177, Comments: 5 pages, contribution to the ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders, CERN, Geneva, Switzerland, 18-22 Mar 2013.
- [126] J. Liouville, *Note sur la théorie de la variation des constantes arbitraires.*, Journal de Mathématiques Pures et Appliquées (1838) 342.
- [127] R. B. Palmer, *Energy Scaling, Crab Crossing, and the Pair Problem*, eConf **C8806271** (1988) 613.
- [128] R. Alemany-Fernandez, F. Follin, and R. Jacobsson, *The LHCb Online Luminosity Control and Monitoring*, in *4th International Particle Accelerator Conference*, TUPFI010, 2013.
- [129] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb VELO (VERtex LOCator): Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.
- [130] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb inner tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2002. revised version number 1 submitted on 2002-11-13 14:14:34.
- [131] M. Pepe Altarelli and F. Teubert, *B Physics at LHCb*, Int. J. Mod. Phys. **A23** (2008) 5117, arXiv:0802.1901, [,299(2008)].
- [132] LHCb collaboration, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (2015) 1530022, arXiv:1412.6352.

- [133] G. Von Holtey, *Aperture requirements around interaction IP8* , CERN, 1997.
- [134] R. Aaij *et al.*, *Performance of the lhcb vertex locator*, Journal of Instrumentation **9** (2014) P09007–P09007.
- [135] U. of Zürich for the LHCb Silicon Tracker Group, *LHCb Silicon Tracker Home Page*, 2013. [Online; accessed 28 December 2020].
- [136] A. Bay *et al.*, *Hybrid design, procurement and testing for the LHCb silicon tracker*, in *11th Workshop on Electronics for LHC and Future Experiments (LECC 2005)*, 10, 2005.
- [137] LHCb collaboration, S. Amato *et al.*, *LHCb magnet: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000.
- [138] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb outer tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.
- [139] J. Van Tilburg, *Matching VELO tracks with seeding tracks* , CERN, Geneva, 2001. revised version number 1 submitted on 2001-08-08 09:00:15.
- [140] R. W. Forty, *Track Seeding* , CERN, Geneva, 2001. revised version number 1 submitted on 2001-10-12 16:38:05.
- [141] R. Hierck, *Track following for LHCb* , CERN, Geneva, 2001.
- [142] M. Benayoun and O. Callot, *The forward tracking, an optical model method* , CERN, Geneva, 2002. revised version number 1 submitted on 2002-02-22 17:19:02.
- [143] Y. Xie, *Short track reconstruction with VELO and TT* , CERN, Geneva, 2003.
- [144] O. Callot, M. Kucharczyk, and M. Witek, *VELO-TT track reconstruction* , CERN, Geneva, 2007.
- [145] D. Hutchcroft, *VELO Pattern Recognition* , CERN, Geneva, 2007.
- [146] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies* , CERN, Geneva, 2007.
- [147] R. W. Forty and M. Needham, *Updated Performance of the T-Seeding* , CERN, Geneva, 2007.
- [148] O. Callot, *Downstream Pattern Recognition* , CERN, Geneva, 2007.
- [149] M. Needham, *Performance of the Track Matching* , CERN, Geneva, 2007.
- [150] M. Needham, *Performance of the LHCb Track Reconstruction Software* , CERN, Geneva, 2008.
- [151] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, [arXiv:1408.1251](#).

- [152] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Journal of Basic Engineering **82** (1960) 35, [arXiv:https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/82/1/35](https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/82/1/35/1134001)
- [153] LHCb collaboration, S. Amato *et al.*, *LHCb RICH: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000.
- [154] P. A. Čerenkov, *Visible radiation produced by electrons moving in a medium with velocities exceeding that of light*, Phys. Rev. **52** (1937) 378.
- [155] LHCb RICH collaboration, 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](https://arxiv.org/abs/1703.08152).
- [156] LHCb collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000.
- [157] LHCb collaboration, R. Antunes-Nobrega *et al.*, *LHCb trigger system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003. revised version number 1 submitted on 2003-09-24 12:12:22.
- [158] C. Abellán Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, [arXiv:2008.11556](https://arxiv.org/abs/2008.11556).
- [159] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.
- [160] LHCb collaboration, B. Schmidt *et al.*, *LHCb muon system: addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003.
- [161] LHCb collaboration, B. Schmidt *et al.*, *LHCb muon system: second addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2005. Submitted on 9 Apr 2005.
- [162] G. Charpak *et al.*, *The Use of Multiwire Proportional Counters to Select and Localize Charged Particles*, Nucl. Instrum. Meth. **62** (1968) 262.
- [163] B. Botchine *et al.*, *Wire pad chamber for LHCb muon system*, LHCb-2000-003, CERN, Geneva, 2000.
- [164] B. Botchine *et al.*, *Wire pad chambers and cathode pad chambers for the LHCb muon system*, LHCb-2000-114, CERN, Geneva, 2002.
- [165] F. Sauli, *Gem: A new concept for electron amplification in gas detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **386** (1997) 531 .
- [166] M. Alfonsi *et al.*, *Advances in fast multi-gem based detector operation for high-rate charged-particle triggering*, in *2003 IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No.03CH37515)*, **1** 247–251 Vol.1, 2003.

