Dissertation

submitted to the

Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

M.Sc. Simon Stemmle

born in: Heidelberg, Germany

Oral examination: 22.07.2019

CP violation and lifetime measurements in the decay $B_s^0 \to J/\psi\,\phi$ with the LHCb experiment

Referees:

Prof. Dr. Stephanie Hansmann-Menzemer

Prof. Dr. Klaus Reygers

Abstract

This thesis presents a study of the decay $B_s^0 \to J/\psi\,\phi$ with the LHCb experiment. Due to the mixing between B_s^0 and $\overline{B}_s^0$ mesons and a final state that is accessible for both species, this decay is sensitive to CP violation originating from the interference between the mixing and the decay process. The CP violating is parametrized by the weak phase difference ϕ_s, which can be precisely constrained within the Standard Model of particle physics based on other measurements. Thus, the measurement of this phase difference constitutes an interesting test for possible contributions of physics phenomena from beyond the Standard Model. A flavour-tagged analysis of the time-dependent decay rates of B_s^0 and $\overline{B}_s^0$ mesons is performed. In addition, an angular analysis is needed to disentangle the different CP components of the final state. The analysis is based on a proton-proton collision data set collected in 2015 and 2016 by the LHCb experiment, corresponding to an integrated luminosity of $1.9\,\mathrm{fb}^{-1}$. The obtained value for the CP violating phase difference is $\phi_s = (0.083 \pm 0.041_{\mathrm{stat}} \pm 0.006_{\mathrm{syst}})\,\mathrm{rad}$, which, combined with a previous analysis of this channel by the LHCb experiment, constitutes the most precise single channel and single experiment measurement of this quantity. No significant deviation from the Standard Model expectation is observed.

Besides the measurement of ϕ_s, the main other determined parameters are the decay width and decay-width splitting of the B_s^0 meson system. In contrast to previous analyses of this channel, the decay width is directly measured with respect to decay width of the B^0 meson.

Kurzfassung

In dieser Arbeit wird eine Studie des Zerfalls $B_s^0 \to J/\psi\,\phi$ mit dem LHCb-Experiment vorgestellt. Aufgrund der Mischung von B_s^0- und $\overline{B}_s^0$-Mesonen und einem Endzustand der diesen beiden zugänglich ist, ist dieser Zerfall sensitiv auf CP-Verletzung die aus der Interferenz zwischen Mischung und Zerfall entsteht. Diese CP-Verletzung wird durch die schwache Phasendifferenz ϕ_s parametrisiert, welche, basierend auf anderen Messungnen, präzise innerhalb des Standardmodells der Teilchenphysik vorhergesagt werden kann. Deshalb stellt die Messung dieser Phasendifferenz einen interessanten Test auf mögliche Beiträge von außerhalb des Standardmodells dar. Eine flavour-abhängige Analyse der zeitabhängigen Zerfallsraten von B_s^0- und $\overline{B}_s^0$-Mesonen wird durchgeführt. Zusätzlich ist eine Winkelanalyse notwendig um die unterschiedlichen CP-Eigenzustände der Zerfallsprodukte voneinander zu trennen. Die Analyse basiert auf einem Proton-Proton Kollisions-Datensatz, der in den Jahren 2015 und 2016 durch das LHCb Experiment aufgezeichnet wurde und der einer integrierten Luminosität von $1.9\,\mathrm{fb}^{-1}$ entspricht. Der gemessene Wert der CP-verletzenden Phasendifferenz liegt bei $\phi_s = (0.083 \pm 0.041_{\mathrm{stat}} \pm 0.006_{\mathrm{syst}})\,\mathrm{rad}$ und bildet, kombiniert mit einer früheren Messung in diesem Kanal, die präziseste Messung in einem einzelnen Kanal und durch ein einzelnes Experiment. Es ist keine signifikante Abweichung von dem im Standardmodell erwarteten Wert zu beobachten.

Zusätzlich zur Messung von ϕ_s sind die wichtigsten anderen gemessen Parameter die Zerfallsbreite und der Zerfallsbreitenunterschied im B_s^0-Meson-System. Im Gegenteil zu frühere Analysen dieses Kanals, wird die Zerfallsbreite hier relative zur Zerfallsbreite des B^0-Mesons gemessen.

Contents

Introduction

Since millennia, mankind is wondering about the most fundamental building blocks of nature and about the cement that keeps them together. Today, the Standard Model of particle physics (SM) summarizes our current understanding of this underlying structure in nature. It describes all known fundamental interactions except gravity, and within the last decades it has been probed and confirmed to amazing precision. Also the Higgs boson, whose discovery in 2012 [1, 2] completed the particle content of the SM, more and more proves to behave like predicted by the SM [3].

However, this tremendous success of the SM puts the scientific community in an unpleasant situation. One the one hand the SM precisely describes elementary processes, on the other hand it lacks explanations for some very fundamental observations. An example is the energy content of the universe. Cosmological and astronomical observations strongly suggest the presence of dark matter, an at most weakly interacting form of matter, with an abundance more than five times larger than the baryonic matter we know [4]. Within the SM, none of the fundamental particles is suited to serve as candidate for this dark matter. Other examples for the shortcomings of the SM are the large matter antimatter asymmetry observed in the universe and the origin of the extremely small neutrino masses.

All this leads to the idea of an even more fundamental underlying theory. Thus, the primary interest in these days lies no longer in the precise determination of the SM parameters alone, but in finding inconsistencies and processes where the theory breaks down. The Large Hadron Collider (LHC) at CERN is currently the most powerful particle accelerator and is therefore predestinated for these kind of searches. While the two largest experiments at the LHC, ATLAS and CMS, primarily aim to detect potential new particles that are directly produced in the proton proton collisions, the LHCb experiment follows an alternative, indirect strategy. Possible extensions of the SM, manifested in new heavy particles or interactions, are likely to enter in quantum loops of SM processes and therefore modify these. Since particles contributing to these quantum loops are not limited by the available energy in the proton-proton collision, one can probe energy scales that are typically not reached in direct searches.

Besides precise measurements of such SM processes, the indirect strategy relies on the availability of predictions or constraints within the SM at a similar or better level. One of the most promising channels at the LHCb experiment is the decay[1] of a B_s^0 meson to a J/ψ and a ϕ meson. Since the final state is accessible for the B_s^0 as well as for the $\overline{B}_s^0$ meson, a CP-violation measurement in this decay is sensitive to the B_s^0-$\overline{B}_s^0$ mixing process, which is primarily governed by the above mentioned quantum loops. In addition, the size of the CP violation can be preciously constrained withing the SM from other measurements.

Currently[2], the LHCb experiment provides the most precise single measurement of CP violation in this channel [5]. This thesis presents an update of this measurement with data recorded in 2015 and 2016. It is based on a decay-time- and angular-dependent analysis of the decay rates of B_s^0 and $\overline{B}_s^0$ mesons. Besides the CP-violation and mixing parameters, also the decay-width difference between the B_s^0 and B^0 mesons is precisely determined.

The thesis is structured as follows: Chapter 1 provides a theoretical introduction to the topic. Afterwards, the LHCb experiment is briefly described in Chapter 2, and the analysis strategy is outlined in Chapter 3. After a short introduction of relevant statistical tools and techniques in Chapter 4, Chapters 5 to 7 provide detailed descriptions of the selection of $B_s^0 \to J/\psi\,\phi$ decays, the determination of the B_s^0 flavour at production and the modeling and extraction of detector acceptances and resolutions, respectively. The heart of the analysis, a maximum likelihood fit to the decay time and angular observables, is presented in Chapter 8. It is used to determine the CP-violation, mixing and decay-width parameters, which are then presented in Chapter 9. After a discussion of the relevant systematic uncertainties in Chapter 10, the thesis will conclude with a summary of the results and their combination with other measurements.

Besides the analysis presented in the following, the author of this thesis was the main contributor to two other projects within the LHCb collaboration. During his master studies and in the first year of his PhD, he worked on the measurement of CP violation in $D^0 \to K^+K^-$ decays, which was published in Ref. [6]. Furthermore, he developed a fast parametrized Kalman filter for the LHCb upgrade tracking system, which will be used in the future trigger. An overview of this study is presented in Ref. [7].

[1] If not stated otherwise, the charge-conjugated decays are implied.
[2] Currently means at the time the analysis, presented here, was first shown to the public.

1 Theoretical background

In this chapter the basic theoretical concepts and formulas that are needed for the analysis of the decay $B_s^0 \to J/\psi\,\phi$ are introduced. It starts with a short overview of the Standard Model of particle physics (SM), which is currently the best available description for fundamental interactions. After that, the concept of CP violation is introduced as a sensitive way to probe the SM. In this context, the chapter continues with a detailed description of the Cabibbo-Kobayashi-Maskawa mechanism, which is the only source of CP violation in the SM. After a general introduction to meson mixing and CP violation in neutral meson systems, these concepts are applied to the decay $B_s^0 \to J/\psi\,\phi$. The chapter concludes with a detailed discussion of the differential decay rate of this decay, which is later used to model the observed data and to extract the underlying physics parameters.

1.1 Standard Model of particle physics

When in 2012 the existence of the Higgs boson was confirmed by the LHC experiments ATLAS and CMS [1, 2], the last elementary particle predicted by the Standard Model of particle physics was discovered. Therefore, we are now in the situation to have a model that can describe every known fundamental elementary particle and whose principle predictions are fully confirmed with high precision by experiment. In the following, a short overview of this model is given[1].

Fundamental particles

The visible matter in our universe is completely made up of elementary particles with spin $\frac{1}{2}$, called fermions. The SM describes these as fields and categorizes them into leptons and quarks. The quarks interact via the strong interaction, while the leptons do not. In the quark sector there are so-called up-type and down-type quarks, which have an electric charge of $\frac{2}{3}\,e$ and $-\frac{1}{3}\,e$, respectively. The fermions can be grouped into leptons with electric charge $-1\,\mathrm{e}$ and neutrinos, which are

[1] For a more detailed introduction to the Standard Model see for example Ref. [8].

Table 1.1: Fermionic content of the Standard Model of particle physics. For each of these particles also the corresponding anti-particle with the opposite charge is implied. The masses are taken from Ref. [3].

	Leptons			Quarks		
Generation	Type	El. charge $[e]$	Mass	Type	El. charge $[e]$	Mass
1^{st}	u	$+\frac{2}{3}$	$2.2\,\text{MeV/c}^2$	e^-	-1	$0.511\,\text{MeV/c}^2$
	d	$-\frac{1}{3}$	$4.7\,\text{MeV/c}^2$	ν_e	0	$< 2\,\text{eV/c}^2$
2^{nd}	c	$+\frac{2}{3}$	$1.275\,\text{GeV/c}^2$	μ^-	-1	$106\,\text{MeV/c}^2$
	s	$-\frac{1}{3}$	$95\,\text{MeV/c}^2$	ν_μ	0	$< 0.19\,\text{eV/c}^2$
3^{rd}	t	$+\frac{2}{3}$	$173\,\text{GeV/c}^2$	τ^-	-1	$1.777\,\text{GeV/c}^2$
	b	$-\frac{1}{3}$	$4.18\,\text{GeV/c}^2$	ν_τ	0	$< 18.2\,\text{eV/c}^2$

neutral. Both, quarks and leptons, always come in pairs of up-type and down-type or charged and neutral, respectively. There are three generations of these pairs in the quark sector as well as in the lepton sector, which are ordered by increasing mass. Table 1.1 provides a list of all SM fermions.

Fundamental interactions

The SM is a quantum field theory that describes the previously introduced elementary particles and their interactions. These interactions are generated by requiring the Lagrangian to be invariant under local gauge transformations of the $\text{SU(3)} \bigotimes \text{SU(2)} \bigotimes \text{U(1)}$ symmetry group. According to Noether's theorem [9], each of these symmetries leads to a charge that is conserved in the respective interaction. In addition, the generators of the gauge groups correspond to so-called gauge bosons that are spin 1 particles and mediate the three fundamental interactions, which are the electromagnetic, the weak and the strong force.

Quantum chromodynamics (QCD) describes the strong force and is generated by the SU(3) group. The eight massless gauge bosons are called gluons and carry the three strong charges, called colors. Besides the gluons themselves, only quarks carry such a color charge and are therefore the only other fundamental particles that participate in the strong interaction. Two important features of QCD are the confinement and the dependence of the coupling constant on the energy scale. The former describes a property of the QCD potential. In contrast to other forces, the strong force between two colored objects does not decrease for large distances but stays constant. The result is that colored particles always form color neutral objects, called hadrons, and never appear barely. Most hadrons can be categorized as either mesons, combinations of a quark and an anti-quark, or baryons, bound systems of three quarks or three anti-quarks.

The coupling constant $\alpha_s(Q^2)$ strongly depends on the transferred four-momentum squared, Q^2. It is increasing with decreasing energy, which prohibits calculations within perturbation theory in this regime. Only at energies significantly above a certain reference scale, $\Lambda_{QCD} \approx 200\,\text{MeV}$, perturbative calculations become reliable.

The electroweak interaction is the unification of the electromagnetic and weak interaction and is generated by the SU(2)⊗U(1) symmetry. The corresponding gauge bosons are indicated with $W_\mu^{1,2,3}$ and B_μ, and the charges are the weak isospin and hypercharge. The left-handed fermions carry weak isospin 1/2 while the right-handed fermions do not couple to the SU(2) part of the electroweak force and have a weak isospin of zero. This means that the W_μ^i bosons couple only to left-handed particles and right-handed anti-particles. However, the B_μ boson, which couples to the weak hypercharge, interacts with all fermions of the SM.

The spontaneous symmetry breaking within the Higgs mechanism gives mass to the fermions and converts the previously introduced gauge bosons of the electroweak interactions to the three massive bosons W^+, W^- and Z of the weak interaction, and the massless photon, which couples to the electric charge. As a superposition of two W_μ^i bosons, the charged weak gauge bosons couple only to left-handed particles and right-handed anti-particles, while the neutral Z retrieves a contribution from the U(1) gauge boson and couples also to particles with the opposite chirality. A further implication of the Higgs mechanism is the prediction of the massive Higgs boson, which couples to all massive fundamental particles and completes the bosonic content of the SM.

Beyond the SM

Although the SM is impressively successful in explaining qualitative and quantitative aspects of fundamental particle physics, there are good reasons to consider it not to be the final theory in this field. An obvious shortcoming is that none of the particles described by the SM can explain the amount and nature of the dark matter that is observed in the universe [4]. In addition, there are theoretical arguments that suggest the existence of a more fundamental theory. An example is the large spread of fermion masses, which are parameters of the SM, across many orders of magnitude. The aim is to develop a theory that contains a mechanism to explain such extreme numbers and, ideally, decreases the number of input parameters.

Therefore, experimental particle physics is continuously searching for fundamental particles or interactions that are not part of the SM (so-called New Physics). The first approach is the direct search for such particles produced in high energetic proton-proton collisions. The maximum mass of new particles that can be discovered

in this way is limited by the energy available in a single proton-proton collision. Another approach that is not directly limited by the energy the accelerators can provide, is the indirect search for such new particles. These particles might be too heavy to be produced as real particles, but they might well enter in quantum loops where they stay virtual. The indirect search is therefore based on three ingredients:

- a process or observable that obtains sizable contributions from such a quantum loop,
- a way to precisely measure this observable in experiment,
- and a precise prediction of it within the SM to which the measurement can be compared to.

One interesting concept that fulfills these requirements in many cases is CP violation. In the following sections, the general concept of CP violation and the specific decay channel chosen in this thesis will be introduced.

1.2 CP violation

Symmetries play an important role in the fundamental description of nature. As an example, for a long time the parity (P) transformation was considered to be such a symmetry. This means that physical systems should be invariant under the inversion of spacial coordinates $\overrightarrow{x}$:

$$\mathbf{P}\overrightarrow{x} = -\overrightarrow{x}, \tag{1.1}$$

where $\mathbf{P}$ is the parity operator. However, in 1956 Lee and Yang postulated [10] and Wu discovered [11] that the weak interaction maximally violates parity by coupling differently to left and right-handed particles. Charge conjugation (C) flips all internal charge like quantum numbers of a system and relates thereby particles and their antiparticles. It is also maximally violated by the weak interaction. This lead to the hypothesis of CP, as the combined transformation of parity and charge, being the actual fundamental matter-antimatter symmetry [12]. This idea was falsified by the observation of CP violation in the neutral kaon system in 1964 [13].

In the following, the basic requirements for a process in particle physics to show CP violation are discussed. A generic process is characterized by its initial state I and final state F. Charge and parity conjugation of an individual amplitude A contributing to this process does not change the absolute value but at most the

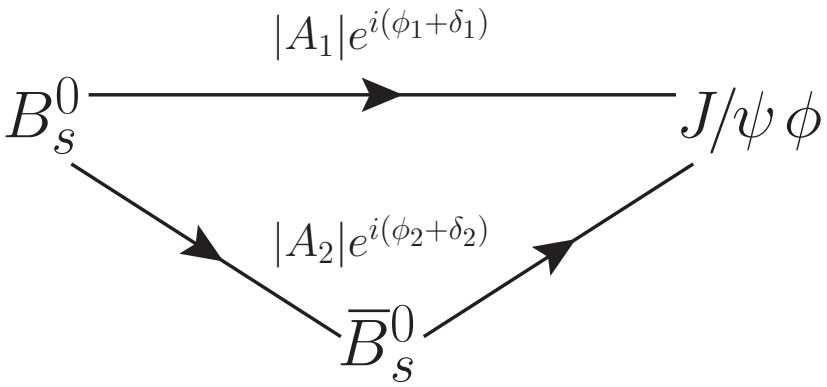


Figure 1.1: Schematic representation of the two main amplitudes responsible for CP violation in the decay $B_s^0 \to J/\psi\,\phi$. The amplitudes follow the definition given in the text.

phase of this amplitude. Typically, the phase of A can therefore be split into a strong phase δ, which does not change sign, and a weak phase ϕ, which changes sign under CP transformation:

$$\mathbf{CP}A = \mathbf{CP}|A|e^{i(\phi+\delta)} = |A|e^{i(-\phi+\delta)}. \tag{1.2}$$

In case the process I→F is governed only by a single amplitude, the transition probability, which is proportional to AA^*, does not change. However, if two (or more) amplitudes, A_1 and A_2, are contributing to the total amplitude A_{tot}, interference can occur and the situation is different. Given the two amplitudes and their respective strong and weak phases,

$$A_1 = |A_1|e^{i(\phi_1+\delta_1)}, \tag{1.3}$$
$$A_2 = |A_2|e^{i(\phi_2+\delta_2)}, \tag{1.4}$$

the difference between the original and the CP-conjugated transition probability is proportional to:

$$\mathbf{CP}[A_{tot}A_{tot}^*] - A_{tot}A_{tot}^* = \mathbf{CP}[(A_1+A_2)(A_1+A_2)^*] - (A_1+A_2)(A_1+A_2)^* \tag{1.5}$$
$$= 4|A_1||A_2|\sin(\phi_1-\phi_2)\sin(\delta_1-\delta_2). \tag{1.6}$$

This means that there can be only CP violation if the strong phases as well as the weak phases differ between the two amplitudes.

Having a look at the decay studied in this thesis, one can identify the two main contributing amplitudes as the decay of the B_s^0 meson to the final state $J/\psi\phi$ and the mixing of a B_s^0 meson into a $\overline{B}_s^0$ meson with the subsequent decay $\overline{B}_s^0 \to J/\psi\,\phi$, see Fig. 1.1.

Typically, at least one of the amplitudes contains an internal quantum loop where non-SM-like particles can enter and alter the strong or weak phase difference. Together with the fact that the SM values of many CP observables can be precisely determined from other measurements, this makes the study of CP violation an excellent probe for non-SM-like interactions or particles.

1.3 Flavour changing currents of the weak interaction

Within the SM, CP violation only occurs when flavour and charge changing currents of the weak interaction are involved. The gauge bosons W^+ and W^- mediate these currents. Their coupling to the fermions is described by the following part of the SM Lagrangian:

$$\mathcal{L} = -\frac{g}{2\sqrt{2}}\left[\sum_i W_\mu^+ \overline{\phi_i^u}\gamma^\mu(1-\gamma^5)\phi_i^d + \sum_i W_\mu^- \overline{\phi_i^d}\gamma^\mu(1-\gamma^5)\phi_i^u\right]. \tag{1.7}$$

Here, g is the coupling constant of the $\mathrm{SU}(2)_L$ gauge group and ϕ^u/ϕ^d are the fields of up and down type quarks or leptons:

$$\phi^u = \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix}, \phi^d = \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \quad or \quad \phi^u = \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}, \phi^d = \begin{pmatrix} e^- \\ \mu^- \\ \tau^- \end{pmatrix}. \tag{1.8}$$

Dirac matrices are denoted by γ^μ and γ^5, and the operator $1/2(1-\gamma^5)$ projects out the left-handed part of the fermion fields.

In Equation (1.8) the quarks are given as electroweak eigenstates (u_i', d_i'), which in general do not coincide with the mass eigenstates. After symmetry breaking, the quarks and charged leptons obtain their mass via the Higgs mechanism. Neutrino masses are ignored at this point, since they are not relevant for this study. The quark mass eigenstates (u_i, d_i) are then given by a rotation of the electroweak eigenstates. When writing the weak interaction Lagrangian in terms of the mass

eigenstates, these rotations of the up and down type quarks can be absorbed to a single unitary matrix V:

$$\mathcal{L}_{quark} = -\frac{g}{2\sqrt{2}} \left[\sum_{i,j} W_\mu^+ \overline{u_i} \gamma^\mu (1-\gamma^5) V_{ij} d_j + \sum_{i,j} W_\mu^- \overline{d_j} \gamma^\mu (1-\gamma^5) V_{ij}^\dagger u_i \right]. \quad (1.9)$$

This means that up and down type quarks from different families i and j couple to each other with a relative strength proportional to V_{ij}.

The matrix V is called Cabibbo-Kobayashi-Maskawa matrix (V_{CKM}) [14] and is strongly related to CP violation in the SM. This can be understood by analyzing the behavior of the Lagrangian under CP transformation. Objects like $\overline{\phi}\gamma^\mu\phi$ change sign and therefore transform as vector under parity transformation. Adding an additional γ^5 changes the transformation behavior, which is why $\overline{\phi}\gamma^\mu\gamma^5\phi$ transforms as axial vector and flips no sign. This reflects the V-A structure of the weak interaction that maximally violates parity. Also under C transformation, which flips all charge-like quantum numbers, the vector and axial vector parts obtain a relative minus sign. Therefore, C is also maximally violated. Using these transformations of Equation (1.9) one can derive that the combination of C and P can only be violated if $V_{CKM} \neq V_{CKM}^*$.

The CKM matrix describes the couplings of the charged weak bosons to the different combinations of an up-type and a down-type quark. As an unitary three by three matrix it has nine degrees of freedom, which are further reduced to four by the free choice of the unobservable quark phases. These four degrees of freedom can be split into three rotation angles and one phase. The latter is the origin of the weak phase discussed in Section 1.2.