- [167] E. Aslanides *et al.*, *The Level-0 muon trigger for the LHCb experiment*, Nucl. Instrum. Meth. A **579** (2007) 989, [arXiv:0705.0310](#).
- [168] ATLAS collaboration, G. Aad *et al.*, *Measurement of the total cross section from elastic scattering in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, Nucl. Phys. B **889** (2014) 486, [arXiv:1408.5778](#).
- [169] ATLAS collaboration, M. Aaboud *et al.*, *Measurement of the total cross section from elastic scattering in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, Phys. Lett. B **761** (2016) 158, [arXiv:1607.06605](#).
- [170] TOTEM collaboration, G. Antchev *et al.*, *First measurement of elastic, inelastic and total cross-section at  $\sqrt{s} = 13$  TeV by TOTEM and overview of cross-section data at LHC energies*, Eur. Phys. J. C **79** (2019) 103, [arXiv:1712.06153](#).
- [171] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb online system, data acquisition and experiment control: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.
- [172] LHCb collaboration, *Trigger Schemes*, 2015. [Online; accessed 17 November 2021].
- [173] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, [arXiv:1211.3055](#).
- [174] LHCb collaboration, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, [arXiv:1812.10790](#).
- [175] T. Head, *The LHCb trigger system*, Journal of Instrumentation **9** (2014) C09015.
- [176] LHCb collaboration, J. Albrecht, *The LHCb Trigger System: Present and Future*, J. Phys. Conf. Ser. **623** (2015) 012003.
- [177] G. Cowan, *Statistical Data Analysis*, Oxford science publications, Clarendon Press, 1998.
- [178] M. Pivk and F. R. Le Diberder, *Plots: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005) 356 .
- [179] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- [180] H. Drucker and C. Cortes, *Boosting decision trees*, in *Advances in neural information processing systems*, **8**, 479–485, 1995.
- [181] B. P. Roe *et al.*, *Boosted decision trees, an alternative to artificial neural networks*, Nucl. Instrum. Meth. A **543** (2005) 577, [arXiv:physics/0408124](#).
- [182] J. R. Quinlan, *Induction of decision trees*, Machine Learning **1** (1986) 81.



- [183] H. Drucker *et al.*, *Boosting and other machine learning algorithms*, in *Machine Learning Proceedings 1994* (W. W. Cohen and H. Hirsh, eds.), pp. 53 – 61. Morgan Kaufmann, San Francisco (CA), 1994. doi: <https://doi.org/10.1016/B978-1-55860-335-6.50015-5>.
- [184] 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) 119.
- [185] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: bounds for k-fold and progressive cross-validation*, in *COLT '99*, 1999.
- [186] T. Bayes, Rev. *An essay toward solving a problem in the doctrine of chances*, Phil. Trans. Roy. Soc. Lond. **53** (1764) 370.
- [187] F. James and M. Roos, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, Computer Physics Communications **10** (1975) 343 .
- [188] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, in *CERN School of Computing (CSC 2000)*, 11–42, 2000.
- [189] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.
- [190] C. Langenbruch, *Parameter uncertainties in weighted unbinned maximum likelihood fits*, Eur. Phys. J. C **82** (2022) 393, [arXiv:1911.01303](https://arxiv.org/abs/1911.01303).
- [191] R. J. Barlow, *Extended maximum likelihood*, Nucl. Instrum. Meth. A **297** (1990) 496.
- [192] R. L. G. F. Ålessio, R. Jacobsson, *Integrated recorded luminosity*, 2019. [Online; accessed 29 January 2021].
- [193] R. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, Wiley, 2017.
- [194] G. Barrand *et al.*, *Gaudi — a software architecture and framework for building hep data processing applications*, Computer Physics Communications **140** (2001) 45 , CHEP2000.
- [195] M. Clemencic *et al.*, *Recent developments in the LHCb software framework gaudi*, Journal of Physics: Conference Series **219** (2010) 042006.
- [196] LHCb collaboration, *The GAUSS Project*, 2019. [Online; accessed 1 March 2020].
- [197] M. Clemencic *et al.*, *The LHCb simulation application, gauss: Design, evolution and experience*, Journal of Physics: Conference Series **331** (2011) 032023.
- [198] LHCb collaboration, *The BOOLE Project*, 2019. [Online; accessed 1 March 2020].
- [199] LHCb collaboration, *The MOORE Project*, 2018. [Online; accessed 1 March 2020].
- [200] LHCb collaboration, *The BRUNEL Project*, 2019. [Online; accessed 1 March 2020].
- [201] LHCb collaboration, *The DAVINCI Project*, 2019. [Online; accessed 1 March 2020].
- [202] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to pythia 8.1*, Computer Physics Communications **178** (2008) 852 .

- [203] T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, [arXiv:1410.3012](#).
- [204] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [205] D. J. Lange, *The evtgen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001) 152 , BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.
- [206] 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](#).
- [207] S. Agostinelli *et al.*, *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250 .
- [208] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [209] J. Allison *et al.*, *Recent developments in geant4*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **835** (2016) 186 .
- [210] K. Bos *et al.*, *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, Technical Design Report LCG, CERN, Geneva, 2005.
- [211] I. Bird *et al.*, *Update of the Computing Models of the WLCG and the LHC Experiments*, CERN-LHCC-2014-014, LCG-TDR-002, CERN, 2014.
- [212] A. Ali, E. Lunghi, C. Greub, and G. Hiller, *Improved model-independent analysis of semileptonic and radiative rare b decays*, Phys. Rev. D **66** (2002) 034002.
- [213] P. Ball and R. Zwicky,  *$B_{d,s} \rightarrow \rho, \omega, K^*, \phi$  decay form factors from light-cone sum rules reexamined*, Phys. Rev. D **71** (2005) 014029.
- [214] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, [arXiv:1306.0249](#).
- [215] LHCb collaboration, D. Derkach *et al.*, *Machine-Learning-based global particle-identification algorithms at the LHCb experiment*, J. Phys. Conf. Ser. **1085** (2018) 042038.
- [216] LHCb collaboration, D. Derkach, M. Hushchyn, and N. Kazeev, *Machine Learning based Global Particle Identification Algorithms at the LHCb Experiment*, EPJ Web Conf. **214** (2019) 06011.
- [217] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction* , CERN, Geneva, 2017.

- [218] LHCb collaboration, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, [arXiv:1812.10790](#).
- [219] M. Williams *et al.*, *The HLT2 Topological Lines*, CERN, Geneva, 2011.
- [220] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, CERN, Geneva, 2011. LHCb-INT-2011-030.
- [221] T. Likhomanenko *et al.*, *LHCb Topological Trigger Reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025, [arXiv:1510.00572](#).
- [222] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, PoS **ACAT** (2007) 040.
- [223] A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.*, [arXiv:physics/0703039](#).
- [224] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
- [225] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015) P02011, [arXiv:1411.5528](#).
- [226] A. Poluektov, *Correction of simulated particle identification response in LHCb using kernel density estimation.*, LHCb-INT-2017-007. CERN-LHCb-INT-2017-007, CERN, Geneva, 2017.
- [227] A. Poluektov, *Unbinned PID resampling using kernel density estimation*, 2020. twiki page, [Online; accessed 17 November 2021].
- [228] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, 2014.
- [229] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, 2016.
- [230] HFLAV collaboration, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and  $\tau$ -lepton properties as of summer 2016*, Eur. Phys. J. C **77** (2017) 895, [arXiv:1612.07233](#).
- [231] L. MOTT and W. ROBERTS, *Rare dileptonic decays of  $\Lambda_b^0$  in a quark model*, International Journal of Modern Physics A **27** (2012) 1250016.
- [232] LHCb collaboration, R. Aaij *et al.*, *Observation of  $J/\psi$  Resonances Consistent with Pentaquark States in  $\Lambda_b^0 \rightarrow J/\psi K^- p$  Decays*, Phys. Rev. Lett. **115** (2015) 072001, [arXiv:1507.03414](#).
- [233] L. Mott and W. Roberts, *Rare dileptonic decays of  $\Lambda_b$  in a quark model*, Int. J. Mod. Phys. A **27** (2012) 1250016, [arXiv:1108.6129](#).