One of the most common parametrizations of the CKM matrix was developed by Wolfenstein [8]:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (1.10)$$

where $\lambda \approx 0.23$, is the sine of the rotation angle between the first two generations and is used as an expansion parameter. A, ρ and η are of $\mathcal{O}(0.1-1)$. This parametrization nicely visualizes the nearly diagonal structure of the matrix, which reflects suppressed

transitions between different quark families. In this parametrization and up to $\mathcal{O}(\lambda^3)$, the only elements with a non-zero phase are V_{ub} and V_{td}.

One of the relations between the CKM elements following from the unitarity of the CKM matrix is:

$$V_{ud}V_{ub} + V_{cd}V_{cb} + V_{td}V_{tb} = 0. \tag{1.11}$$

This equation can be presented by a so-called unitarity triangle in the complex plane with the angles

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.12}$$

where β and γ are, using the Wolfenstein parametrization, to first order the phases of V_{td} and V_{ub}.

As discussed later, these two CKM elements do not occur in the dominant amplitudes of B_s^0 mixing or the decay $B_s^0 \to J/\psi\,\phi$. However, the element V_{ts} enters in the mixing and at $\mathcal{O}(\lambda^4)$ acquires a complex part. Via another unitarity relation, namely

$$V_{us}V_{ub} + V_{cs}V_{cb} + V_{ts}V_{tb} = 0, \tag{1.13}$$

one can define another angle that is of particular interest in the scope of this thesis:

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{1.14}$$

Due to the relatively small complex part of V_{ts}, this angle and any CP violating effect in the B_s^0 system is expected to be small. To probe β_s is one of the central aspects of this thesis, and its relation to the decay $B_s^0 \to J/\psi\,\phi$ will be discussed in more detail in the next sections.

1.4 Neutral meson phenomenology

In the following section, the phenomenology of mixing and CP violation in neutral meson systems is summarized[2]. Although the focus is on the B_s^0 system, the presented formalism is also valid for all other neutral mesons. A discussion of the specific features of the decay $B_s^0 \to J/\psi\,\phi$ will follow in the next section.

[2]This section is based on Ref. [3, 15].

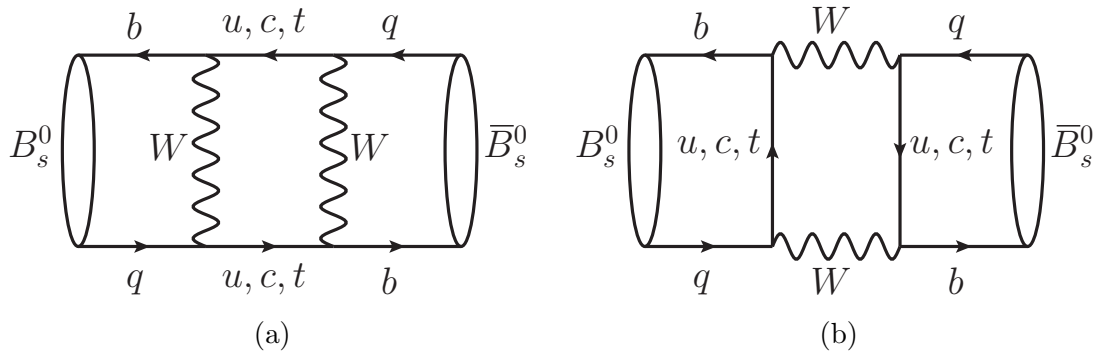


Figure 1.2: Leading order diagrams for the transition between B_s^0 and $\overline{B}_s^0$ mesons.

1.4.1 Mixing

The two neutral mesons B_s^0 and $\overline{B}_s^0$ are flavour eigenstates with the quark content $\left|\bar{b}s\right\rangle$ and $|b\bar{s}\rangle$. Due to the presence of flavour changing currents in the weak interaction, transitions between these two states are possible. Examples of such amplitudes are the so-called *box-diagrams* that are shown in Figure 1.2. These transitions imply that the flavour eigenstates do not coincide with the mass eigenstates of the system. A general state in this system is given by a superposition of the two flavour eigenstates, $|\Psi(t)\rangle = a(t)|B_s^0\rangle + b(t)|\overline{B}_s^0\rangle$. Thus, the mass eigenstates are given by the eigenstates of the effective Hamiltonian describing the two-state B_s^0-$\overline{B}_s^0$ system. The according Schrödinger equation can be written as:

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \left(\boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma}\right)\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{1.15}$$

where the effective Hamiltonian is split into a hermitian matrix $\boldsymbol{M}$ and an anti-hermitian matrix $\frac{i}{2}\boldsymbol{\Gamma}$. The off diagonal elements $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$ represent the *short range* and *long range* contributions to the transitions between the flavour eigenstates. The former originate from the *box diagrams*, which involve heavy virtual particles, while the latter accounts for transitions via on-shell light states, like for a example a $\pi^+\pi^-$ pair. The CPT invariance theorem constrains the diagonal elements:

$$M_{11} = M_{22} \text{ and } \Gamma_{11} = \Gamma_{22}. \tag{1.16}$$

The mass eigenstates are labeled as $B^0_{s,L}$ (light) and $B^0_{s,H}$ (heavy) according to their masses M_L and M_H, have the decay widths Γ_L and Γ_H and are given by:

$$|B^0_{s,H}\rangle = p|B^0_s\rangle + q|\overline{B}^0_s\rangle, \tag{1.17}$$

$$|B^0_{s,L}\rangle = p|B^0_s\rangle - q|\overline{B}^0_s\rangle, \tag{1.18}$$

with

$$\frac{q}{p} = -\sqrt{\frac{M^*_{12} - \frac{i}{2}\Gamma^*_{12}}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{1.19}$$

These states evolve in time as:

$$|B^0_{s,H/L}(t)\rangle = e^{-iM_{H/L}t - \Gamma_{H/L}t/2}|B^0_{s,H/L}\rangle. \tag{1.20}$$

One defines the following important quantities describing the mixing:

$$\Delta m = M_H - M_L, \qquad m = \frac{M_H + M_L}{2} = M_{11}, \tag{1.21}$$

$$\Delta\Gamma = \Gamma_L - \Gamma_H, \qquad \Gamma = \frac{\Gamma_H + \Gamma_L}{2} = \Gamma_{11}. \tag{1.22}$$

The mass and lifetime differences can be related to the effective Hamiltonian via:

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4\left|M_{12}\right|^2 - \left|\Gamma_{12}\right|^2, \quad \Delta m\,\Delta\Gamma = -4\Re(M_{12}\Gamma^*_{12}). \tag{1.23}$$

Using these variables, and the Equations (1.17), (1.18) and (1.20), one can derive the time evolution of the two initial flavour eigenstates:

$$|B^0_s(t)\rangle = g_+(t)|B^0_s\rangle + \frac{q}{p}g_-(t)|\overline{B}^0_s\rangle, \tag{1.24}$$

$$|\overline{B}^0_s(t)\rangle = \frac{p}{q}g_-(t)|B^0_s\rangle + g_+(t)|\overline{B}^0_s\rangle. \tag{1.25}$$

Here, $g_+(t)$ and $g_-(t)$ are given by:

$$g_+(t) = e^{-imt}e^{-\Gamma t/2}\left(\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} - i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right), \tag{1.26}$$

$$g_-(t) = e^{-imt}e^{-\Gamma t/2}\left(-\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right). \tag{1.27}$$

Given that a B_s^0 ($\overline{B}_s^0$) meson was produced at time $t = 0$, the probability that it has mixed to a $\overline{B}_s^0$ (B_s^0) meson after time t can be expressed by:

$$\left|\langle \overline{B}_s^0 | B_s^0(t) | \rangle\right|^2 = \left|\frac{q}{p}\right|^2 |g_-(t)|^2 , \tag{1.28}$$

$$\left|\langle B_s^0 | \overline{B}_s^0(t) | \rangle\right|^2 = \left|\frac{p}{q}\right|^2 |g_-(t)|^2 , \tag{1.29}$$

$$\tag{1.30}$$

with

$$|g_-(t)|^2 = \frac{e^{-\Gamma t}}{2}\left(\cosh\frac{\Delta\Gamma t}{2} - \cos\Delta m t\right). \tag{1.31}$$

The mass splitting Δm plays the role of the frequency of an oscillation modifying an exponential decay. In the B^0 system the current world average for this frequency is $\Delta m_d = (0.5065 \pm 0.0019)\,\mathrm{ps}^{-1}$ [16], which is significantly smaller than in the B_s^0 system, $\Delta m_s = (17.757 \pm 0.021)\,\mathrm{ps}^{-1}$ [16]. The latter was measured most precisely by the LHCb collaboration with $B_s^0 \to D_s^- \pi^+$ decays [17]. In addition, the pure exponential decay is modified by the decay-width splitting $\Delta\Gamma$, which is negligible for the B^0 system but sizable for B_s^0 mesons: $\Delta\Gamma_s = (0.085 \pm 0.006)\,\mathrm{ps}^{-1}$ [16].

The difference between Equations (1.28) and (1.29), namely the absolute value of the ratio q/p, indicates that the mixing probabilities do not have to be the same for an initial B_s^0 and $\overline{B}_s^0$ meson. This is one of the types of CP violation that can occur in neutral meson systems. All of them will be discussed in the next section.

1.4.2 CP violation

The two neutral B_s^0 mesons transform into each other under CP as:

$$\mathrm{CP}|B_s^0\rangle = -|\overline{B}_s^0\rangle \text{ and } \mathrm{CP}|\overline{B}_s^0\rangle = -|B_s^0\rangle. \tag{1.32}$$

Therefore, any observable difference in the decay rates of the two mesons is equivalent to CP violation in this system. Of special interest is the case when the B_s^0 mesons decay to a CP eigenstate that is accessible for the B_s^0 as well as for the $\overline{B}_s^0$ state. An example for such a decay is the mode $B_s^0 \to J/\psi\,\phi$ which is studied in this thesis.

In the following, the decay amplitudes of a B_s^0 and $\overline{B}_s^0$ mesons to such a final state f will be denoted as:

$$A_f = \langle f|B_s^0\rangle \text{ and } \overline{A}_f = \langle f|\overline{B}_s^0\rangle. \tag{1.33}$$

These amplitudes allow to define a further quantity that can be used to parametrize CP violation:

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{1.34}$$

The time dependence of the decay rates of B_s^0 and $\overline{B}_s^0$ mesons are given by

$$\frac{d\Gamma_{B_s^0\to f}(t)}{dt} \propto \left|\langle f|B_s^0(t)\rangle\right|^2, \tag{1.35}$$

$$\frac{d\Gamma_{\overline{B}_s^0\to f}(t)}{dt} \propto \left|\langle f|\overline{B}_s^0(t)\rangle\right|^2. \tag{1.36}$$

Using Equations (1.24) to (1.27) and the definition of λ_f this gives:

$$\frac{d\Gamma_{B_s^0\to f}(t)}{dt} \propto |A_f|^2 \quad \frac{1}{1+C_f}e^{-\Gamma_s t}\Big[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + D_f\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + C_f\cos\left(\Delta m_s t\right) - S_f\sin\left(\Delta m_s t\right)\Big] \tag{1.37}$$

$$\frac{d\Gamma_{\overline{B}_s^0\to f}(t)}{dt} \propto |A_f|^2\left|\frac{p}{q}\right|^2 \frac{1}{1+C_f}e^{-\Gamma_s t}\Big[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + D_f\sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - C_f\cos\left(\Delta m_s t\right) + S_f\sin\left(\Delta m_s t\right)\Big], \tag{1.38}$$

with

$$C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2},\ S_f = \frac{2\Im(\lambda_f)}{1+|\lambda_f|^2} \text{ and } D_f = -\frac{2\Re(\lambda_f)}{1+|\lambda_f|^2}. \tag{1.39}$$

CP violation is equivalent with Equations (1.37) and (1.38) being not identical. This is the case when either $|q/p|$ or λ_f deviate from unity. There are three types of CP violation that alter these quantities in different ways.

CP violation in the decay

The probably most intuitive type of CP violation occurs when the absolute decay amplitude $|A_f|$ of a B_s^0 to a final state f is different from $|\overline{A}_{\overline{f}}|$, the absolute decay amplitude of a $\overline{B}_s^0$ to the final state $\overline{f}$. This results in a difference in the partial

decay width of the CP conjugated decays $B_s^0 \to f$ and $\overline{B}_s^0 \to \overline{f}$. Given that the CP eigenstate f has the CP eigenvalue η_{CP},

$$\langle \overline{f}| = CP\langle f| = \eta_{CP}\langle f|, \tag{1.40}$$

λ_f can be written in terms of the two amplitudes A_f and $\overline{A}_{\overline{f}}$:

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f} = \eta_{CP}\frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_f}. \tag{1.41}$$

Assuming that $|q/p| = 1$, CP violation in the decay is equivalent with a deviation from unity of the absolute value of λ_f.

CP violation in mixing

Another type of CP violation occurs when the probability of the transition $B_s^0 \to \overline{B}_s^0$ is different from the one of $\overline{B}_s^0 \to B_s^0$. This causes the decay rates $\Gamma(B_s^0(\to \overline{B}_s^0) \to \overline{f})$ and $\Gamma(\overline{B}_s^0(\to B_s^0) \to f)$ to flavour specific final states f and $\overline{f}$ to be different. According to Equations (1.28) and (1.29) this is equivalent to:

$$\left|\frac{q}{p}\right| \neq 1. \tag{1.42}$$

Assuming no CP violation in decay this leads as well to a modification of the absolute value of λ_f.

CP violation in the interference between mixing and decay

In case the final state f is a CP eigenstate there can be CP violation even if the two previously discussed sources are not present. While they modify the absolute value of λ_f, CP is also violated if λ_f has a non-zero imaginary part:

$$\lambda_f = |\lambda_f|\, e^{i\phi} \text{ with } \phi \neq 0. \tag{1.43}$$

According to Equations (1.37) and (1.38) this would cause a difference in the time-dependent decay rates of B_s^0 and $\overline{B}_s^0$ mesons. Since CP violation in mixing and in the decay are expected to be negligible in the case of $B_s^0 \to J/\psi\,\phi$, CP violation in the interference of these two is of special interest in the scope of this thesis, and measuring the phase ϕ in Equation (1.43) will be a central aspect.

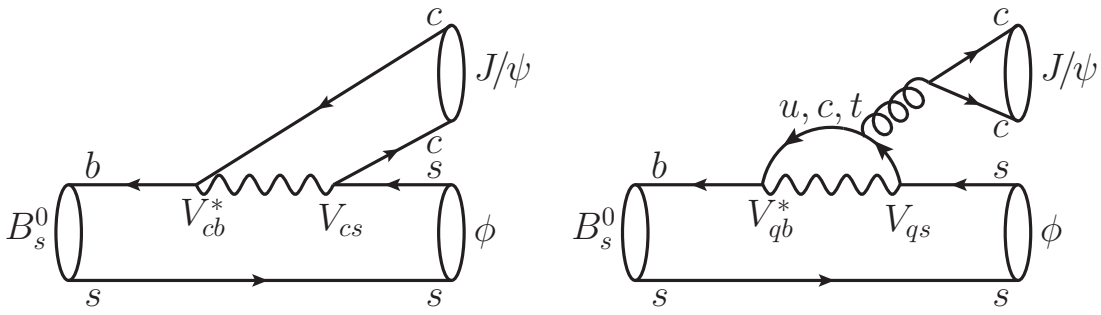


Figure 1.3: Feynman diagrams for the decay $B_s^0 \to J/\psi\,\phi$. The dominant tree amplitude (a) and the higher order penguin contributions (b) are shown.

1.5 The decay $B_s^0 \to J/\psi\,\phi$

As discussed in the previous section, decays of neutral mesons to CP eigenstates offer a rich phenomenology to study CP violation. One of the most interesting decay channels is the mode $B_s^0 \to J/\psi\,\phi$ for which the most relevant Feynman diagrams are shown in Figure 1.3. The ϕ meson is reconstructed via the decay to two charged kaons. As discussed later, there are also non-resonant $B_s^0 \to J/\psi K^+K^-$ decays contributing, but for simplicity in the following $B_s^0 \to J/\psi\,\phi$ will be used for both categories. In this section, the concepts presented in the previous chapter are applied to this decay, and further specific aspects of it are discussed.

1.5.1 CP violation

According to Figure 1.3 the amplitude $A_{J/\psi\phi}$ of a B_s^0 meson decaying to the $J/\psi\phi$ state can be written as:

$$\begin{aligned} A_{J/\psi\phi} &\approx V_{cs}V_{cb}^*\,T + V_{us}V_{ub}^*\,P_u + V_{cs}V_{cb}^*\,P_c + V_{ts}V_{tb}^*\,P_t \\ &= V_{cs}V_{cb}^*\,(T + P_c - P_t) + V_{us}V_{ub}^*\,(P_u - P_t), \end{aligned} \tag{1.44}$$

where T is the tree level and P_u, P_c and P_t are the respective penguin contributions. The unitarity of the CKM matrix, i.e. Equation (1.13), is exploited from the first to the second row. Contributions proportional to the small factor $V_{us}V_{ub}^*$ ($\mathcal{O}(\lambda^4)$)

are typically neglected, which allows then to give the following expression for the ratio $\overline{A}_f/A_f$:

$$\frac{\overline{A}_f}{A_f} = \eta_{CP}\frac{\overline{A}_{\overline{f}}}{A_f} \approx -\eta_{CP}\left|\frac{\overline{A}_f}{A_f}\right|\frac{V_{cs}V_{cb}^*}{V_{cs}^*V_{cb}} = -\eta_{CP}\left|\frac{\overline{A}_f}{A_f}\right| e^{2i\phi_D}, \tag{1.45}$$

where f denotes the final state $J/\psi\,\phi$ and η_{CP} its CP eigenvalue. The phase $\phi_D = \arg(V_{cs}V_{cb}^*)$ is called the weak decay phase.

For the other part of λ_f, namely the ratio q/p, Equation (1.19) can be used. Given that in the B_s system $|\Gamma_{12}| \ll |M_{12}|$ this equation can be simplified to:

$$\frac{q}{p} \approx -e^{-i\arg(M_{12})} = -e^{-i\phi_M}, \tag{1.46}$$

with ϕ_M being the weak mixing phase that is defined by the CKM elements of the box diagram in Figure 1.2. Since the top quark gives by far the dominant contribution, the phase can be approximated as:

$$\frac{q}{p} \approx -e^{-\phi_M} \approx -\frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}}. \tag{1.47}$$

The absolute value of q/p is assumed to be unity, which is equivalent to the assumption of no CP violation in mixing. This is supported by dedicated measurements of semileptonic asymmetries in the B_s^0 system [18].

Combining Equations (1.45) and (1.47) allows to give the following expression for $\lambda_{J/\psi\phi}$, which will be in the following only denoted as λ:

$$\begin{aligned}\lambda = \frac{q}{p}\frac{\overline{A}_f}{A_f} &\approx \eta_{CP}\left|\frac{\overline{A}_f}{A_f}\right|\frac{V_{cs}V_{cb}^*}{V_{cs}^*V_{cb}}\frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}} \\ &= \eta_{CP}\,|\lambda|\,e^{-i\phi_M+2\phi_D} \\ &= \eta_{CP}\,|\lambda|\,e^{-i\phi_s}.\end{aligned} \tag{1.48}$$

The phase difference $\phi_s = \phi_M - 2\phi_D$ can be related to the angle β_s of one of the CKM unitarity triangles, see Equation (1.14), as $\phi_s = -2\beta_s$. Within the SM the angle β_s can be precisely determined from global fits of the CKM matrix: $2\beta_s = (0.03686^{+0.00096}_{-0.00068})$ rad [19]. This precise prediction makes the measurement of β_s via the CP-violating phase difference ϕ_s an excellent probe of the SM.

Using the expression for λ of Equation (1.48) for Equations (1.37) and (1.38) yields the following time-dependent decay rates for the mode $B_s^0 \to J/\psi\,\phi$:

$$\frac{d\Gamma_{B_s^0\to J/\psi\phi}(t)}{dt} \propto \left|A_{J/\psi\phi}\right|^2 \frac{1}{1+C} e^{-\Gamma_s t}\Big[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + \eta_{J/\psi\phi}\, D \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + C\cos\left(\Delta m_s t\right) - \eta_{J/\psi\phi}\, S \sin\left(\Delta m_s t\right)\Big] \tag{1.49}$$

$$\frac{d\Gamma_{\bar{B}_s^0\to J/\psi\phi}(t)}{dt} \propto \left|A_{J/\psi\phi}\right|^2 \frac{1}{1+C} e^{-\Gamma_s t}\Big[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + \eta_{J/\psi\phi}\, D \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - C\cos\left(\Delta m_s t\right) + \eta_{J/\psi\phi}\, S \sin\left(\Delta m_s t\right)\Big], \tag{1.50}$$

with

$$C = \frac{1-|\lambda|^2}{1+|\lambda|^2},\quad S = -\frac{2\,|\lambda|}{1+|\lambda|^2}\sin\phi_s \quad \text{and} \quad D = -\frac{2\,|\lambda|}{1+|\lambda|^2}\cos\phi_s. \tag{1.51}$$

These time-dependent decay rates build the core of the measurement presented in this thesis. Figure 1.4 shows these rates when assuming realistic values for Δm_s, $\Delta\Gamma_s$ and Γ_s, a CP eigenvalue of 1 for the final state and CP violating parameters as $\lambda = 1$ and $\phi_s = 0.3\,\text{rad}$. The latter value is an order of magnitude larger than the SM prediction in order to visualize the oscillations. For the same reason, also an ideal reconstruction of the decay time and the initial B_s^0 flavour is assumed.

1.5.2 Polarization amplitudes

As shown in Equation (1.50), the time-dependent decay rates depend on the CP eigenvalue of the final state. The decay $B_s^0 \to J/\psi\,\phi$ is the decay of a pseudo scalar to two vectors and allows therefore for three different relative angular momenta (l=0,1,2) between the J/ψ and the ϕ meson. This leads to different CP eigenvalues according to:

$$\eta_{J/\psi\phi} = (-1)^l. \tag{1.52}$$

In order to be able to correctly describe the time-dependent decay rates, a disentangling of the different CP eigenstates is necessary. Since these components are related to different angular momentum states, the angular distributions of the final state particles allows such a separation. Three different combinations of polarization states of the J/ψ and ϕ meson are used as basis for this decomposition. They are

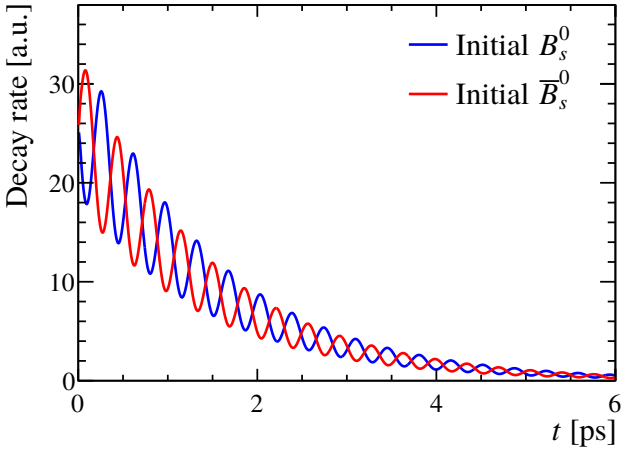


Figure 1.4: Simplified time-dependent decay rates of B_s^0 and $\overline{B}_s^0$ mesons to the final state $J/\psi\,\phi$. A final state CP eigenvalue of 1 is assumed. While the mass and decay-width splitting of the B_s^0 system are set to the current world averages [16], the CP-violating phase difference is set to 0.3 rad, which is one order of magnitude larger than the SM prediction. No experimental effects are taken into account that would dilute the oscillations.

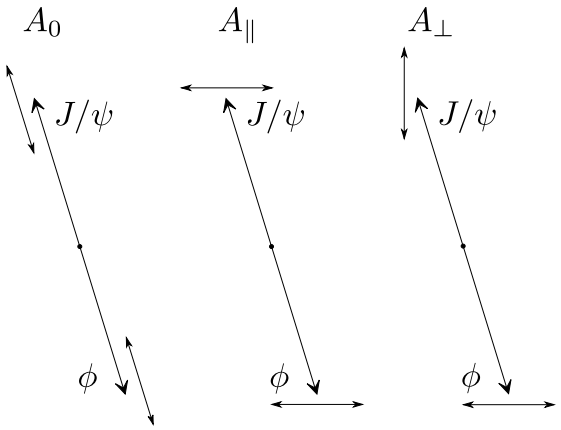


Figure 1.5: The three different polarization amplitudes of the $J/\psi\,\phi$ system in the decay $B_s^0 \to J/\psi\,\phi$. The short arrows indicate the spin orientation of the two vector mesons.

shown in Figure 1.5. The amplitude A_0 represents the case when both mesons have a longitudinal polarization, while $A_\perp$ and $A_\parallel$ label the states of perpendicular and parallel transverse polarizations, respectively. A_0 and $A_\parallel$ are even under CP, while $A_\perp$ has a relative angular momentum of 1 and is therefore CP-odd.

Each of the three amplitudes comes with a potentially different phase, which are labeled as δ_0, $\delta_\perp$ and $\delta_\parallel$. They originate from QCD interactions of the final

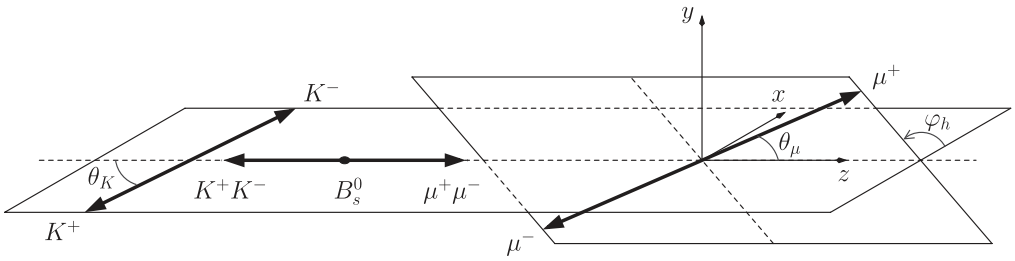


Figure 1.6: Definition of the angles in the helicity basis, which is used to describe the angular distribution of the $B_s^0 \to J/\psi\,\phi$ decay. Figure taken from Ref. [20].

states and are therefore also called strong phases and do not change sign under CP transformation. Since only phase differences are observable, the three phases reduce to two phase differences that were chosen to be $\delta_\perp - \delta_0$ and $\delta_\parallel - \delta_0$. These are the parameters that will be later quoted in the results.

The final state angular distributions are described by three angles that form the so-called helicity basis. Their definition is shown in Figure 1.6. The angle θ_μ is defined as the angle between the positively charged muon and the negative flight direction of the B_s meson in the center-of-mass system of the J/ψ meson. Similarly, the angle θ_K is defined for the positively charged kaon and the ϕ meson. Comparing the orientations of the two decay planes of the resonances in the B_s^0 center-of-mass system defines the angle φ_h. By convention, the angle between the side of the negatively charged kaon and the side of the positively charged muon is chosen.

S-wave

Besides the three polarization states of the $J/\psi\phi$ system, there is one more component that contributes to the data samples used in this analysis. Although the ϕ resonance dominates the two kaon system in the chosen mass range, there is a small scalar contribution that consists out of a non-resonant part and the scalar $f_0(980)$. This CP-odd component is called S-wave and comes with an additional strong phase that is labeled with δ_S. As for the other components, the absolute phase is not measurable, and only the phase difference to the phase of the perpendicular polarization amplitude will be later quoted as a result.

1.5.3 Time-dependent decay rate

Following the previous discussion of the four different components contributing to the decay $B_s^0 \to J/\psi\,\phi$, one can write down the full time- and angle-dependent decay

Table 1.2: Definition of prefactors in the time-dependent functions $h_k(t)$ of the differential $B_s^0 \to J/\psi\,\phi$ and $\overline{B}_s^0 \to J/\psi\,\phi$ decay rates, Equations (1.55) and (1.56).

k	a_k	b_k	c_k	d_k
1	1	D	C	$-S$
2	1	D	C	$-S$
3	1	$-D$	C	S
4	$C\sin(\delta_\perp - \delta_\parallel)$	$S\cos(\delta_\perp - \delta_\parallel)$	$\sin(\delta_\perp - \delta_\parallel)$	$D\cos(\delta_\perp - \delta_\parallel)$
5	$\cos(\delta_\parallel - \delta_0)$	$D\cos(\delta_\parallel - \delta_0)$	$C\cos(\delta_\parallel - \delta_0)$	$-S\cos(\delta_\parallel - \delta_0)$
6	$C\sin(\delta_\perp - \delta_0)$	$S\cos(\delta_\perp - \delta_0)$	$\sin(\delta_\perp - \delta_0)$	$D\cos(\delta_\perp - \delta_0)$
7	1	$-D$	C	S
8	$C\cos(\delta_\parallel - \delta_S)$	$S\sin(\delta_\parallel - \delta_S)$	$\cos(\delta_\parallel - \delta_S)$	$D\sin(\delta_\parallel - \delta_S)$
9	$\sin(\delta_\perp - \delta_S)$	$-D\sin(\delta_\perp - \delta_S)$	$C\sin(\delta_\perp - \delta_S)$	$S\sin(\delta_\perp - \delta_S)$
10	$C\cos(\delta_0 - \delta_S)$	$S\sin(\delta_0 - \delta_S)$	$\cos(\delta_0 - \delta_S)$	$D\sin(\delta_0 - \delta_S)$

rate [21]. Including interference between the components, there are in total ten different terms:

$$\frac{\mathrm{d}\Gamma(B_s^0 \to J/\psi\,\phi)}{\mathrm{d}t\,\mathrm{d}\theta_\mu\,\mathrm{d}\theta_K\,\mathrm{d}\varphi_h} \propto \sum_{k=1}^{10} A_k\, h_{k,+1}(t)\, f_k(\theta_\mu, \theta_K, \varphi_h), \tag{1.53}$$

$$\frac{\mathrm{d}\Gamma(\overline{B}_s^0 \to J/\psi\phi)}{\mathrm{d}t\,\mathrm{d}\theta_\mu\,\mathrm{d}\theta_K\,\mathrm{d}\varphi_h} \propto \sum_{k=1}^{10} A_k\, h_{k,-1}(t)\, f_k(\theta_\mu, \theta_K, \varphi_h). \tag{1.54}$$

The time-dependent terms $h_{k,+1}(t)$ and $h_{k,-1}(t)$ are derived from Equations (1.49) and (1.50) and can be written as:

$$h_{k,+1} = \frac{1}{1+C} e^{-\Gamma_s t}\Big[a_k \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + b_k \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + c_k \cos\left(\Delta m_s t\right) + d_k \sin\left(\Delta m_s t\right)\Big], \tag{1.55}$$

$$h_{k,-1} = \frac{1}{1+C} e^{-\Gamma_s t}\Big[a_k \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + b_k \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - c_k \cos\left(\Delta m_s t\right) - d_k \sin\left(\Delta m_s t\right)\Big]. \tag{1.56}$$

where the prefactors a_k, b_k , c_k and d_k are defined in Table 1.2. The angular functions $f_k(\theta_\mu, \theta_K, \varphi_h)$ and amplitudes A_k are given in Table 1.3.

1.5.4 Current experimental and theoretical status

Figure 1.7 shows the current experimental status for ϕ_s and $\Delta\Gamma_s$ together with the SM prediction. Especially for the CP-violating phase difference ϕ_s, the world

Table 1.3: Definition of the angular functions $f_k(\theta_\mu, \theta_K, \phi)$ and amplitudes of the differential $B_s^0 \to J/\psi\,\phi$ and $\overline{B}_s^0 \to J/\psi\,\phi$ decay rates, Equations (1.53) and (1.54).

k	A_k	$f_k(\theta_\mu, \theta_K, \varphi_h)$
1	$\lvert A_0\rvert^2$	$2\cos^2\theta_K \sin^2\theta_\mu$
2	$\lvert A_\parallel\rvert^2$	$\sin^2\theta_k(1-\sin^2\theta_\mu\cos^2\varphi_h)$
3	$\lvert A_\perp\rvert^2$	$\sin^2\theta_k(1-\sin^2\theta_\mu\sin^2\varphi_h)$
4	$\lvert A_\parallel A_\perp\rvert$	$\sin^2\theta_k \sin^2\theta_\mu \sin 2\varphi_h$
5	$\lvert A_0 A_\parallel\rvert$	$\frac{1}{2}\sqrt{2}\sin 2\theta_k \sin 2\theta_\mu \cos\varphi_h$
6	$\lvert A_0 A_\perp\rvert$	$-\frac{1}{2}\sqrt{2}\sin 2\theta_k \sin 2\theta_\mu \sin\varphi_h$
7	$\lvert A_S\rvert^2$	$\frac{2}{3}\sin^2\theta_\mu$
8	$\lvert A_S A_\parallel\rvert$	$\frac{1}{3}\sqrt{6}\sin\theta_k \sin 2\theta_\mu \cos\varphi_h$
9	$\lvert A_S A_\perp\rvert$	$-\frac{1}{3}\sqrt{6}\sin\theta_k \sin 2\theta_\mu \sin\varphi_h$
10	$\lvert A_S A_0\rvert$	$\frac{4}{3}\sqrt{3}\cos\theta_K \sin^2\theta_\mu$

average is dominated by the LHCb measurements using data corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$. These measurements contain the previous version of the analysis presented here [5] but also the study of the complementary decay $B_s^0 \to J/\psi\,\pi^+\pi^-$ [22]. Other relevant contributions come from the ATLAS and CMS collaborations [23, 24]. The numeric values for the world averages are given in Ref. [16]:

$$\Delta\Gamma_s^{\text{w.a.}} = (0.085 \pm 0.006)\,\mathrm{ps}^{-1},$$
$$\phi_s^{\text{w.a.}} = (-0.021 \pm 0.031)\,\mathrm{rad}.$$

These values are consistent with the theoretical prediction in the case of $\Delta\Gamma_s$ [25] and the indirect determination from global fits of the CKM matrix parameters in the case of ϕ_s [19]:

$$\Delta\Gamma_s^{\text{theo.}} = (0.088 \pm 0.020)\,\mathrm{ps}^{-1},$$
$$\phi_s^{\text{theo.}} = (-0.03686^{+0.00096}_{-0.00068})\,\mathrm{rad}.$$

For the latter, the experimental uncertainties are still significantly larger. Thus, any improvement of the experimental precision directly translates to an increased sensitivity for effects beyond the SM.

1.5.5 Impact of higher order diagrams

As discussed above, the interest in the channel $B_s^0 \to J/\psi\,\phi$ is, amongst others, justified by the precise constrain on the angle β_s that can be related to the CP-

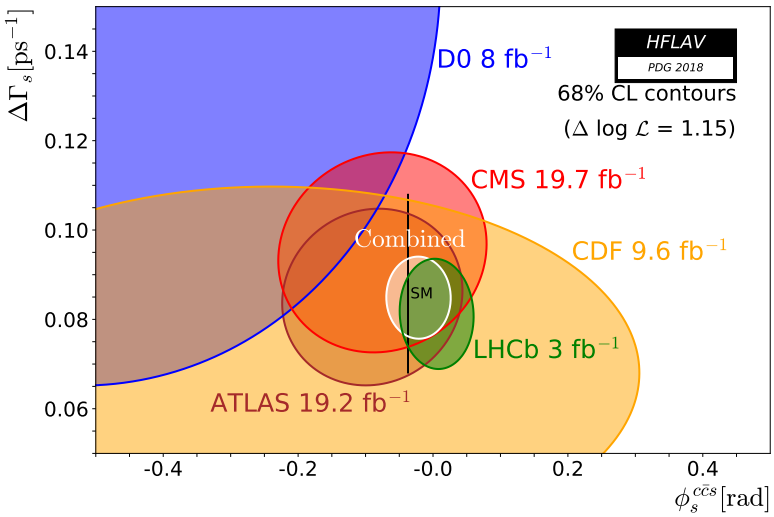


Figure 1.7: The experimental status of ϕ_s and $\Delta\Gamma_s$ measurements shortly before the analysis presented in this thesis was published. The black bar indicates the SM calculation and constraint of $\Delta\Gamma_s$ and ϕ_s, respectively. Figure taken from Ref. [16].

violating phase difference ϕ_s, which is observable in this channel. This relation, $\phi_s = -2\beta_s$, is only valid when terms proportional to $V_{us}V_{ub}^*$ can be neglected in Equation (1.44). In order to be able to compare the measured value of ϕ_s to the value obtained from indirect measurements, it is necessary to determine the contribution to ϕ_s of these higher order penguin diagrams, $\Delta\phi_s^{peng.}$:

$$\phi_s = -2\beta_s + \Delta\phi_s^{peng.} \tag{1.57}$$

It is not possible to reliably calculate the necessary hadronic quantities within QCD, but measurements of direct CP violation and branching fractions in related decay channels allow to put constraints on $\Delta\phi_s^{peng.}$. Following the strategy proposed in Ref. [26], the LHCb experiment employed the decays $B_s^0 \to J/\psi\,K^*(892)^0$ and $B^0 \to J/\psi\,\rho^0$ [27]. In these channels, the relative contributions of similar higher order penguin diagrams are enhanced with respect to the channel $B_s^0 \to J/\psi\,\phi$, which enables a precise determination of $\Delta\phi_s^{peng.}$. The obtained value is close to 0 with an uncertainty, depending on the polarization state, between 0.014 and 0.016 radian [27]. Although this is still significantly smaller than the experimental uncertainty on ϕ_s, a more precise determination of the penguin contributions will become necessary to keep up with future measurements of ϕ_s.

1.5.6 Possible non-Standard-Model contributions

The measured CP violation in the decay $B_s^0 \to J/\psi\,\phi$ can be affected by processes of New Physics contributing to the $B_s^0/\overline{B}_s^0$ mixing diagrams shown in Figure 1.2. An overview of New Physics models affecting this mixing processes can be found in [25]. Any such modification of the meson mixing can be parametrized by a change of the off-diagonal matrix element M_{12} in Equation (1.15). This is the starting point of a model-independent way to constrain New Physics contributions described in [28]. The change can be written as:

$$M_{12} = M_{12}^{SM} \cdot \Delta_s \text{ with } \Delta_s = |\Delta_s| e^{i\phi_s^{\Delta}}. \tag{1.58}$$

where M_{12}^{SM} is the SM value and any deviation from $\Delta_s = 1$ corresponds to a contribution of New Physics. The absolute value of Δ_s affects the mass splitting Δm_s and any nonzero phase ϕ_s^{Δ} modifies the value of ϕ_s according to:

$$\phi_s = \phi_s^{SM} + \phi_s^{\Delta}. \tag{1.59}$$

Current available measurements of ϕ_s, Δm_s and $\Delta\Gamma_s$ were used, together with other measurements, to perform a global fit that puts constrains on the real and imaginary part of Δ_s [28]. The corresponding contour plot is shown in Figure 1.8. The SM value of $\Delta_s = (1, 0)$ is in good agreement with the current best fit value, which is dominated by the constrains due to the mass splitting Δm_s and the phase difference $-2\beta_s \approx \phi_s$. While the width of the Δm_s band originates mainly from the uncertainty of the theoretical quantities, the width of the ϕ_s contribution is dominated by experimental uncertainties. This is a strong motivation to improve the measurement of ϕ_s.

1.5.7 Polarization-dependent CP violation

The time-dependent decay rates presented above assume that the CP-violation parameter λ is the same for all polarization amplitudes. However, some New Physics models may induce a polarization-dependent shift of the phase difference ϕ_s, see Ref. [29]. Therefore, an alternative parametrization of the decay rates with polarization-dependent CP violation is useful. In this parametrization, each

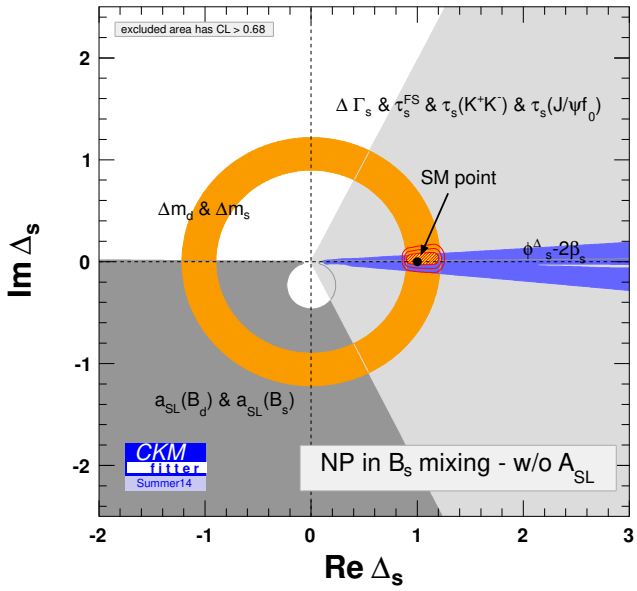


Figure 1.8: Constrains of the parameter Δ_s from a global fit to observables of the B_s^0 system. The red contours correspond to the 1, 2 and 3 sigma regions around the best fit value. The figure is taken from [28].

polarization, including the S-wave component, comes with a separate parameter λ^p that is given by:

$$\lambda^p = \frac{q}{p}\frac{\overline{A^p}}{A^p}, \tag{1.60}$$

where A^p and $\overline{A^p}$ are the decay amplitudes of a B_s^0 and a $\overline{B}_s^0$ meson to the polarization p, respectively. This means that there are separate values for $|\lambda^p|$ and ϕ_s^p for every polarization $p \in [0, \perp, \|]$ and the S-wave component $p = S$. Following Ref. [30], the time-dependent decay rate with polarization-dependent CP violation is given in Appendix A. While the polarization-independent CP-violation parametrization is used as nominal model throughout the analysis presented here, a crosscheck is performed using the more complex parametrization to test for any sign of polarization-dependent CP violation.

1.5.8 The B_s^0 lifetime

Besides the CP-violating phase difference ϕ_s, which can be precisely predicted from global fits of the CKM parameters, the lifetime-related parameters $\Delta\Gamma_s$ and Γ_s are two of the main physical quantities measured in the presented analysis. In

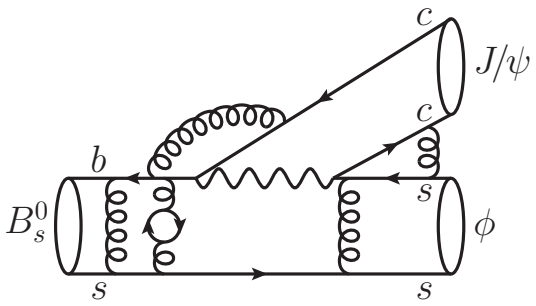


Figure 1.9: One of the Feynman diagrams representing possible QCD iterations within and between the initial and final state particles.

contrast to ϕ_s, these parameters lack a precise theoretical prediction. While the decay width of the most simple weak decay, the decay of a muon to an electron and two neutrinos, can be precisely calculated considering only the pure weak process and the phase space, decay-widths predictions of processes involving strongly bound objects in the initial and final state suffer from a range of QCD effects.

Figure 1.9 shows a version of Figure 1.3b where possible QCD interactions within and between the initial and final state particles are added. Such diagrams can not be calculated perturbatively. However, the Heavy Quark Expansion (HQE) formalism uses an expansion in $1/m(B_s^0)$ of an effective Hamiltonian to describe such processes [31]. Especially lifetime or decay-width ratios of different B mesons can be predicted with high precision within this formalism. Many corrections and related uncertainties cancel, which yields the following value for the decay-width ration of the B_s^0 and B^0 meson: $\Gamma_s/\Gamma_d = 1.0006 \pm 0.0025$ [32]. This precise prediction is consistent with the current experimental value, $\Gamma_s/\Gamma_d = 1.007 \pm 0.004$ [16]. Reaching now an experimental precision comparable to the prediction, makes this quantity an excellent testing case of HQE, which is used for many theoretical predictions in the B-meson systems.

The analysis presented here, is directly sensitive to the decay-width difference $\Gamma_d^s = \Gamma_s - \Gamma_d$, which can be easily related to the corresponding ratio when using the current experimental world average for the B^0 lifetime, $\tau_{B^0} = (1.520 \pm 0.004)$ ps [16].

2 The LHCb experiment

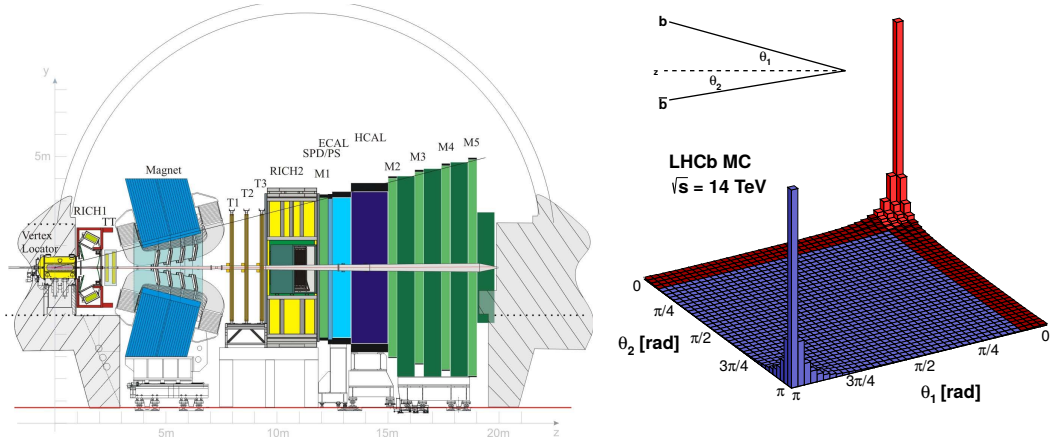
This chapter provides an overview of the LHCb experiment at the Large Hadron Collider. The purpose of this chapter is not to give a full-length review of all components of the experimental apertures, but to briefly introduce the elements and concepts that are relevant for the analysis presented in this thesis.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) operated by the European Organization for Nuclear Research (CERN) near Geneva is up-to-date the most powerful particle accelerator [33]. Its controlled operation started in 2010 with symmetric proton-proton collisions at a center-of-mass energy of 7 TeV. The circular accelerator has a circumference of 26.7 km and while the most runtime is reserved for proton-proton collisions, a sizable fraction is also spent on the acceleration and collision of lead nuclei.