- [234] W. Detmold and S. Meinel,  $\Lambda_b \rightarrow \Lambda \ell^+ \ell^-$  form factors, differential branching fraction, and angular observables from lattice QCD with relativistic  $b$  quarks, Phys. Rev. D **93** (2016) 074501, [arXiv:1602.01399](#).
- [235] F. James, in *Statistical methods in experimental physics*, ch. 8.2, World Scientific, 2006.
- [236] F. Beaujean, M. Chruszcz, N. Serra, and D. van Dyk, *Extracting Angular Observables without a Likelihood and Applications to Rare Decays*, Phys. Rev. D **91** (2015) 114012, [arXiv:1503.04100](#).
- [237] LHCb collaboration, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [238] LHCb collaboration, M. P. Whitehead, *A palette of fast simulations in LHCb*, PoS **ICHEP2018** (2019) 271.
- [239] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in  $B_s^0 \rightarrow J/\psi K^+ K^-$  decays*, Phys. Rev. Lett. **114** (2015) 041801, [arXiv:1411.3104](#).
- [240] LHCb collaboration, R. Aaij *et al.*, *Updated measurement of time-dependent CP-violating observables in  $B_s^0 \rightarrow J/\psi K^+ K^-$  decays*, Eur. Phys. J. **C79** (2019) 706, [arXiv:1906.08356](#), [Erratum: Eur. Phys. J.C80,no.7,601(2020)].
- [241] I. Dunietz, R. Fleischer, and U. Nierste, *In pursuit of new physics with  $B_s$  decays*, Phys. Rev. D **63** (2001) 114015, [arXiv:hep-ph/0012219](#).
- [242] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. A **764** (2014) 150, [arXiv:1312.5000](#).
- [243] A. Bharucha, D. M. Straub, and R. Zwicky,  *$B \rightarrow V \ell^+ \ell^-$  in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, [arXiv:1503.05534](#).
- [244] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  observables using form factors from lattice QCD*, Phys. Rev. Lett. **112** (2014) 212003, [arXiv:1310.3887](#).
- [245] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Rare B decays using lattice QCD form factors*, PoS **LATTICE2014** (2015) 372, [arXiv:1501.00367](#).
- [246] LHCb collaboration, S. Sheldon, R. Lindner *et al.*, *Letter of Intent for the LHCb Upgrade*, CERN, 2011.
- [247] LHCb collaboration, I. Bediaga *et al.*, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN, 2012.
- [248] LHCb collaboration, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, [arXiv:1808.08865](#).
- [249] D0 collaboration, V. M. Abazov *et al.*, *Search for the rare decay  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  with the D0 detector*, Phys. Rev. D **74** (2006) 031107, [arXiv:hep-ex/0604015](#).
- [250] CDF collaboration, D. Acosta *et al.*, *Search for the Decay  $B_s \rightarrow \mu^+ \mu^- \phi$  in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8\text{-TeV}$* , Phys. Rev. D **65** (2002) 111101.

# Acknowledgements

Finally, I want to thank everybody who made this thesis possible. First of all, I would like to thank Dr. Christoph Langenbruch for his continued support and for giving me the possibility to write this dissertation and to travel to and work at CERN for six months. I also want to thank Prof. Dr. Michael Krämer who agreed to serve as a second referee for my thesis.

Thanks also to all members of the LHCb Aachen group, current and former, for the great work environment and atmosphere. Thanks to Dr. Eluned Smith for the supervision of my work, answering every question and being available whenever help was needed. Without your help I would never have been able to get the analysis to this point. Thanks to all member of the office, Sophie Kretzschmar, Stephan Escher, Dr. Eluned Smith and Leon Carus, for being open to discussions about physics and any other topic. I already miss working in the office with you. Thanks to everybody who contributed to this analysis, thanks to Johannes Heuel, Sebastian Schmitt, Hendrik Jage, Thomas Oeser, Jan-Marc Basels, Simon Nieswand, Mark Whitehead, Gedas Sarpis, Andreas Güth and many more. Thanks to everybody in the LHCb collaboration especially my review committee, Carla and Conor, and everyone in the RD and EWP working groups, your input to this analysis was very helpful and much appreciated.

Thanks also to my friends who took my mind also off of physics, thanks to Stephan Matzke, Felix Eicke, Tabea Buss, Christian Pohen, Sebastian Wiedenbeck, Leonie Moerschbacher, Fabian Machate and many more. Thanks to my family, for all the love and support. Thanks to my wife, Federica Dematté, for always having my back and supporting me. Finally, I would like to thank my parents, without whom this thesis would not have been possible. Thank you.