Although so far not reached, the LHC is designed for a maximum proton-proton center-of-mass energy of 14 TeV with proton beams that each consist of up to 2808 bunches of approximately 10^{11} protons. This corresponds to a collision rate of up to 40 MHz. While the LHC was run at 7 to 8 TeV center-of-mass energy in the years 2010 to 2012, the following upgrade phase allowed to operate it at 13 TeV in the years 2015 to 2018. Currently the second upgrade phase is taking place, and in 2021 the LHC will finally resume operation with the design center-of-mass energy of 14 TeV.

In total, the LHC provides 4 interaction points at which the four main experiments are located: ALICE [34], ATLAS [35], CMS [36] and LHCb [37]. The analysis presented here is based on a proton-proton sample recorded in 2015 and 2016 by the LHCb experiment.



(a) (b)

Figure 2.1: Schematic view of the LHCb detector (a) and distribution of the angles between produced $b\bar{b}$ quark pairs and the beam axis in simulated proton-proton collisions (b). In (b) the LHCb acceptance is indicated in red. The figures are taken from Ref. [38, 39].

2.2 The LHCb detector

In comparison to the other three main experiments at the LHC, the LHCb experiment features a unique detector geometry. Is is shown in Figure 2.1a. While especially ATLAS and CMS are multipurpose detectors that cover a symmetric area around the proton-proton interaction point and mainly concentrate on the reconstruction of heavy particles that are produced with high momentum transverse to the beams, the LHCb detector is realized as a forward spectrometer covering an angle between 10 and approximately 300 mrad with respect to the beam line. This forward geometry is motivated by the research focus of the LHCb experiment, namely the study of heavy mesons and baryons containing bottom or charm quarks. These are produced mainly in the forward[1] direction as can be seen in Figure 2.1b. Each of the detector components is specifically designed to efficiently reconstruct and select decays of such heavy mesons and baryons and will be shortly described in the following.

[1] Forward represents here both directions along the beam pipe. For practical reasons the LHCb detector is equipped only in one direction.

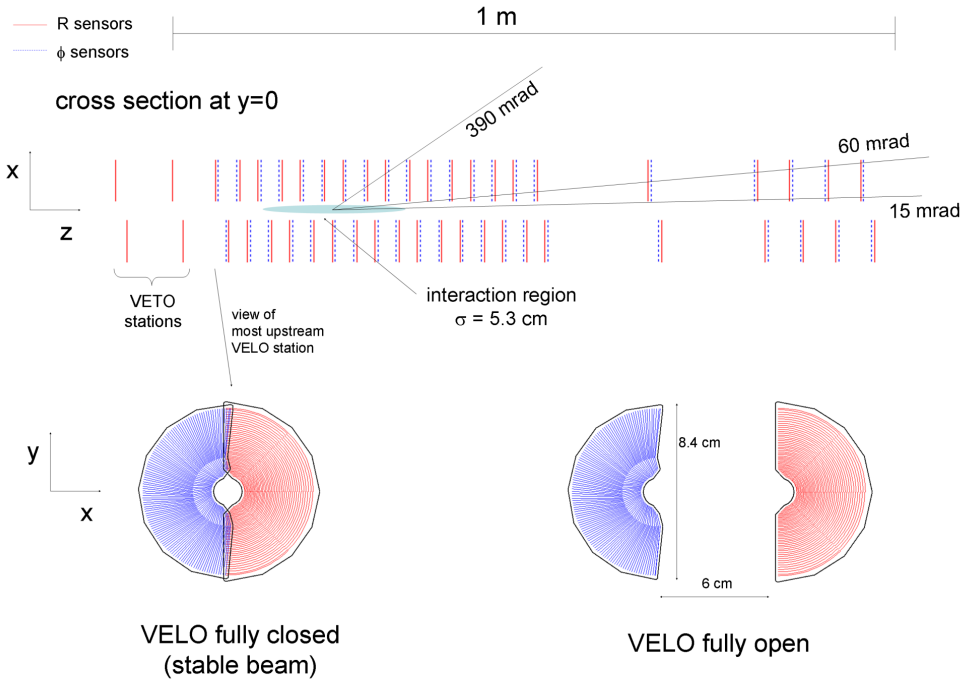


Figure 2.2: Schematic view of the Vertex Locator. The overall arrangement of the stations is shown on the top and the position and structure of the two halves of one station are shown at the bottom. The figure is taken from Ref. [44].

2.2.1 Tracking system

Most of the produced heavy mesons and baryons decay within picoseconds after their production and can therefore not be directly detected. Thus, their identification heavily relies on the precise reconstruction of their decay products, which originate from a common vertex and have an invariant mass matching the mass of the mother particle. This requires a precise measurement of the trajectory and the momentum of these particles. Within the LHCb detector and for charged particles, both is provided by the tracking system. It consists of a Vertex Locator [40] around the interaction point, two tracking stations before (Trigger-Tracker or TT) and three tracking stations behind [41, 42] (T1-T3) a dipole magnet [43], see Figure 2.1a.

The Vertex Locator

The Vertex Locator (VELO) is responsible for the reconstruction of particle trajectories (tracks) near the proton-proton interaction point. It consists of 21 circular stations arranged along the beam line, see Figure 2.2. Each of them contains two types of silicon-strip detectors. The so-called R sensors are oriented tangentially

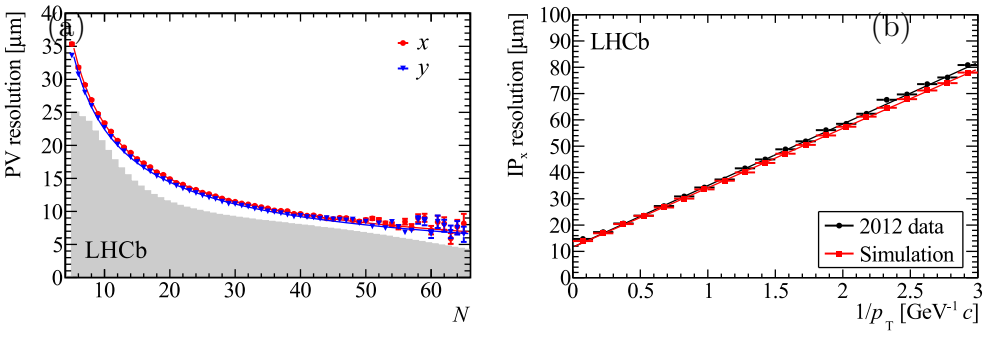


Figure 2.3: Resolution plots of the Vertex Locator. In (a) the primary vertex resolution perpendicular to the beam axis is shown as a function of the number of associated tracks. The gray histogram indicates the distribution of this number in data. In (b) the resolution of the impact parameter perpendicular to the beam axis is shown as a function of the transverse momentum of a particle. The figures are taken from Ref. [46].

and measure the distance to the beam line. In contrast, the ϕ sensors are arranged in radial direction. The sensors have a minimal distance of 7 mm to the beams and only a thin aluminum pipe separates the LHC vacuum and the detector. During the injection phase of the LHC, the VELO has to be removed from its nominal position in order to protect it from the unstable beams. For this purpose, the complete Vertex Locator consists of two halves that can be moved to and away from the beams, see Figure 2.2.

In a first step during the reconstruction of tracks inside the VELO, only the information of the R sensors is used [45]. Given that most of the particles passing the detector are produced directly at or close to the primary interaction point, their projection of the radial distance to the beam axis as a function of their z position can be approximated by a straight line. This allows to find track candidates by searching for four or three hits building a straight line in this projection. These track candidates are then further extended by adding additional hits that match to the extrapolation. Only after that, hits from the ϕ sensors are associated to the tracks.

Besides the reconstruction of individual tracks, the main purpose of the VELO is the identification and precise position measurement of vertices. The reconstruction of the primary vertex (PV) of the proton-proton collision and secondary vertices of the decays of relatively long-living particles, such as B-mesons, is crucial to identify the latter and to measure their decay time. Figure 2.3 shows the PV resolution in the two directions transverse to the beam axis and the impact parameter (IP) resolution in one of these directions. The IP is defined as the minimal distance of a track to the PV, and requiring a large IP is a way to identify tracks originating

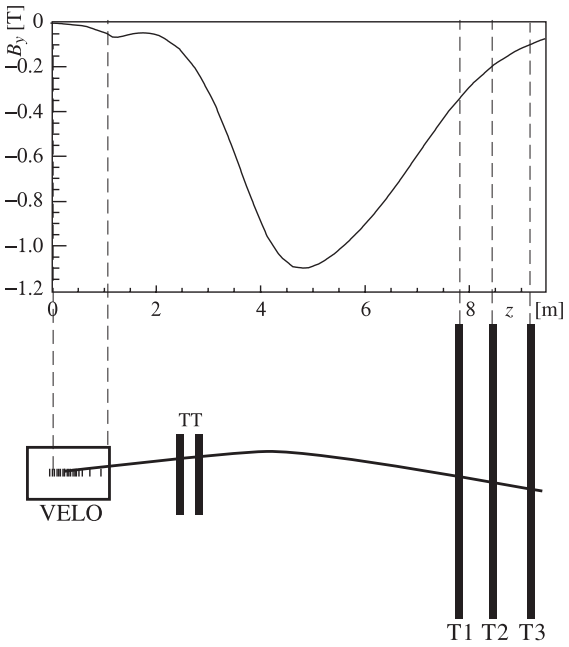
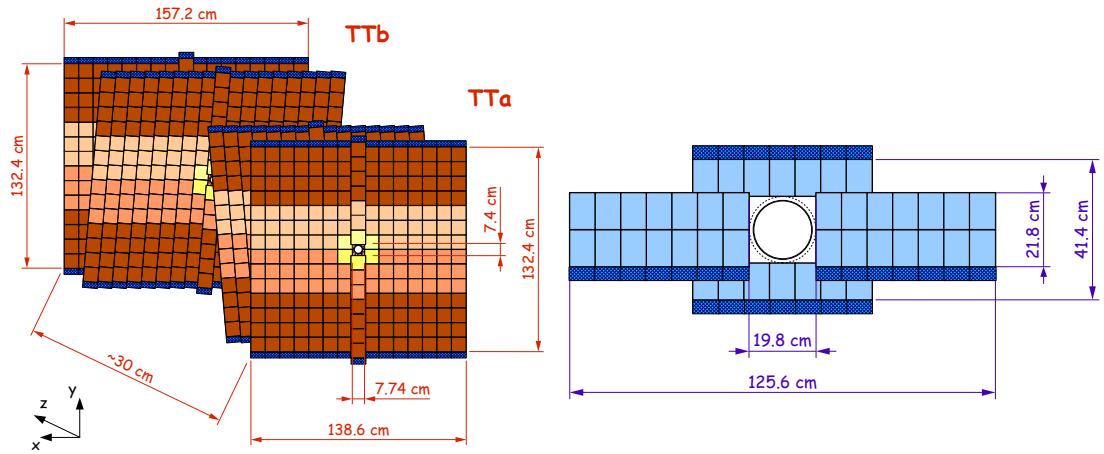


Figure 2.4: Main component of the magnetic field as a function of the z position. The figure is taken from Ref. [47].

from a secondary vertex. Due to the forward geometry of the LHCb detector, the respective spatial resolution along the beam axis are typically a factor 5-8 worse.

The magnet

While the position of charged particles near the collision point is precisely measured by the VELO, their momenta is so far completely unknown. For this, the curvature of the particle trajectories inside a magnetic field has to be measured. In the LHCb detector, this magnetic field is provided by a dipole magnet and is aligned with the y-axis, see Figure 2.1a. The polarity is flipped on a regular basis, which allows to control effects arising from potential asymmetries of the detector. In Figure 2.4, the magnetic field is shown as a function of the z position. For particles transversing the complete detector, the field adds up to 4 Tm. Inside the VELO the magnetic field is nearly negligible, and therefore no momentum measurement is possible using only hits in this detector. For this purpose, further tracking stations are located in front and especially behind the magnet, see Figure 2.4. They will be described in more detail in the following subsections.



(a) (b)

Figure 2.5: Schematic view of the four layers of the TT (a) and of one layer of the IT (b). The figures are taken from Ref. [48, 49].

Silicon trackers

The tracking stations before (TT) and the inner part of the tracking stations behind (IT) the dipole magnet are based on silicon micro-strips. These strips are oriented perpendicular to the bending plane and have a pitch of 183 µm and 197 µm for the TT and IT, respectively. Both sub-detectors are shown in Figure 2.5. While the IT covers only the most inner LHCb acceptance and is surrounded by the outer tracker, which will be described later, the TT detector covers the full acceptance. It is arranged in two stations of two layers each, of which the respective inner ones are tilted by 5° around the z axis in opposite directions. This allows a more precise position measurement along the y axis.

The TT stations are mainly needed to detect very low-momentum particles ($\lesssim 2\,\mathrm{GeV/c}$), which are bent out of the detector by the dipole magnet, and particles that originate from long-living neutral resonances that decay outside the VELO, like K_s^0 mesons or λ^0 baryons. TT measurements also improve the momentum resolution by further constraining the slope of the trajectories before the main magnetic field.

Outer tracker

Together with the IT, the outer tracker (OT) forms the tracking stations behind the dipole magnet, which are crucial for the momentum measurement. While the highly occupied inner region is covered by the high-resolution IT, the OT does not

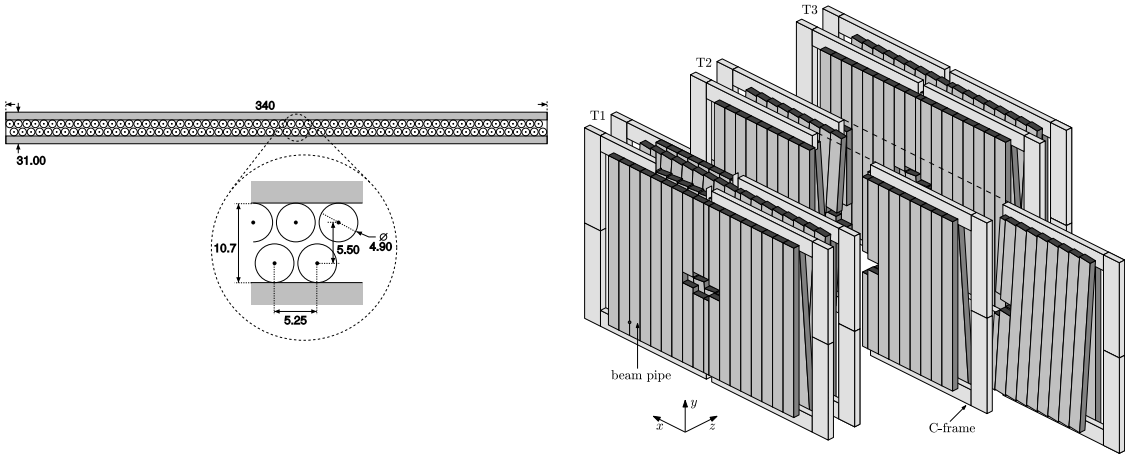


Figure 2.6: Schematic view of the cross section of a single OT layer on the left, and of the arrangement of the three OT stations on the right. The figures are taken from Ref. [50].

have to deal with these high multiplicities and provides a coarser granularity in the outer regions with lower activity. The OT is based on straw tubes arranged in two rows per layer. They have a pitch of 5.25 mm and a resolution of less than 200 µm is achieved in the direction perpendicular to the tubes. Figure 2.6 shows the cross section of a single layer and the arrangement of the full OT in three stations with four layers each. The OT covers an area of approximately $5 \times 6\,\mathrm{m}^2$ and as for the TT station, the respective inner layers are tilted by 5° around the z axis in opposite directions to allow the position measurement in the y direction.

Combining the tracking sub-detectors

The information of the tracking sub-detectors are combined by several algorithms to identify tracks, see e.g. [46]. There are different track types, describing from which sub-detectors the track is receiving hits. In the scope of this analysis, only the so-called long tracks are of interest. They have at least hits in the VELO and in the tracking stations behind the magnet and show the highest momentum resolution. After the hits associated to a track have been identified, a Kalman filter is performed to identify and remove outliers, and to obtain the optimal estimate of the track parameters including the final momentum estimate. For long tracks a resolution $\Delta p/p$ of 5 to 8 per mill for momenta between 20 and 100 GeV/c is achieved [46]. In addition, the Kalman filter provides a χ^2_{track} for the track fit describing the fit quality. This is the main input to a combination of the information from the tracking systems that is used to reject bad-quality track candidates, which are likely to not correspond to a real particle passing the detector.

2.2.2 Particle-species identification system

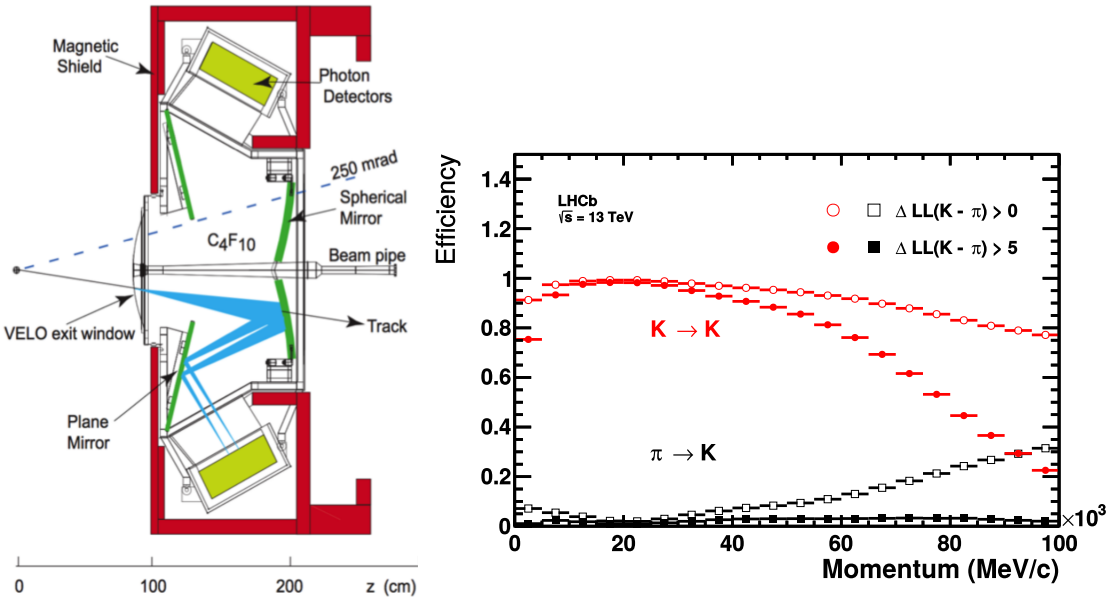
Especially for experimental flavour physics, the possibility discriminate between different particle species is essential. In order to identify and study a wide range of possible decays of heavy mesons and hadrons, it is crucial to correctly identify the particle types of the decay products. The LHCb detector provides a dedicated particle identification system that allows to distinguish between electrons, muons, charged pions and kaons, protons and photons. In the following, the different detectors, relevant for the particle identification, are shortly presented.

Ring-Imaging Cherenkov detectors

Two Ring-Imaging Cherenkov detectors (RICH1 and RICH2) are responsible for the identification of charged hadrons. Their functional principle is based on Cherenkov radiation, which occurs when charged particles transverse a material with a velocity that is larger than the speed of light in this medium. Depending on the velocity β of the particle and the refractive index n of the material, the photons are emitted under a specific angle that is given by $\cos\theta = 1/(\beta n)$. By measuring the angle θ, the velocity of the particle can be determined. When matching a signal in the RICH detectors to a track, its momentum estimate can be related to the mass and thereby to the species of the respective particle.

The RICH detectors are placed before and after the magnet and its surrounding tracking stations. They contain different radiator materials with different refractive indices. While the RICH1 detector in front of the magnet uses the gas C_4F_{10} and is able to discriminate particles in the momentum range between 2 and 50 GeV/c, the RICH2 detector is located behind the magnet and contains the gas CF_4, which allows to discriminate charged hadrons up to a momentum of 100 GeV/c. Besides the radiator material, both RICH detectors contain mirrors to guide the Cherenkov light away from the beam pipe and position resolving photon detectors that detect the light rings. Figure 2.7a shows the structure of the RICH1 detector.

For each of the RICH detectors, the detected light patterns are matched to the traversing tracks, and mass hypothesis are assigned. For each track and possible mass hypothesis a likelihood to match the observed pattern is calculated. Differences in this likelihood for the respective particle hypothesis are then used to discriminate between the species. Typically, the likelihood of the pion hypothesis is chosen as reference. Figure 2.7b shows the efficiency of the kaon hypothesis for two different cuts on the respective logarithmic likelihood difference (ΔLL). It is shown as a function of the momentum and separately for kaons and pions.

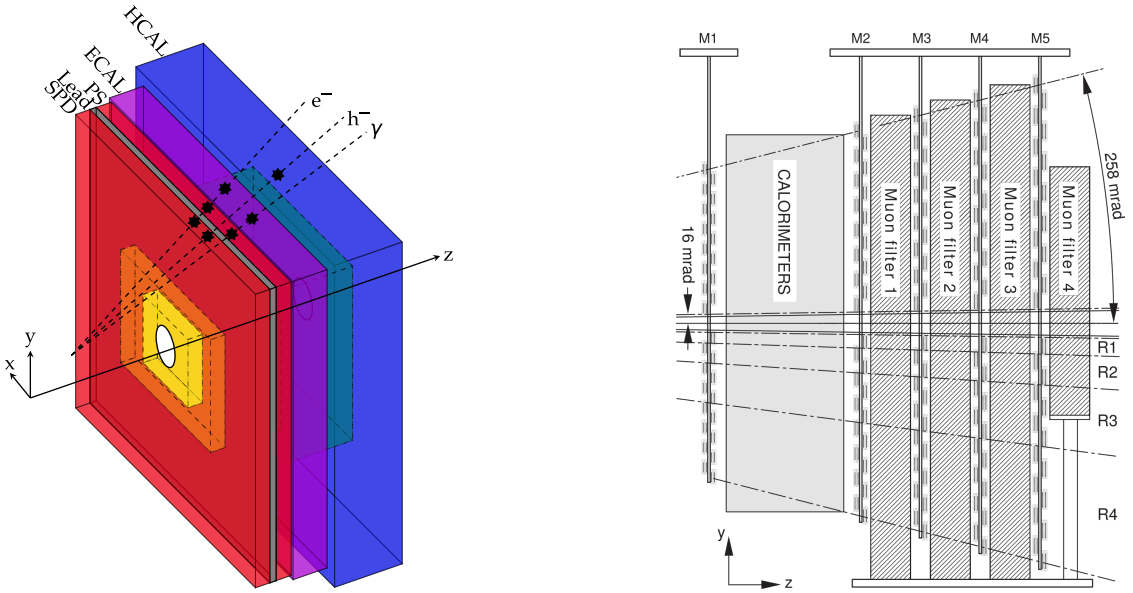


(a) (b)

Figure 2.7: Schematic view of the RICH1 detector (a) and efficiency of the kaon (miss)identification for different particles as a function of their momentum and for different cuts on the RICH response (b). The figures are taken from Ref. [51, 52].

Calorimeters

The calorimeter system at LHCb fulfills three main tasks. It measures energies of charged and neutral particles, it allows to discriminate light hadrons, electrons and photons, and it is a crucial part of the trigger system, see Section 2.2.3. The four calorimeter sub-systems are placed behind the RICH2 detector and are all based on alternating layers of absorber and scintillation material, see Figure 2.8a. Inside the Electromagnetic Calorimeter (ECAL) electrons and photons produce electromagnetic showers via bremsstrahlung and e^+e^- pair production and are typically fully absorbed. Their energy can be measured by collecting the light inside the scintillator layers and reading it out with photomultipliers. The Hadronic Calorimeter (HCAL) is located behind the ECAL and creates and detects hadronic showers of neutral and charged hadrons entering it. A Scintillating Pad Detector (SPD) and a Pre-Shower (PS) detector are located in front of the ECAL and are separated by an additional lead plate. Since the SPD only gives signal for charged particles, it helps to distinguish between electromagnetic showers created from photons and electrons. The PS detector supports the separation of electrons and light hadrons by their different shower behavior inside the lead layer.



(a) (b)

Figure 2.8: Schematic views of the calorimeter system in (a) and the muon detectors in (b). The figures are taken from Ref. [37, 53].

Muon system

In contrast to other charged particles, muons at the typical LHC energies behave like minimum ionizing particles and therefore pass through the whole calorimeter system without depositing a sizable amount of energy. Their identification is based on a dedicated muon system, which is located mostly behind the calorimeters, see Figure 2.8a. It consists mainly out of multi-wire proportional chambers, which are separated by 80 cm thick layers of iron aiming to stop high energy hadrons that might pass the calorimeter system. Charged particles passing the chambers ionize the gas inside, and the produced charge is collected and amplified.

Since the muon chambers can be read out at 40 MHz and provide a clean signal of high energy muons, they are a crucial part of the LHCb trigger, see Section 2.2.3. The muon chamber that is positioned in front of the calorimeter system significantly improves the transverse momentum resolution in the trigger.

Combining the sub-systems

As shortly describe for the RICH detectors, for every particle identification sub-detector per-track likelihoods for specific particle hypothesis can be calculated. Again, the pion hypothesis is chosen as reference since these are the most abundant

particles. The differences in logarithmic likelihood are then linearly combined and global particle identification variables are defined for every particle species X [46]:

$$\Delta\ln\mathcal{L}_{X\pi} = \ln\mathcal{L}(X) - \ln\mathcal{L}(\pi). \tag{2.1}$$

Here, $\ln\mathcal{L}(X)$ is the sum of logarithmic likelihoods of all sub-detectors for the particle hypothesis X.

A more advanced set of global particle identification variables was developed in the last years [54]. It is based on machine learning and employs neural networks to identify particle species. In addition to the logarithmic likelihood differences discussed above, further information coming from the particle identification and tracking detectors are used as input. The output variables of the neural nets are labeled as `ProbNN`X, where X represents one of the particle species.

2.2.3 Trigger system and data flow

As described above, the LHC provides proton bunch collisions at a rate of 40 MHz. With the current LHCb detector it is not possible to fully read out and record the information from all sub-detectors at this rate. Anyhow, only in one of approximately 1000 proton-proton collisions a $b\bar{b}$ quark pair is produced inside the LHCb acceptance. While this rate is fairly high in comparison to the production rate of heavier particles like the Higgs boson, the decay products of the produced B-hadrons have a similar momentum distribution as the underlying event. Therefore, it is challenging to identify and record only events that are likely to contain the decays of interest. In the following, the LHCb trigger system, which is performing this task, is shortly presented.

The strategy for triggering events within the LHCb experiment is based on three consecutive stages, which stepwise reduce the rate while including the information of increasingly more sub-detectors, see Figure 2.9. Only the muon chambers and the calorimeters can be read out at 40 MHz and are therefore the only detectors available for the first trigger stage. This level-zero (L0) hardware trigger is looking for high transverse components of the energy deposits in the hadronic or electromagnetic calorimeter, or for single muons or pairs of muons with a high transverse momentum. Typical calorimeter thresholds are $E_T \gtrapprox 3\,\text{GeV}$ for hadrons and $E_T \gtrapprox 2\,\text{GeV}$ for photons and electrons. For the muon triggers, typical threshold on the transverse momentum of a single particle are around $p_T \gtrapprox 1.2\,\text{GeV/c}$, and for a pair of muons the threshold is defined as $\sqrt{p_T^1\, p_T^2} \gtrapprox 1.1\,\text{GeV/c}$. This reduces the total rate to around 1 MHz, which is low enough to read out the full tracking system and enter

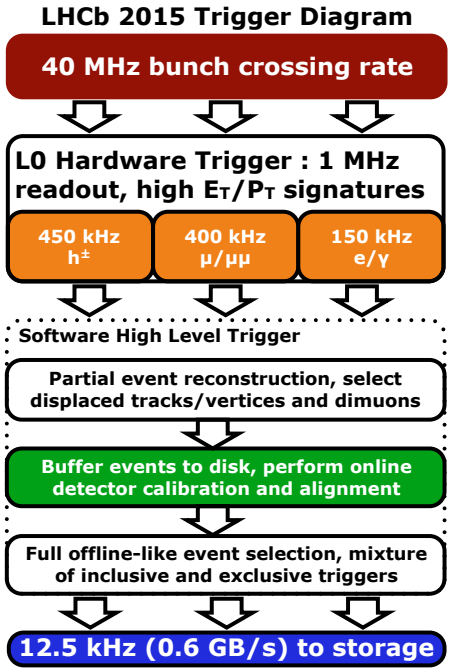


Figure 2.9: Flow diagram describing the structure of the LHCb trigger system. The figure is taken from Ref. [55].

the first stage of the software high level trigger (HLT1). In the HLT1 stage, tracks and primary vertices (PVs) are reconstructed. A more detailed discussion of the HLT1 trigger strategy, relevant for this thesis, can be found in Section 5.1. In general, the reconstruction of tracks and vertices allows to identify high momentum particles that have a large distance to any PV and are therefore likely to come from a long-living B or charm hadron. In addition, also two muons forming a good vertex and providing a high invariant mass are used to select an event. After the HLT1 stage, the output rate is reduced to approximately 150 kHz, which can then be written to a buffer. This allows to perform detector calibrations and alignments before entering the second and final stage of the software high level trigger (HLT2). Here, the full information of all sub-detectors is available and a comprehensive reconstruction of the event is performed. This allows to trigger on a range of inclusive and exclusive signatures of multi body decays of B and charm hadrons. This ranges from the identification of secondary vertices, requirements on the invariant mass of particle combinations as far as to the complete reconstruction of decay chains. A more detailed summary of the HLT2 selection that is relevant for the analysis presented here, is given in Section 5.1. The events passing the HLT2 stage correspond to a rate of 12.5 kHz, which can then be finally written to disk.

For a large fraction of the events passing the trigger, the complete detector output is stored and another offline reconstruction is performed. The other events are processed in the so-called `turbo stream` [56] where the reconstruction in the HLT2 stage is directly further processed. This allows to reduce the event size by an order of magnitude. Although the `turbo stream` is not used in the presented analysis, it will become more and more important and the default option after the current upgrade of the LHCb detector [57, 58].

Simulated samples

In particle physics, simulations of the decays and processes of interest and their signature in the detector are often a crucial ingredient for the analysis of experimental data. Key aspects of an analysis, like the detector efficiencies and resolutions can be studies and understood using these simulated samples. Therefore, it is important that the generated samples match the real data samples as closely as possible. The generation process can be split in two parts. The first is the simulation of the proton-proton collision with the Monte-Carlo event generator Pythia [59] and of the subsequent decays of unstable particles with the EvtGen library [60]. The second and more time consuming part is the simulation of the LHCb detector. It is realized with the Geant4 toolkit [61] as described in Ref. [62]. With the detector response at hand, the simulated events are processed in the exact same way as real data. This includes an emulation of the digitization of the analogous detector responses and of the complete trigger system.

3 Analysis overview

3.1 Analysis strategy

The analysis presented in this thesis aims to measure properties of the decay $B_s^0 \to J/\psi\,\phi$, where the J/ψ and ϕ mesons decay further to two charged muons and two charged kaons, respectively. Of main interest are the CP-violating phase difference ϕ_s, the decay width Γ_s and the decay-width splitting $\Delta\Gamma_s$. These parameters are all defined in Chapter 1. To determine them requires a measurement of the time-dependent decay rate separated according to the initial flavour of the $B_s^0/\overline{B}_s^0$ mesons. To disentangle the different CP components of the $J/\psi\,\phi$ final state, in addition an angular analysis is necessary. In the following, a rough outline of the structure of this thesis is given:

- Before starting with the actual analysis, Chapter 4 introduces some statistical tools and techniques that are employed throughout the following chapters.
- In order to be able to extract the physics parameters, it is necessary to have a pure sample of fully reconstructed $B_s^0 \to J/\psi\,\phi$ decays without any background originating from other particles being produced in the proton-proton collisions. In Chapter 5, the corresponding reconstruction, signal selection and background removal strategies are discussed.
- As mentioned above, the analysis requires a separation of the data set according to the initial flavour of the $B_s^0/\overline{B}_s^0$ mesons. Since the final state does not allow to deduce the initial flavour, dedicated methods were developed that allow to obtain this information by a reconstruction of other particles being produced in the respective proton-proton event. Chapter 6 summarizes these methods and their calibration.
- Given that a time-dependent angular analysis is performed, several detector acceptance and resolution effects have to be considered. They originate from the geometrical detector acceptance, the selection strategy and the intrinsic

measurement uncertainties. The determination of these detector effects is a central part of this thesis and is described in Chapter 7.

- Once a clean flavour-tagged signal sample with a correct description of the resolution and acceptance effects is obtained, in Chapter 8 a maximum likelihood fit to the decay-time and angular distributions is employed to extract the physics parameters of interest. In the same chapter, a more detailed description of the treatment of the S-wave component is given. As a consequence of this study, the maximum likelihood fit is simultaneously performed in bins of the invariant mass of the two kaon system.

- After presenting the results of this fit in Chapter 9, a detailed summary of the considered systematic uncertainties is given in Chapter 10.

This analysis is based on previous studies of the channel $B_s^0 \to J/\psi\,\phi$ with data collected in Run I by the LHCb experiment [5]. Several aspects were improved or completely redeveloped. An important change is the way the decay-time acceptance is determined. It is obtained from the topological similar channel $B^0 \to J/\psi\,K^{*0}$, which leads to the interesting fact that a direct measurement of the decay-width difference between the B_s^0 and B^0 meson, $\Delta\Gamma_d^s = \Gamma_s - \Gamma_d$, is possible. This difference can be related to the ratio Γ_s/Γ_d, which is of much more theoretical interest than the individual decay width of the B_s^0 system. Details are given in Section 7.2.

3.2 Own contribution to the official LHCb analysis

The content of this thesis is part of an official LHCb analysis that is close to be published in Ref. [63], was first shown at [64] and is documented in detail in the internal note Ref. [65]. For such a complex study it is common that a team of several people contribute and that not every part of it can be done by a single person. Nevertheless, it is crucial that important parts of the analysis are performed independently as a cross-check by at least two scientists. Given that everything is consistent, the choice of the exact numbers that will then enter in the final official result is arbitrary.

The situation discussed above also applies to the analysis presented here. Starting after the selection of signal decays and exploiting the flavour tagging that was developed and crosschecked by the LHCb collaboration over the last years, this thesis presents an autonomous study of the decay $B_s^0 \to J/\psi\,\phi$. The decay-time and angular

acceptance and the decay-time resolution determined here are close to the ones used in the official analysis, while the final fit result and the statistical subtraction of background candidates were only used as a cross-check. For consistency reasons, it was decided to present the results of this autonomous study rather than the one of the official LHCb analysis. A comparison of the two sets of results is shown in Appendix G.3.

3.3 Probability density function of the decay $B_s^0 \to J/\psi\,\phi$

The extraction of the physics parameters is strongly based on a correct description of decay-time and angular distributions of the selected $B_s^0 \to J/\psi\,\phi$ decays. The parametrization of this distribution is called probability density function (PDF) and will be developed throughout the following chapters. The starting point is the underlying PDF as it would be present if the complete information about all $B_s^0 \to J/\psi\,\phi$ decays would be available. It is essentially given by Equations (1.53) and (1.54):

$$\mathrm{PDF}(t, \Omega|q) = \frac{1}{N_q} \sum_{k=1}^{10} A_k\, h_{k,q}(t)\, f_k(\Omega), \tag{3.1}$$

where Ω represents the three angles of the helicity basis and $q = \pm 1$ corresponds to the initial flavour of the $B_s^0/\overline{B}_s^0$ meson. The absolute amplitudes squared $A^2_{\perp,0,\parallel,S}$ are parametrized by a S-wave fraction, $F_S^j = A_S^2$, for every bin in $m(K^-K^+)$ and values for $|A_\perp|^2$ and $|A_0|^2$ representing the respective fraction of the resonant component. The parallel component is fixed by $A_\perp^2 + A_0^2 + A_\parallel^2 = 1$. N_q is a normalization factor that is given by:

$$N_q = \int_{t=0}^{\infty} \int_{\Omega} \mathrm{PDF}(t, \Omega|q)\, \mathrm{d}\Omega\, \mathrm{d}t. \tag{3.2}$$

This PDF has to be modified when introducing flavour tagging, detector acceptances and resolution effects. At the end of the respective chapter or section, the relevant modifications of the PDF are given.

4 Analysis tools

4.1 Efficient selection: Boosted decision trees

The task to solve

A typical situation at the beginning of an analysis based on data from a collider experiment is a sample of signal candidates that is swamped by background processes. It is crucial to effectively discriminate between these two contributions and to obtain a signal sample as large and pure as possible. While the classical approach is based on the optimization of a set of rectangular cuts on some of the properties of the candidates, the method presented here allows to automatically consider correlations between these properties. In addition, it reduces the final optimization decision between signal efficiency and background rejection rate to the simple choice of a cut value on a single classification variable.

Machine learning

The algorithms presented here fall in the class of supervised machine learning. This means that the algorithm is trained to discriminate between signal and background candidates using a set of labeled candidates of these two categories. Typically, these training data sets are obtained from control regions in data or from simulation. An independent set of such samples can then be used to get an unbiased estimate of the performance of the algorithm.

Decision trees

We consider a training sample N that consists of two species, labeled as $y = 1$ and $y = -1$, which have a set of properties $\boldsymbol{x}$. A decision tree (DT) aims to create regions, called leafs, in the property space and classifies them as either $y = 1$ or $y = -1$. These leafs are defined in an iterative procedure that is based on binary decisions in one of the properties x_i. Figure 4.1 shows a simple example of a DT. In every step k, the algorithm searches for the optimal combination of one property

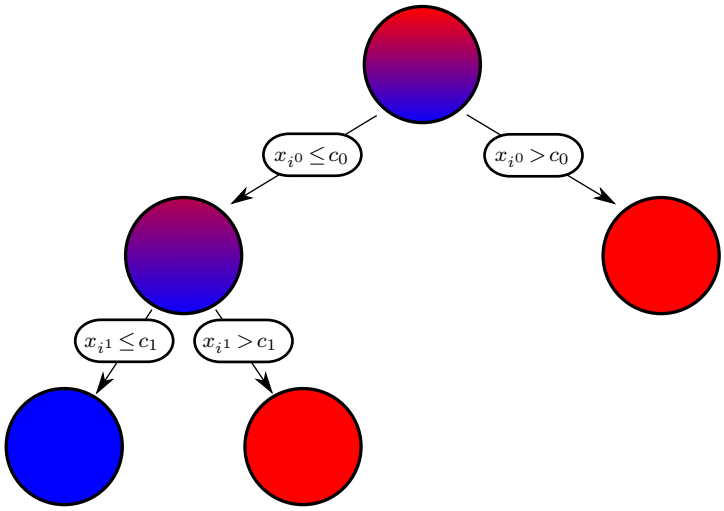


Figure 4.1: A simple example of a decision tree. The colors blue and red represent the two species that are separated by splits in the variables x_{i^0/i^1} at values $c_{0/1}$.

x_{i^k} and a respective cut value c_k such that the chosen metric is minimized. An example for such a metric is the sum of squares of the difference between the predicted species and the true species for all elements n in the leaf $\tilde{N} \subset N$ that is currently processed:

$$\sum_{\substack{n \in \tilde{N} \\ x_{i^k} \leq c_k}} (\tilde{y}_1 - y_n)^2 + \sum_{\substack{n \in \tilde{N} \\ x_{i^k} > c_k}} (\tilde{y}_2 - y_n)^2. \tag{4.1}$$

Here, $y_n \in \{1, -1\}$ is the actual species of the element n and $\tilde{y}_{1/2}$ are the predicted species in the specific leaf. This prediction is typically defined as the species that is more abundant in this region. In this case, the metric is therefore directly proportional to the number of wrongly assigned species hypothesis.

The same concepts can be also applied to regression trees, which try to predict a continuous variable y instead of a binary classification. Typically, the average y value of the entries of the training sample in a leaf is chosen as predicted value. A decision and regression tree is then completely defined by the parameters $\{(i^0, c_0), ..., (i^l, c_l)\}$, i.e. the cut values c_k and the properties x_{i^k} to cut on. The number of those cuts, l, depends on the depth of the tree.

Boosting and gradient boosting

The decision trees presented previously can in principle perfectly solve the task of classifying a training sample. However, they suffer from instability under small variations of this training sample. To mitigate this effect and ensure a good per-

formance also on an independent sample, the method of boosting is employed [66]. The concept of boosting involves the sequential combination of many relatively weak classification or regression algorithms, called weak learners, to obtain a more powerful, but still robust, overall algorithm. One way to formulate a boosting algorithm, is the gradient boosting method [67], which is discussed in more detail below.

We consider again a training data set with N entries that have the properties $\boldsymbol{x}$ and y. The aim is to find a function F, such that $F(\boldsymbol{x})$ infers the variable y of an entry based on its other properties $\boldsymbol{x}$. Given a general loss function $L(F(\boldsymbol{x}_1), y_1, ..., F(\boldsymbol{x}_N), y_N)$ that measures the deviation between the predicted and true values of y, the gradient boosting algorithm tries to minimize L in terms of a gradient descent method, in which the gradients are approximated by weak learners. An example for such a loss function is the metric given in Equation (4.1):

$$L(F(\boldsymbol{x}_1), y_1, ..., F(\boldsymbol{x}_N), y_N) = \sum_{i=1}^{N} (F(\boldsymbol{x}_i) - y_i)^2, \tag{4.2}$$

but in general any loss function can be used. In the case discussed here, the weak learners, $\phi(\boldsymbol{x}, \boldsymbol{\theta})$, are the previously introduced regression trees that are described by the parameters $\boldsymbol{\theta} = \{(i^0, c_0), ..., (i^l, c_l)\}$, see Figure 4.1.

The first step of the boosting is to fit a weak learner, $\phi(\boldsymbol{x}, \boldsymbol{\theta_0})$, to the training data, which is then the first estimate $F_0(\boldsymbol{x})$ of the desired relation between y and $\boldsymbol{x}$. The following three steps, see Figure 4.2 for illustration, are then repeated M times to sequentially improve this approximation:
For $m = 1$, $m < M$:

1. Calculate the gradient $\boldsymbol{r^m}$ of the loss function L with respect to the prediction of the current model:

$$r_i^m = - \left[\frac{\partial L(F(\boldsymbol{x}_1), y_1, ..., F(\boldsymbol{x}_N), y_N)}{\partial F(\boldsymbol{x}_i)} \right]_{F=F_{m-1}} \tag{4.3}$$

   In the case of the loss function given in Equation (4.2), these residuals are given for every element i of the training sample by:

$$r_i^m = -2[F_{m-1}(\boldsymbol{x}_i) - y_i]. \tag{4.4}$$

2. Fit another weak learner, $\phi(\boldsymbol{x}, \boldsymbol{\theta_m})$, to this gradient

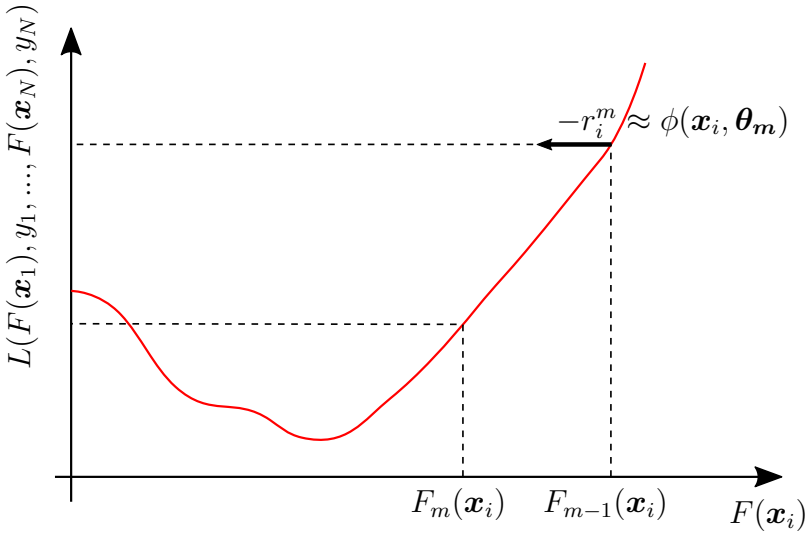


Figure 4.2: Schematic view of one step during the gradient boosting technique. The red line indicates the value of the loss function L evaluated for the training sample. The x axis represents one dimension of the high dimensional space $F(\boldsymbol{x}_i)$, with $i \in \{1, ..., N\}$.

3. Update the estimate of the relation between y and x:

$$F_m(\boldsymbol{x}) = F_{m-1}(\boldsymbol{x}) + \nu_m \phi(\boldsymbol{x}, \boldsymbol{\theta_m}), \tag{4.5}$$

where ν_m is a real parameter that can be determined using line search to minimize the loss function.

In this way, the final prediction of y based on $\boldsymbol{x}$ is given by the linear combination of the output of many weak learners:

$$F(\boldsymbol{x}) = \sum_{m=1}^{M} \nu_m \phi(\boldsymbol{x}, \boldsymbol{\theta_m}), \tag{4.6}$$

and minimizes the defined loss function L. During the boosting iterations, the step parameters ν_m are typically scaled by a number of the interval $(0, 1]$. This procedure is called shrinkage and, although more weak learners have to be combined, makes the boosting more robust.

In the analysis presented in this thesis, the implementation of gradient boosting within the TMVA framework [68] is used to discriminate between signal and background candidates in data. Although the classifier consists solely out of regression trees, such a classifier is usually called a boosted decision tree (BDT).

4.2 Boosted weighting

The task to solve

Several components of the analysis presented in this thesis rely on the weighting of one data set such that the distributions of some of its variables agree with the respective ones of another data set. This is for example needed when correcting simulated samples to better describe data. To obtain such weights for one or two variables at the same time is trivial and can be achieved by dividing histograms containing the respective distributions. However, higher dimensional agreement can often not be achieved with this method, since the histograms are increasingly sparsely populated. To overcome this limitation, a method is presented that is based on boosted regression trees [69].

Boosted weighting

The principle idea to overcome the problem of high dimensional weighting is to sequentially apply weak weighting algorithms, and thereby gradually correct the high dimensional distribution. This is implemented as a boosted regression tree, see Section 4.1, where each regression tree is trained to minimize the symmetrized χ^2 of its leafs. Given the normalized distributions of two data sets D_1 and D_2, the symmetrized χ^2 is defined as:

$$\chi^2 = \sum_{\text{leafs}} \frac{\left(N_{\text{leaf}}^{D_1} - N_{\text{leaf}}^{D_2}\right)^2}{N_{\text{leaf}}^{D_1} + N_{\text{leaf}}^{D_2}}, \tag{4.7}$$

where $N_{\text{leaf}}^{D_{1/2}}$ are the normalized number of entries in the respective leaf for the two data sets. The prediction of the regression tree in the respective leaf is then given by the ratio $N_{\text{leaf}}^{D_2}/N_{\text{leaf}}^{D_1}$. Once the first regression tree is trained, its prediction is used to weight the data sample D_1. This procedure is iteratively repeated with replacing the previous version of the sample D_1 with the newly weighted version of it. Again, as for the gradient boosting method, a shrinkage factor η is applied such that the applied weights are modified to be $(N_{\text{leaf}}^{D_2}/N_{\text{leaf}}^{D_1})^{\eta}$. A shrinkage factor significantly smaller than 1 helps to stabilize the weighting algorithm. If not explicitly stated differently, the boosted weighting technique is employed as default weighting tool throughout this thesis.

4.3 Parameter estimate: Maximum likelihood method

The task to solve

The heart of this thesis is the extraction of certain physics parameters from the measured decay times and angles of a sample of $B_s^0 \to J/\psi\,\phi$ decays. Given a set of values for these physics parameters, the expected distribution of these observables is known, see for example Equation (3.1). However, the opposite direction, i.e. to infer the values of the physics parameters from the observed distributions, is more challenging. In the following, the concept of maximum likelihood estimation is introduced[1] as a way to achieve this inversion.

Maximum likelihood estimation

Given a sample $\boldsymbol{X}$ of N independent measurements of some set of observables $\boldsymbol{x}$ that are distributed according to a normalized probability density function $\mathrm{PDF}(\boldsymbol{x}|\boldsymbol{\theta})$ with some unknown parameters $\boldsymbol{\theta}$, the likelihood $\mathcal{L}$ of this sample given the parameters $\tilde{\boldsymbol{\theta}}$ is defined as:

$$\mathcal{L}(\boldsymbol{X}|\tilde{\boldsymbol{\theta}}) = \prod_{i=1}^{N} \mathrm{PDF}(\boldsymbol{X}_i|\tilde{\boldsymbol{\theta}}). \tag{4.8}$$

According to Bayes theorem

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}, \tag{4.9}$$

which connects the conditional probabilities $p(A|B)$ and $p(B|A)$, this likelihood, interpreted as probability of the observed data $\boldsymbol{X}$ given the parameters $\tilde{\boldsymbol{\theta}}$, can be transformed to an estimator of the parameters given the data:

$$p(\tilde{\boldsymbol{\theta}}|\boldsymbol{X}) = \frac{\mathcal{L}(\boldsymbol{X}|\tilde{\boldsymbol{\theta}})\,p(\tilde{\boldsymbol{\theta}})}{p(\boldsymbol{X})}. \tag{4.10}$$

The prior $p(\tilde{\boldsymbol{\theta}})$ is typically chosen to be flat and $p(\boldsymbol{X})$ is given by the normalization. The best estimates of the parameters $\tilde{\boldsymbol{\theta}}$ are therefore obtained by maximizing the

[1] For a detailed discussion of this concept see e.g. Ref. [70].

likelihood function or, what is typically done, by minimizing the negative logarithmic likelihood:

$$\log \mathcal{L}(\boldsymbol{X}|\tilde{\boldsymbol{\theta}}) = \sum_{i=1}^{N} \log \left[\mathrm{PDF}(\boldsymbol{X}_i|\tilde{\boldsymbol{\theta}}) \right]. \tag{4.11}$$

It can be shown that in the limit of large N, the likelihood function converges to a normal distribution that can be used to estimate the true underlying parameters $\boldsymbol{\theta}$. Therefore, a natural uncertainty estimate corresponding to a one σ interval is given by the parameter value where the negative logarithmic likelihood has changed by halve a unit with respect to the minimum. This also allows to determine asymmetric uncertainties in case the likelihood shape is not symmetric. In case that each element $\boldsymbol{X}_i$ of the sample $\boldsymbol{X}$ features a weight w_i, the likelihood function is modified according to:

$$\mathcal{L}(\boldsymbol{X}|\tilde{\boldsymbol{\theta}}) = \prod_{i=1}^{N} \mathrm{PDF}(\boldsymbol{X}_i|\tilde{\boldsymbol{\theta}})^{w_i}. \tag{4.12}$$

To ensure that the uncertainty estimate is not biased by the overall scale of these weights, they have to be normalized such that [71]:

$$\frac{\sum_{i=1}^{N} w_i}{\sum_{i=1}^{N} w_i^2} = 1. \tag{4.13}$$

In case of absolute yield estimations, the maximum likelihood method has to be extended [70]. The normalized PDF is replaced by the one that contains the yields $\boldsymbol{\mathcal{N}}$ of every species (e.g. signal and background) present in the PDF as parameters and a Poisson factor is added:

$$\mathcal{L}(\boldsymbol{X}|\tilde{\boldsymbol{\theta}}) = \frac{e^{-\sum \boldsymbol{\mathcal{N}}}}{N!} \prod_{i=1}^{N} \mathrm{PDF}(\boldsymbol{X}_i|\tilde{\boldsymbol{\theta}}, \boldsymbol{\mathcal{N}}), \tag{4.14}$$

where $\sum \boldsymbol{\mathcal{N}}$ is the sum of all yields. The term $N!$ is constant and thus can be dropped during the minimization.

The maximum likelihood method also allows to include external constraints on the parameters $\tilde{\boldsymbol{\theta}}$. Given a set of parameters $\tilde{\boldsymbol{\theta}}' \subset \tilde{\boldsymbol{\theta}}$, for which a prior knowledge, represented by the values $\boldsymbol{\nu}$ and the according covariance matrix $\boldsymbol{C_\nu}$, should

be considered, the likelihood can be adjusted by adding a Gaussian constraint according to:

$$\mathcal{L}(\boldsymbol{X}|\tilde{\boldsymbol{\theta}}) = \frac{1}{\tilde{N}} e^{-\frac{1}{2}(\tilde{\boldsymbol{\theta}}'-\boldsymbol{\nu})^{\mathrm{T}}\boldsymbol{C}^{-1}(\tilde{\boldsymbol{\theta}}'-\boldsymbol{\nu})} \prod_{i=1}^{N} \mathrm{PDF}(\boldsymbol{X}_i|\tilde{\boldsymbol{\theta}}), \tag{4.15}$$

where $\tilde{N}$ is the normalization of the multidimensional Gaussian.

Throughout this thesis, the minimization of the negative logarithmic likelihood and the uncertainty estimates are based on the MINUIT [72] implementation within the ROOT framework [73].

4.4 Statistical background subtraction: ${}_s\mathcal{P}lot$

The task to solve

Although the boosted decision trees described in Section 4.1 are typically quite powerful in reducing the amount of background candidates, a pure signal sample is practically never obtained. Under certain conditions the remaining background candidates can be statistically subtracted, in order to anyhow extract the distributions of the pure signal sample. The ${}_s\mathcal{P}lot$ technique [74] constitutes the optimal way to achieve this and is presented in the following.

The concept of background subtraction.

We assume a data set with N entries consisting of two species, each of them having values for two uncorrelated observables x and y. Their respective distributions are shown in Figure 4.3. Species 1, in red and called signal, is distributed according to a normal distribution with width 1 and mean value 0 in y and according to a normal distribution with width 2 and mean value 1 in x. Species 2, in green and called background, is distributed uniformly in y and according to a normal distribution with width 2 and mean -1 in x. The aim is to get an estimate of the x distribution of the signal without knowing the underlying true distribution of the two species in this variable. However, we assume that the respective distributions in the control variable y are known. This information allows to statistically subtract the background component from the overall x distribution. The most straight forward way to achieve this in this simple case is the *sideband subtraction*. In y, there are regions ($|y| > 5$) that are to very good approximation purely populated by background. When adding the entries of these regions to the remaining entries with

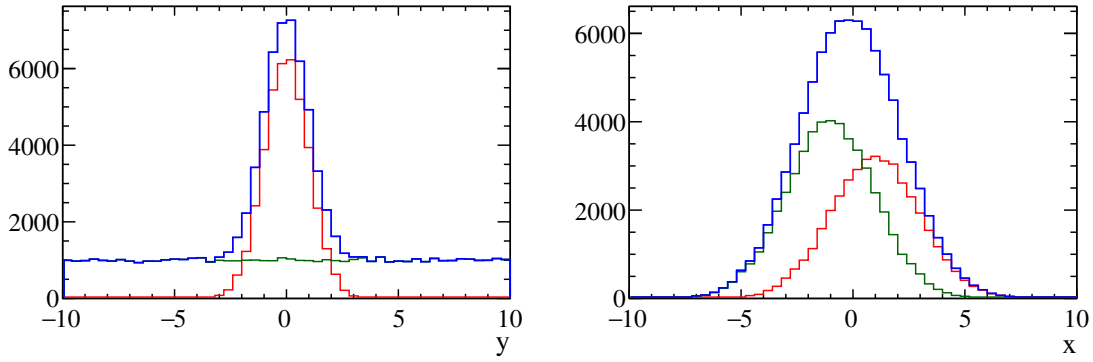
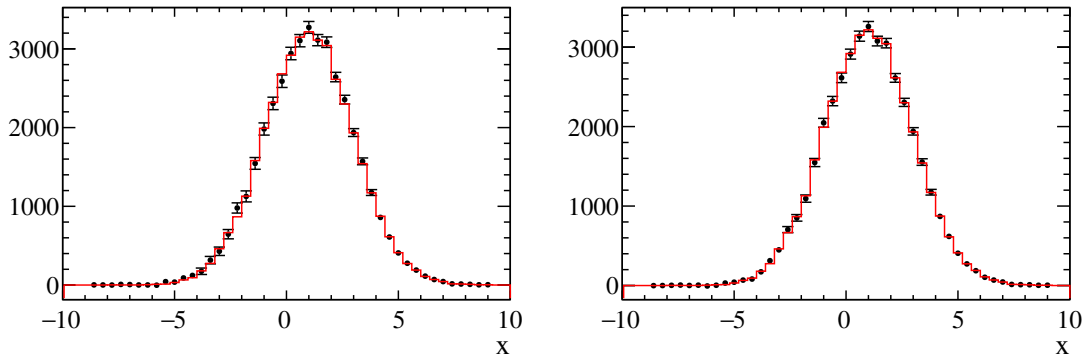


Figure 4.3: Toy distributions used to demonstrate the techniques of statistical background subtraction. The signal and background distributions in the two variables x and y are shown in red and green, respectively. The combined distribution is shown in blue.



(a) (b)

Figure 4.4: Signal distributions in the variable x. The red histogram show the true generated distribution, while the black points are obtained by the respective statistical background-subtracted full sample. In (a) and (b) the results of the *sideband subtraction* and ${}_s\mathcal{P}lot$ technique are shown, respectively.

a weight of -1, the resulting distribution in x approximates the true distribution of signal in this variable. This is shown in Figure 4.4a.

The *sideband subtraction* technique does only work if there are regions in the control variable that are populated purely by species 2. On top, there is no unique way how to apply it. The exact choice of these regions is arbitrary and every choice leads to a slightly different background-subtracted signal sample.

The ${}_s\mathcal{P}lot$ technique

The ${}_s\mathcal{P}lot$ technique generalizes the method of *sideband subtraction* by providing a weight w_i for every entry of the data set such that the weighted distribution resamples the background-subtracted distribution. In addition, the weights are chosen such that effective size N_{eff} defined as [71]:

$$N_{eff} = \frac{\left(\sum_{i=1}^{N} w_i\right)^2}{\sum_{i=1}^{N} w_i^2}, \tag{4.16}$$

is maximized. Given an estimate of the number N_s of entries of each of the S species and an according covariance matrix $\boldsymbol{C}$, the weights w_i^n to obtain a background-subtracted sample for species n are given by:

$$w_i^n = \frac{\sum_{s=1}^{S} \boldsymbol{C}_{ns}\, f_s(y_i)}{\sum_{s=1}^{S} N_s\, f_s(y_i)}. \tag{4.17}$$

Here, $f_s(y_i)$ is the know distribution of species s in the control variable(s) y_i. One way to obtain the yields N_s and their covariance matrix is to perform an extended maximum likelihood fit, see Section 4.3, to the control variables. In the simple example introduced above, the ${}_s\mathcal{P}lot$ technique results in the background-subtracted signal distribution shown in Figure 4.4b.

It is crucial to note that the assumption of no correlation within one species between the control variables and the variable of interest x, for which a background-subtracted sample has to be obtained, has to hold for the *sideband subtraction* as well as for the ${}_s\mathcal{P}lot$ technique.

5 Selecting $B_s^0 \to J/\psi\,\phi$ decays

In the proton-proton collisions at the LHCb interaction point many different particles are produced. Most of them are pions and other light particles, and only a small fraction of the events contains a B_s^0 meson. Of these B_s^0 mesons, only approximately one in a thousand decays to the final state $J/\psi\,\phi$ [3]. In this chapter, the strategy to obtain a large and background free sample of $B_s^0 \to J/\psi\,\phi$ decays is presented. The first step of the selection is happening in the LHCb trigger system. Only events that pass all stages of this trigger are stored and can be further analyzed. In total, roughly 60 billion complete proton-proton events were recorded in the years 2015 and 2016 and build the basis from which $B_s^0 \to J/\psi\,\phi$ candidates have to be selected. Those events are filtered in a LHCb wide offline selection procedure, which is afterwards further refined. A multivariate classifier is trained and applied to reject most of the remaining background candidates. In the last section of this chapter, the ${}_s\mathcal{P}lot$ technique is applied to statistically remove the remaining fake $B_s^0 \to J/\psi\,\phi$ candidates.

5.1 Triggering

According to the description in Section 2.2.3, the LHCb experiment has a three stage trigger system. In principle it is not necessary to specify which exact trigger line has sparked the respective trigger stage. However, the requirements in the trigger strongly influence the acceptances, see Chapter 7, and a fixed set of trigger lines that are required to have triggered the events helps to understand and describe these detector effects. In the following, the chosen configurations are summarized.

The L0 trigger selection

In the L0 trigger roughly 87% of the finally selected events containing $B_s^0 \to J/\psi\,\phi$ candidates were triggered by the muons originating from the J/ψ meson. For those, either one of the muons fulfilled a minimum requirement on the transverse momentum or the product of the transverse momentum of both muons was above

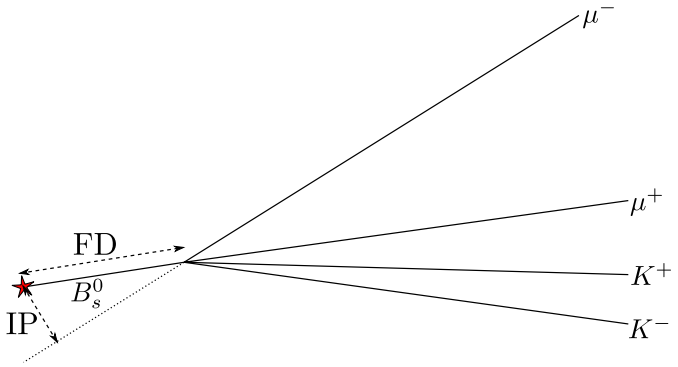


Figure 5.1: Decay topology of the decay $B_s^0 \to J/\psi\,\phi$. The red star indicates the origin vertex of the B_s^0 meson that typically is directly produced in the primary proton-proton collision point (PV). In addition, the flight distance (FD) of the B_s^0 meson and the impact parameters (IP) of one of the muons with respect to the PV are indicated.

a certain threshold. Since the L0 trigger requirements are typically rather loose compared to the requirements in the succeeding selection steps and are therefore not expected to have large impact on the acceptances, it was decided to do not select any specific set of L0 trigger lines, but keep all events, independently of how they were triggered. The remaining 13% of the selected $B_s^0 \to J/\psi\,\phi$ decays are therefore reconstructed in events that are triggered by either the kaons of the ϕ decay or by any other particle produced in the underlying proton-proton collision.

The HLT1 selection

Already during the first online stage of the LHCb trigger system, a complete reconstruction of tracks with high transverse momentum is performed. This allows to access information about momenta and trajectories of individual particles. In addition, the primary proton-proton interaction points (PVs) are reconstructed, and muons can be identified by tracks that have associated hits in the muon stations. Based on this, general selection strategies were developed to effectively trigger on a range of inclusive signatures.

Figure 5.1 provides a schematic view of a $B_s^0 \to J/\psi\,\phi$ decay at the track level. There are two principle ways that are used to select such a decay topology. The first is based on the two muons that form a vertex and have a relatively high invariant mass. Muons typically leave a clear signature in the detector and are much less often produced directly in the proton-proton collision than kaons or pions. The according trigger line is called `Hlt1DiMuonHighMass` and besides requirements on the vertex quality of the two muons and their invariant mass, it contains cuts on

the momentum and transverse momentum of the two muons. Detailed values are given in Appendix B.

The other strategy is based on the relatively long lifetimes of weakly decaying particles. A B_s^0 meson with a momentum of 100 GeV has an average flight distance of 8.5 mm, which can be resolved by the VELO and is used to discriminate such decays from particles originating directly from the PV. Although the B_s^0 decay vertex is not completely reconstructed in HLT1, properties of the final state tracks are related to the flight distance. For a single track the impact parameter IP is defined as its minimal distance to the PV, see Figure 5.1. A related quantity is the impact parameters significance χ^2_{IP}, which is defined as the difference in χ^2 when fitting a PV with and without this track.

The trigger line `Hlt1TrackMuon` requires a muon with high transverse momentum and large χ^2_{IP} with respect to any of the PVs in this event. In contrast to that, the line `Hlt1TwoTrackMVA` relies on a vertex reconstructed from two tracks with high transverse momentum. The minimal significance of the separation to any PV and other vertex and track properties are combined in a multivariate classifier. In Appendix B a detailed list of the variables used in these two trigger lines is given.

Since the last two trigger lines are more efficient for B_s^0 mesons with a long flight distance, they introduce a decay-time-dependent efficiency, which will bias the measured lifetime. This is not the case for the first trigger line. Therefore, throughout the analysis many data sets will be split in two categories:

- `Unbiased` candidates that are triggered by the `Hlt1DiMuonHighMass` line,
- `Biased` candidates that are triggered by either the `Hlt1TrackMuon` or the `Hlt1TwoTrackMVA` line.

Separate decay-time and angular acceptances will be determined for these two categories.

The HLT2 selection

In the second HLT2 trigger stage the rate is low enough to enable a nearly complete reconstruction of the events. This allows to trigger directly on J/ψ candidates, which constitute a relatively clean signature for the decay $B_s^0 \to J/\psi\,\phi$. Only a very loose cut on the significance of the vertex separation to any PV is applied to suppress J/ψ mesons that are directly produced in the primary proton-proton collision. As will be shown in Section 7.2.1, this requirement on the vertex separation significance has nearly no impact on the decay-time-dependent efficiency observed for the candidates

Table 5.1: Selection criteria used to identify $B_s^0 \to J/\psi\,\phi$ candidates.

	Variable	Cut
all tracks	$\chi^2_{\text{track}}/\text{nDoF}$	< 4
$J/\psi \to \mu^+\mu^-$	$m(\mu^+\mu^-)$	$\in [3017, 3177]\,\text{MeV/c}^2$
	$\Delta\ln\mathcal{L}_{\mu\pi}$ $(\mu^\pm)$	> 0
	$\chi^2_{\text{vtx}}/\text{nDoF}$	< 16
	χ^2_{DOCA}	< 20
	p_T $(\mu^\pm)$	> 500 MeV/c
$\phi \to K^+K^-$	$m(K^+K^-)$	$\in [990, 1050]\,\text{MeV/c}^2$
	$\Delta\ln\mathcal{L}_{K\pi}$ (K^+)	> 0
	$\chi^2_{\text{vtx}}/\text{nDoF}$	< 25
	χ^2_{DOCA}	< 30
	p_T (ϕ)	> 500 MeV/c
$B_s^0 \to J/\psi\,\phi$	$m(J/\psi\,K^+K^-)$	$\in [5150, 5570]\text{MeV/c}^2$
	$\chi^2_{\text{vtx}}/\text{nDoF}$	< 20
	$m_{\text{DTF}}(J/\psi\,K^+K^-)$	$\in [5200, 5550]\text{MeV/c}^2$
	t_{DTF}	$\in [0.3, 15]$ ps

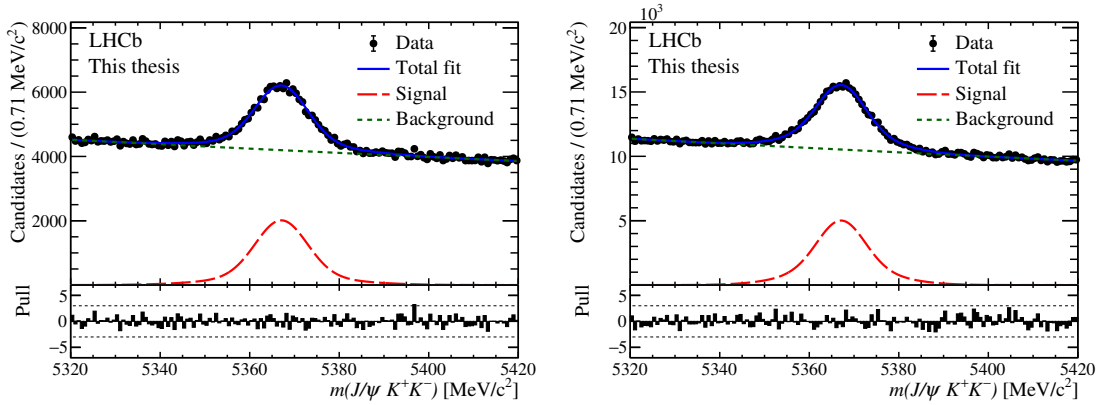
passing the final offline selection. Again, in Appendix B a more detailed summary of the requirements is given.

5.2 Offline selection and reconstruction

All events selected by the described LHCb trigger configurations are written to disk, and the full event reconstruction is repeated. This includes the reconstruction of tracks, particle identification information and PVs. Starting from these objects, $B_s^0 \to J/\psi\,\phi$ candidates are reconstructed and a first loose cut based selection is applied. These criteria are summarized in Table 5.1.

Two muon candidates are combined to form a J/ψ candidate with a mass that lies in a $160\,\text{MeV/c}^2$ window around the nominal J/ψ mass [3]. A loose cut on the difference of the logarithmic likelihood for the muon versus the pion hypothesis, see Section 2.2.2, is applied to suppress misidentified particles. To reject J/ψ candidates formed from random tracks, the χ^2 per degree of freedom of the vertex fit, $\chi^2_{\text{vtx}}/\text{nDoF}$, and the significance of the distance of closest approach of the two muon tracks, χ^2_{DOCA}, are required to be small.

In full analogy, a ϕ candidate is reconstructed. The respective invariant mass is required to lie in a $60\,\text{MeV/c}^2$ window around the nominal ϕ mass [3]. To further clean the sample, minimum transverse momentum cuts are placed on the ϕ meson and the two muons. Combining J/ψ and ϕ candidates that form a vertex with



(a) 2015 data (b) 2016 data

Figure 5.2: Invariant mass distribution of $B_s^0 \to J/\psi\,\phi$ candidates after the first offline selection. Projections of a fit to the data are overlaid.

reasonable χ^2_{vtx}/nDoF and that have an invariant mass that is close to the nominal B_s^0 mass, allows to build candidates of the decay $B_s^0 \to J/\psi\,\phi$.

These candidates are then associated to one of the PVs in the event. First, all PVs are reconstructed again after removing the tracks that are used to build the B_s^0 candidate. This avoids any bias of the PV position towards the B_s^0 decay vertex position. Afterwards, the χ^2_{IP} of the B_s^0 candidate with respect to every PV is calculated and the one with the smallest value is assigned. Thereafter, the four final state tracks are again combined using the so-called Decay Tree Fitter (DTF) [75]. The DTF allows to include additional constraints in this combination that help to increase the mass, momentum and vertex resolution. In the case of the decay $B_s^0 \to J/\psi\,\phi$, the mass of the J/ψ meson is constrained to the world average [3], and the B_s^0 meson is required to point to the associated PV. In addition to a more precise B_s^0 mass estimate, the DTF uses the vertex separation $|\boldsymbol{x}_{B_s^0} - \boldsymbol{x}_{PV}|$ and the momentum estimate $p_{B_s^0}$ of the B_s^0 meson to determine a value for its decay time t according to:

$$t = m_{B_s^0} \frac{|\boldsymbol{x}_{B_s^0} - \boldsymbol{x}_{PV}|}{p_{B_s^0}}. \tag{5.1}$$

To suppress prompt background, this decay time is required to be larger than 0.3 ps. An upper bound on the reconstructed decay time of 15 ps is set because no significant amount of true B_s^0 meson decays are expected beyond this.

Figure 5.2 shows the invariant mass distribution of the selected $B_s^0 \to J/\psi\,\phi$ candidates for the two years of data taking. Besides a clear peak from true

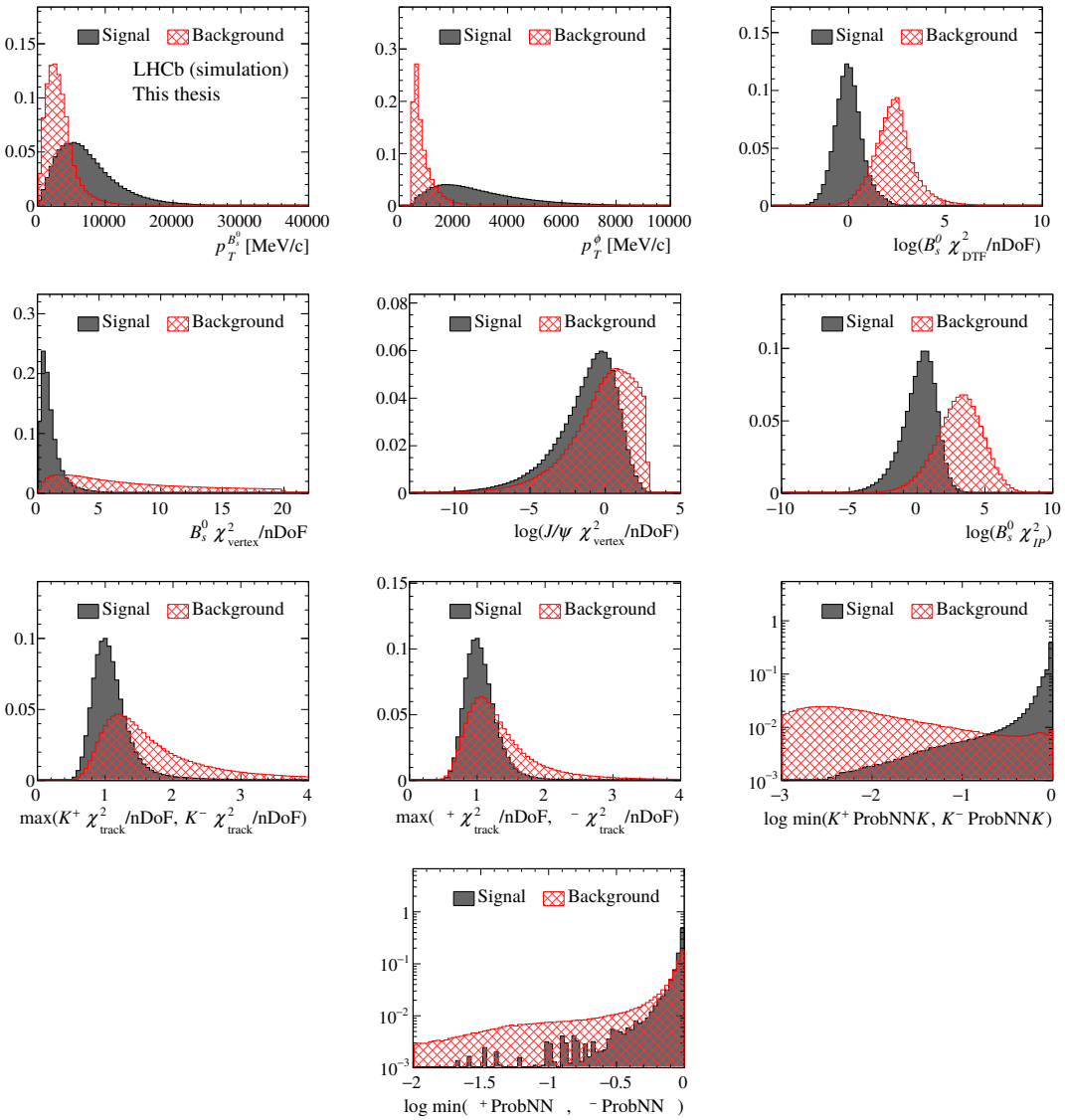


Figure 5.3: Signal and background distributions of the variables included in the multivariate classifier. Their definition is given in the text.

$B_s^0 \to J/\psi\,\phi$ decays, a large background component is visible. The main source of this background consists of random tracks that are wrongly combined to form a decay candidate. In order to further suppress this so-called combinatorial background, variables are identified that have a significant discrimination power between true signal decays and these background candidates. This is achieved by comparing a sample of simulated $B_s^0 \to J/\psi\,\phi$ decays to a pure background sample obtained from selected $B_s^0 \to J/\psi\,\phi$ candidates in data that have an invariant mass above 5450 MeV/c^2. Figure 5.3 shows the respective distributions for the selected variables. They are given by:

1. The transverse momenta $p_T^{B_s^0}$ and p_T^{ϕ} of the B_s^0 and ϕ candidates.
2. The χ^2 per degree of freedom of the DTF , χ^2_{DTF}/nDoF.
3. The χ^2 per degree of freedom of the B_s^0 and J/ψ vertex fits , B_s^0 χ^2_{vertex}/nDoF and J/ψ χ^2_{vertex}/nDoF.
4. The impact parameter significance of the B_s^0 candidate with respect to the PV, B_s^0 χ^2_{IP}.
5. The respective maximum χ^2 per degree of freedom of the track fits, χ^2_{track}/nDoF, of the two kaons and muons.
6. The minimum particle identification variable `ProbNN`K of the two kaons.
7. The minimum particle identification variable `ProbNN`μ of the two muons.

Variables such as the flight distance of the B_s^0 candidate or the impact parameter of the daughter particles have a high separation power but are strongly correlated to the decay time and would cause strong decay-time acceptance effects. Therefore, these variables are not considered.

Based on the chosen discriminating variables, a boosted decision tree (BDT), see Section 4.1, is trained to achieve a high signal efficiency while removing most of the background. This training relies on a sample of events that are clearly labeled as either signal or background. Again, the high mass sideband is used as a proxy for the overall combinatorial background, while simulated signal decays are representing the signal decays in data. To achieve the optimal separation power on real data, the simulated signal sample is required to be as similar as possible to the signal sample in data. This is ensured in terms of a weighting of the simulated sample to a statistically background-subtracted signal sample. This relatively clean signal data sample is obtained with the ${}_s\mathcal{P}lot$ technique, see Chapter 4, which is based on an extended maximum likelihood fit to the invariant mass distribution of the $B_s^0 \to J/\psi\,\phi$ data sample. The projection of this fit can be seen in Figure 5.2. Two Gaussian distributions with the same mean and an exponential function are used to model the signal and background component, respectively.

Figure 5.4 shows the background-subtracted data and the simulated sample for a set of variables for which these two do not perfectly agree. In the same figure, the distributions are shown after a multidimensional weighting in these variables is applied. No significant differences are visible afterwards. It is important to note that any differences between the simulated signal sample and the real decays could

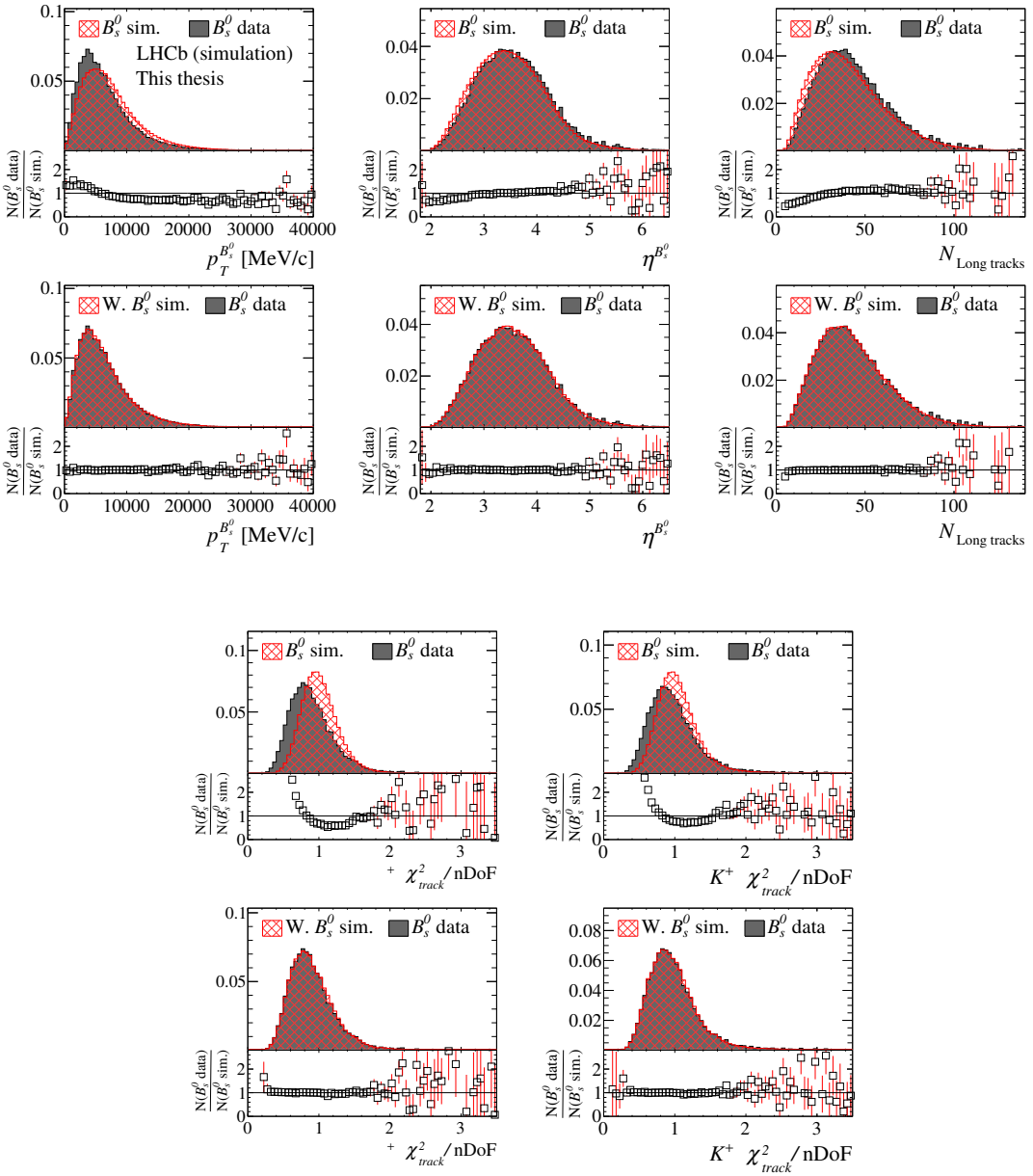


Figure 5.4: Distributions of $B_s^0 \to J/\psi\,\phi$ decays in background-subtracted data and simulation. The respective lower plot show the distributions after the simulated sample is weighted to match data in these variables.

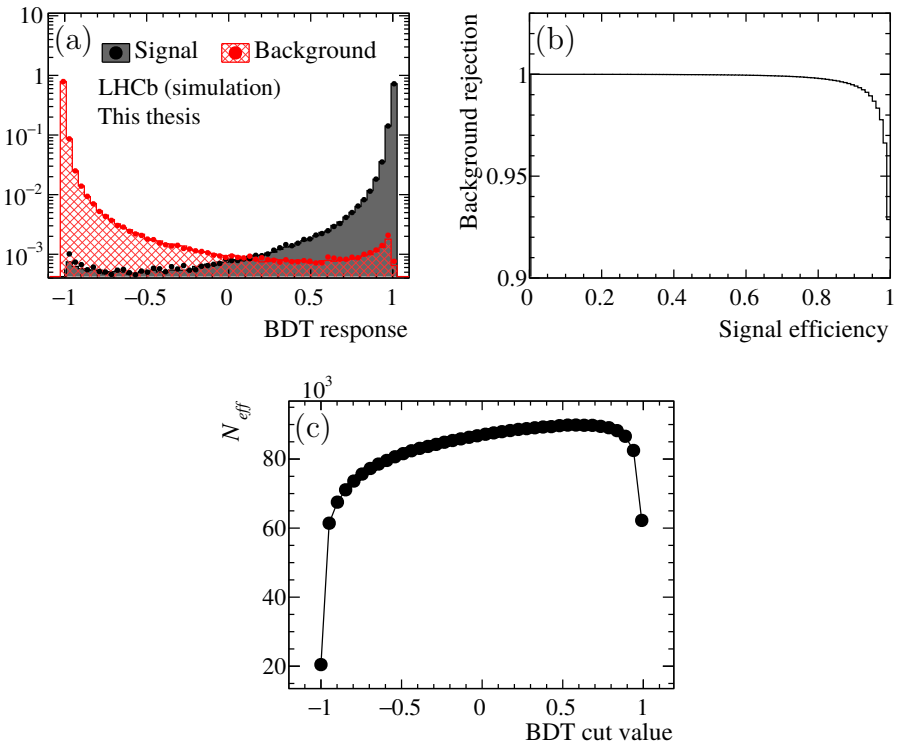
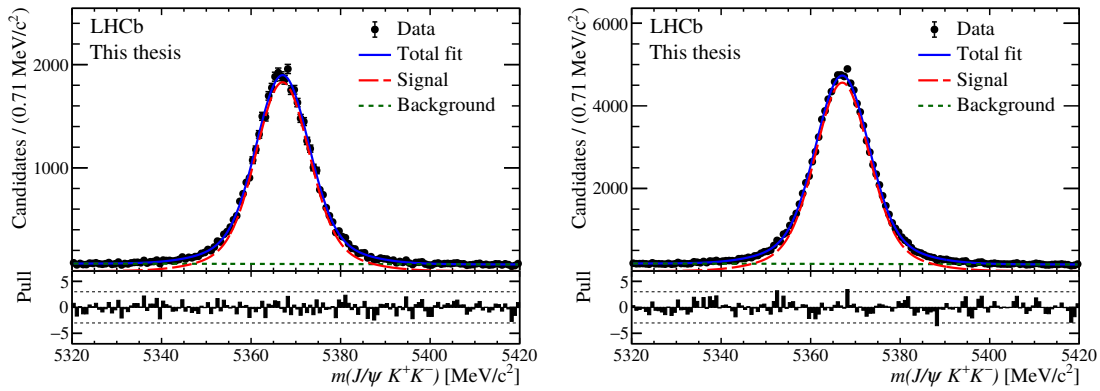


Figure 5.5: Performance plots of the Boosted Decision Tree (BDT) employed to separate $B_s^0 \to J/\psi\,\phi$ decays from background candidates. In (a), the BDT response is shown for the signal and background training (shaded area) and testing (points) samples. The background rejection rate as a function of the signal efficiency when placing a cut on the BDT response can be seen in (b). Finally, (c) shows the effective signal sample size as a function of the BDT cut value.

cause differences in the separation power of the BDT in simulated and in real data, but that this would not automatically cause any biases on the physics parameters extracted later. Remaining differences between data and simulation can therefore also be dealt with at a later stage when the decay-time and angular acceptances are determined.

The signal and background proxy samples are split into large training and smaller testing samples. After the training of the BDT, the testing samples are used to evaluate the separation power and to test for possible overtraining. Figure 5.5a shows the distribution of the BDT response for the training and testing samples of the signal and background proxy. No significant differences between the training and testing samples are observed. In Figure 5.5b, the background rejection rate and signal efficiency are shown when scanning the cut value on the BDT response. A high signal efficiency can be retained while at the same time most of the background



(a) 2015 data (b) 2016 data

Figure 5.6: Invariant mass distribution of $B_s^0 \to J/\psi\,\phi$ candidates after the BDT based selection. Projections of a fit to the data are overlaid.

is rejected. To choose the optimal cut on the BDT response, the effective signal yield N_{eff}, defined in Section 4.4, of the background-subtracted data sample is employed as figure of merit. A scan of the BDT cut value is performed, and at every step the background in the data sample is subtracted with the ${}_s\mathcal{P}lot$ technique. The obtained effective yields are shown in Figure 5.5c, and the optimal value was found to be 0.58. This corresponds to a signal efficiency of roughly 96% on the signal testing sample. In Figure 5.6, the mass distributions of the 2015 and 2016 data samples are shown after applying the cut on the BDT response.

Peaking backgrounds

Besides random combination of tracks, other B meson or hadron decays can contribute to the background if one of their final state particles is misidentified and thus gets assigned to a wrong mass hypothesis. The thereby wrongly reconstructed mass of the fake B_s^0 candidate can fall into the selected mass window of real B_s^0 candidates. Two decays are identified that can have a significant contribution. The first is the decay $B^0 \to J/\psi\,K^{*0}$, where the K^{*0} meson decays into a charged kaon and a charged pion. In case the pion is reconstructed as a kaon, the lower mass of the B^0 meson with respect to the B_s^0 meson is partially compensated. The resulting mass distribution of such wrongly reconstructed decays is shown in Figure 5.7a for simulated $B^0 \to J/\psi\,K^{*0}$ decays.

The second non-combinatorial background component arises from $\Lambda_b^0 \to J/\psi\,pK$ decays where the proton is misidentified as a kaon. In analogy to the decay

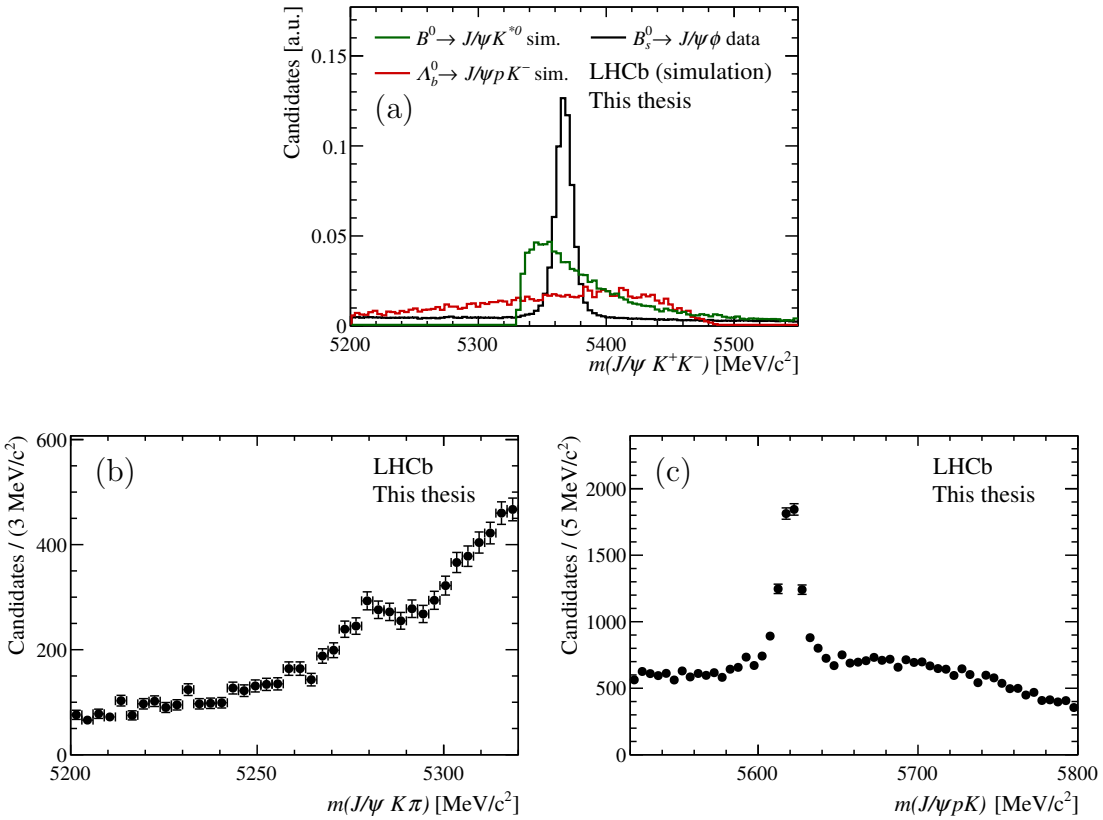


Figure 5.7: Mass distributions illustrating wrongly reconstructed $B^0 \to J/\psi K^{*0}$ and $\Lambda_b^0 \to J/\psi pK$ decays in the $B_s^0 \to J/\psi\phi$ data sample. In (a), the distributions of simulated $B^0 \to J/\psi K^{*0}$ and $\Lambda_b^0 \to J/\psi pK$ decays that are reconstructed as $B_s^0 \to J/\psi\phi$ are shown on top of the data distribution. In (b) and (c) the presence of these wrongly reconstructed backgrounds in the selected data sample is shown by determining the invariant mass under respective alternative mass hypotheses of the final state particles. Details are given in the text.

$B^0 \to J/\psi K^{*0}$, the wrongly assigned mass hypothesis leads to a reconstructed mass that lies in the selected B_s^0 mass region, see Figure 5.7a.

Figure 5.7b shows the selected $B_s^0 \to J/\psi\phi$ candidates in data when assigning the pion mass hypothesis to the kaon with the larger `ProbNN`π value. To suppress the dominant contribution of real $B_s^0 \to J/\psi\phi$ decays, only candidates with a reconstructed B_s^0 mass that is at least 20 MeV/c^2 larger than the nominal one are considered. The peaking structure at approximately 5280 MeV/c^2 corresponds to the expected misidentified B^0 mesons [3]. In analogy, Figure 5.7c shows the selected $B_s^0 \to J/\psi\phi$ candidates in data when assigning the proton mass hypothesis to the kaon with the larger `ProbNN`p value. This time, both B_s^0 mass sidebands are considered by selecting candidates with B_s^0 masses that are at least 20 MeV/c^2 away from

the nominal one. A clear peak at the expected Λ_b^0 meson mass of approximately $5620\,\mathrm{MeV/c^2}$ [3] is visible.

In the following, the strategy to suppress the $\Lambda_b^0 \to J/\psi\,pK$ background is presented. The $B^0 \to J/\psi\,K^{*0}$ component is treated in a similar way, but due to its relatively low yield it is less relevant. As a first step, a veto is applied to the selected $B_s^0 \to J/\psi\,\phi$ data to reject a large fraction of the misidentified Λ_b^0 decays. In principle, any candidate with a matching reconstructed Λ_b^0 mass could be rejected. However, this would remove a significant fraction of the $B_s^0 \to J/\psi\,\phi$ signal. Instead, such candidates are only rejected if the maximum `ProbNN`p value of the kaons is larger than 0.7. The reconstructed Λ_b^0 mass has to lie within $30\,\mathrm{MeV/c^2}$ around the nominal Λ_b^0 mass.

To estimate the remaining Λ_b^0 background, the candidates of Figure 5.7c are split according to the `ProbNN`p criteria of the veto. Figure 5.8 shows the two categories and projections of fits from which the yields of $\Lambda_b^0 \to J/\psi\,pK$ candidates are determined. The background component is modeled by a fourth order polynomial and the signal peak is described by an Ipatia function, see Ref. [76], whose parameters are fixed from simulation. From the extracted yields, the signal shape in combination with the mass window of the veto, and a simulation driven correction that accounts for the fact that only the B_s^0 mass sidebands are considered, the number of $\Lambda_b^0 \to J/\psi\,pK$ decays remaining after the veto can be estimated. For both years of data taking combined, this number is 1626 ± 100.

In the case of the $B^0 \to J/\psi\,K^{*0}$ background component, an analog veto is applied. The only difference is that the PID requirement is given by `probNN`$\pi > 0.7$ || `probNN`$K < 0.35$. The estimated number of remaining background candidates after the veto is determined to be approximately 200 ± 100. While this number is relatively small, the remaining Λ_b^0 decays can not be neglected. Instead, simulated $\Lambda_b^0 \to J/\psi\,pK$ decays are injected into the data sample with negative weights to compensate the estimated Λ_b^0 yield.

5.3 Subtraction of the remaining background

After removing the components of the background that arise from other B hadron decays, the remaining combinatorial background is statistically subtracted in means of the ${}_s\mathcal{P}lot$ technique, see Section 4.4. The mass of the B_s^0 candidates is chosen as discriminating variable.

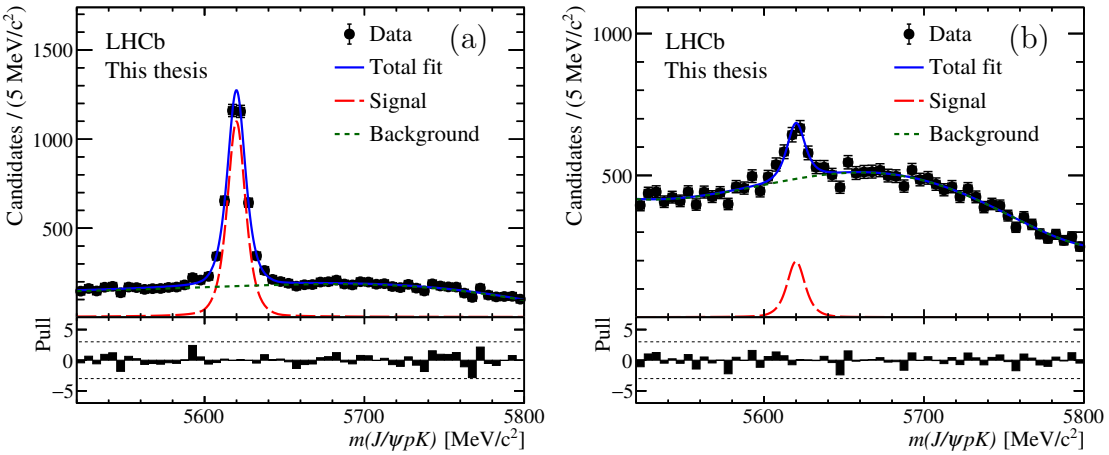


Figure 5.8: Invariant mass distribution of the selected $B_s^0 \to J/\psi\,\phi$ candidates when assigning the proton mass hypothesis to the kaon with the larger `ProbNN`p value. The sample is split according to this `ProbNN`p value being lager (a) or smaller (b) than 0.7.

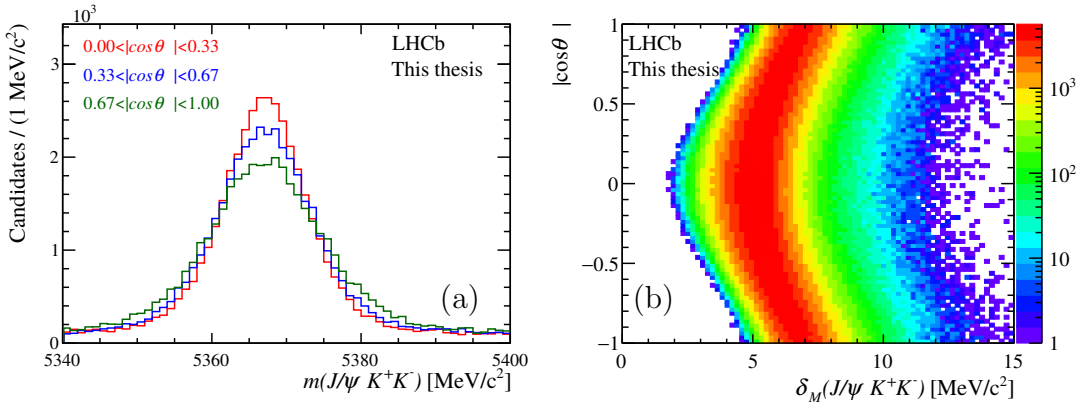


Figure 5.9: Dependence of the mass resolution on the helicity angle of the muons. The left plot shows the invariant mass distribution of the selected $B_s^0 \to J/\psi\,\phi$ data sample in bins of $|\cos\theta_\mu|$. The two dimensional plot at the right shows the correlation between this variable and the estimated mass uncertainty δ_M.

Per-event mass uncertainty

However, the mass of the B_s^0 candidate is correlated with the cosine of the helicity angle of the muons, $\cos\theta_\mu$, which is one of the variables that is needed to extract the physics parameters. This correlation is shown in Figure 5.9a, where the signal mass region in data is shown for different regions of the absolute value of $\cos\theta_\mu$. It

originates from a correlation of this variable with the momenta and opening angles of the muons in the rest frame of the detector and therefore also with the mass resolution. This is taken into account by using the Decay-Tree-Fitter estimated `per-event mass uncertainty` δ_M as a conditional variable when describing the signal mass shape. In Figure 5.9b, the two dimensional distribution of δ_M and $\cos\theta_\mu$ is shown. The same dependence as in Figure 5.9a is observed. This means that the correlation between the helicity angle of the muons and the mass resolution is correctly taken into account by the usage of δ_M.

The mass model

In the fit to the mass distribution that is needed for the ${}_s\mathcal{P}lot$ technique, the combinatorial background is modeled with a single exponential, while the signal is described by a double sided Crystal Ball function [77]. The latter consists of a Gaussian core and asymmetric power law tails. It is commonly used in high energy physics and can model possible radiative tails in the mass shape. Its definition is given by:

$$CB(x|\sigma,\mu,\alpha_1,n_1,\alpha_2,n_2) = N\begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, & \text{for } \alpha_1 < \frac{x-\mu}{\sigma} < \alpha_2 \\ A_1\cdot\left(B_1 - \frac{x-\mu}{\sigma}\right)^{-n^1}, & \text{for } \frac{x-\mu}{\sigma} < \alpha_1 \\ A_2\cdot\left(B_2 - \frac{x-\mu}{\sigma}\right)^{-n^2}, & \text{for } \frac{x-\mu}{\sigma} > \alpha_2 \end{cases}, \quad (5.2)$$

where N is the normalization constant, and A_i and B_i ensure the continuity of the function and its first derivative. The tails are parametrized by the parameters α_1, n_1, α_2, n_2, and μ and σ are the mean and the width of the Gaussian core. In the fit, σ is parametrized as a function of the estimated mass uncertainty as:

$$\sigma = s_1\delta_M + s_2\delta_M^2, \quad (5.3)$$

with the two scaling parameters s_1 and s_2, which are freely floating. This scaling of the estimated `per-event mass uncertainty` is necessary to correct for possible differences between the estimated and actual mass resolution.

In addition to the combinatorial background, a further background component is added to account for the suppressed decay $B^0 \to J/\psi\,\phi$, which has the same final state as the signal decay and lies in the lower mass sideband. Its contribution is accounted for by a Gaussian function with a mean value μ_{B^0} that is shifted with respect to the mean value of the signal peak by the mass difference between the B_s^0

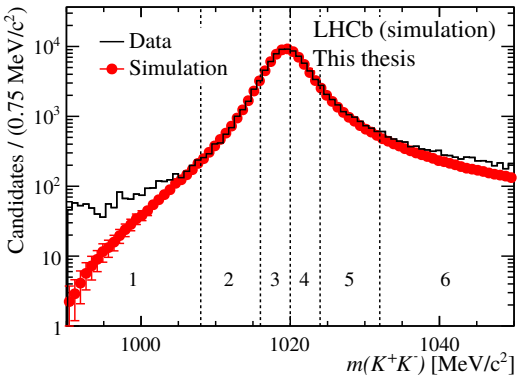


Figure 5.10: Distribution of the invariant two kaon mass in background-subtracted data (black) and simulation (red). The binning employed in the analysis is indicated by the dotted lines: [990, 1008, 1016, 1020, 1024, 1032, 1050] MeV/c^2.

and the B^0 meson [3]. The width σ_{B^0} is fixed to the width observed for the decay $B^0 \to J/\psi\, K^{*0}$, see Section 7.2.2.

The total PDF used in the extended maximum likelihood fit to the mass is then given by:

$$\begin{aligned}\mathrm{PDF}(m|\delta_M) =& f_{sig}\, CB(m|\sigma(s_1, s_2, \delta_M), \mu, \alpha_1, n_1, \alpha_2, n_2) \\ &+ (1 - f_{sig}) \left[f_{B^0}\, G(m|\mu_{B^0}, \sigma_{B^0}) + (1 - f_{B^0})\, N\, e^{-\gamma m} \right] . \end{aligned} \tag{5.4}$$

The parameter $\mu_{B_s^0}$ represents the peak position of the signal component, γ is the exponential constant of the combinatorial background component, N is the normalization of this background and f_{sig} and f_{B^0} parametrize the fractions of the three components.

Fit strategy

The final decay-time and angular fit is performed simultaneously for the two years of data taking, for the two trigger categories `Unbiased` and `Biased`, and in six bins of the invariant mass of the two kaons, see Section 8.1. The latter splitting is shown in Figure 5.10 for the data and the simulated sample. Especially in bin 1 and 6 the presence of the S-wave component in data can be seen, as it is not present in the simulated sample. Since separate PDFs are used in the 24 different data categories, also the weights for the background subtraction, calculated with the ${}_s\mathcal{P}lot$ technique, have to be obtained separately in each of them.

For every category, first a fit to the respective simulated sample is performed to fix the tail parameters α_1, n_1, α_2, n_2 of the Crystal Ball function. With these

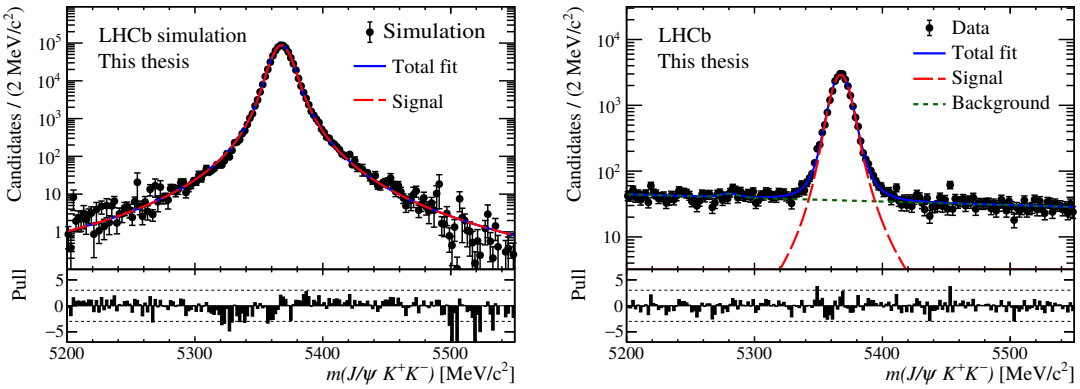


Figure 5.11: Fit projections of the invariant $J/\psi\ K^+K^-$ mass. On the left, the simulated sample is shown from which the tail parameters are obtained. On the right, the according fit projection for the data sample is shown. In both cases, only the 2016 `Unbiased` category in the fourth $m(K^+K^-)$ bin is shown.

fixed parameters, the fit to data is performed, in which the remaining parameters, including the scale parameters of `per-event mass uncertainty`, are floating. Figure 5.11 shows the fit projections of this two step fit procedure for the category with the highest signal yield, and Figure 5.12 shows combined projections of the fits for 2016 data sample split by trigger category. The pull distributions of these fits show some minor structures at the left tail of the signal peak, indicating that either the background or the signal distribution is not perfectly modeled. A range of systematic studies are presented in Section 10.1 that aim to account for this and possible bias due to the usage of the ${}_s\mathcal{P}lot$ technique. In Appendix C, the parameters of the mass fits in all data categories are given.

Figure 5.13 shows fit projections of the 2016 `Unbiased` category in three bins of $|\cos\theta_\mu|$. As expected, the varying width of the signal peak is well modeled by the usage of the `per-event mass uncertainty`.

Table 5.2 gives the estimated signal yields in each of the 24 categories. In total, approximately 117.9 k $B_s^0 \to J/\psi\,\phi$ decays are selected. Due to the weights applied by the ${}_s\mathcal{P}lot$ technique, this number reduces to an effective sample size, see Section 4.4, of $N_{eff} \approx 107.6$ k.

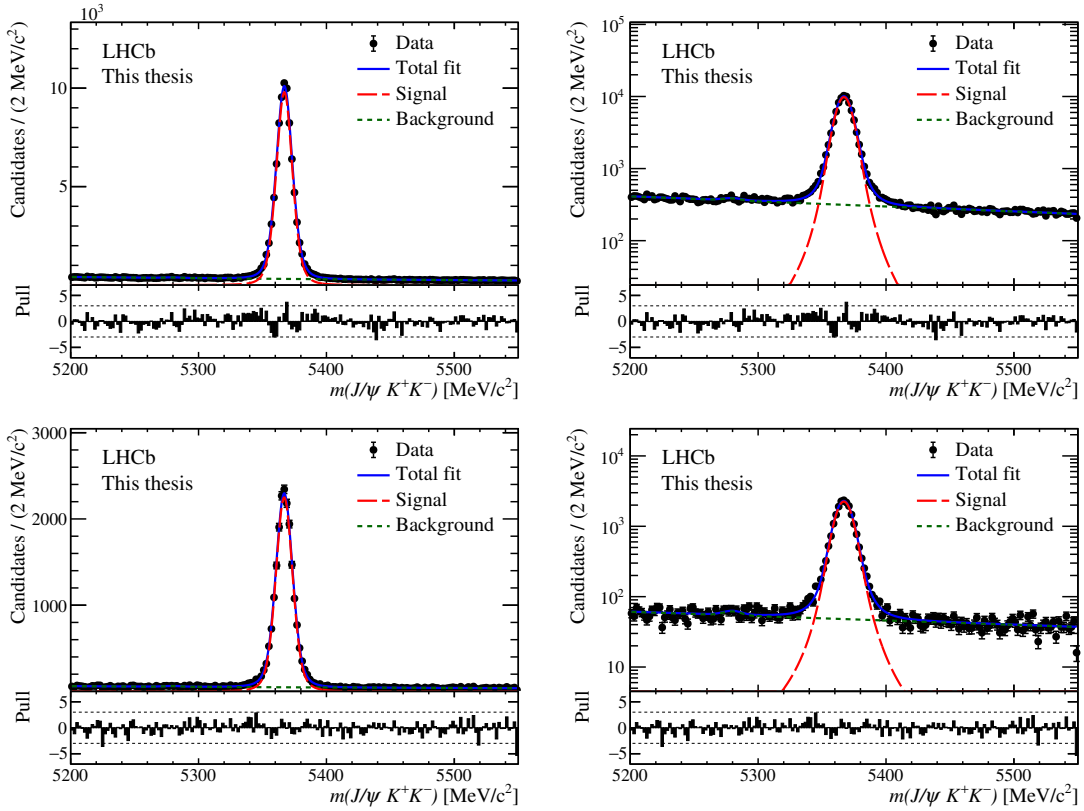
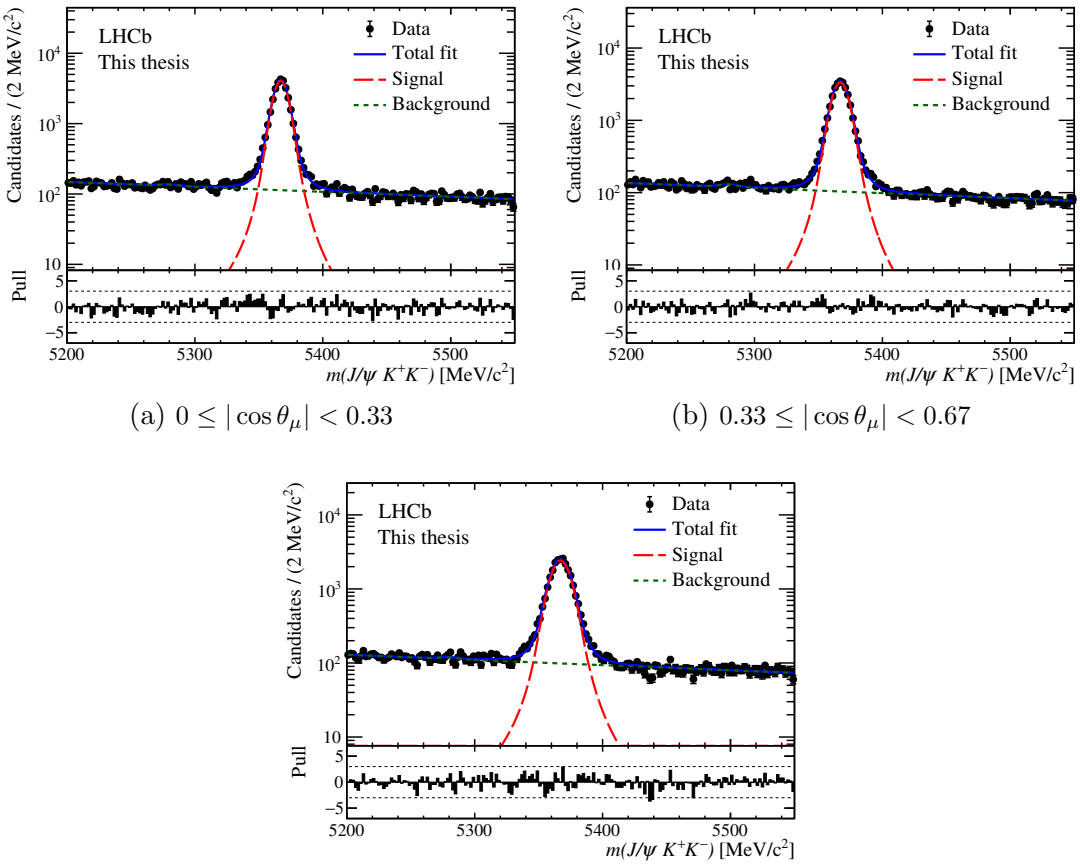


Figure 5.12: Combined fit projections of the two trigger categories in the 2016 data set. The top and bottom plots show the `Unbiased` and `Biased` trigger category, respectively. At the right, the same plots with a logarithmic y-axis are shown.

Table 5.2: Approximate signal yields of the $B_s^0 \to J/\psi\,\phi$ data sample split in the 24 different categories. They are obtained by the extended maximum likelihood fit that is performed in the context of the background subtraction with the ${}_s\mathcal{P}lot$ technique.

	2015		2016	
	`Unbiased`	`Biased`	`Unbiased`	`Biased`
m_{KK} bin 1	237	74	1803	429
m_{KK} bin 2	1270	404	8885	2193
m_{KK} bin 3	4238	1383	30607	7418
m_{KK} bin 4	3556	1071	24592	5888
m_{KK} bin 5	1570	461	10640	2581
m_{KK} bin 6	804	257	6119	1450



(a) $0 \leq |\cos\theta_\mu| < 0.33$

(b) $0.33 \leq |\cos\theta_\mu| < 0.67$

(c) $0.67 \leq |\cos\theta_\mu| \leq 1$

Figure 5.13: Distributions and fit projections of the invariant mass in the $B_s^0 \to J/\psi\,\phi$ data sample split in three bins of $|\cos\theta_\mu|$. The 2016 `Unbiased` category is shown.

6 Determination of the initial B_s^0 flavour

An essential component of measuring mixing-related CP violation in neutral-meson systems is the determination of the initial flavour of the meson, called tagging. In particular, the heart of the analysis presented in this thesis is the measurement of the difference between the decay rates of initial B_s^0 and $\overline{B}_s^0$ mesons. In this chapter, the tagging strategies developed by the LHCb collaboration to determine the initial flavour of neutral B_s mesons are presented. After a short introduction to the two available classes of tagging algorithms, their calibration for the channel $B_s^0 \to J/\psi\,\phi$ and the way their information is used in the final maximum likelihood fit is discussed. Since the tagging algorithms are centrally developed and optimized by the LHCb collaboration, this chapter is rather brief and focuses mainly on the determination of the calibration parameters specific for the channel $B_s^0 \to J/\psi\,\phi$ and the general usage of tagging information in a flavour-tagged maximum likelihood fit.

6.1 Flavour tagging at LHCb

At the LHCb experiment, the determination of the flavour of neutral B mesons is based on the reconstruction and classification of the underlying event in which the B meson is produced. The available tagging algorithms (taggers) can be divided into two classes. The opposite-side (OS) taggers rely on the fact that b quarks are predominantly produced in $b\bar{b}$ pairs and try to infer the initial flavour from the decay chain of the respective other b quark. In contrast, the same-side (SS) taggers exploit the charge of particles that are created in association with the fragmentation of the signal b quark. For a detailed description of the current versions of the tagging algorithms see Ref. [78–81]. Figure 6.1 illustrates the concepts of the taggers relevant for this analysis and divides them in the OS and SS categories.

Three of the OS algorithms try to identify kaons or leptons from either $b \to c \to s$ or semileptonic decays of the opposite-side b quark. The charge of the kaon or lepton determines then the flavour of the signal b quark. In addition, the other two OS taggers aim to reconstruct the charge of the opposite-side b decay vertex

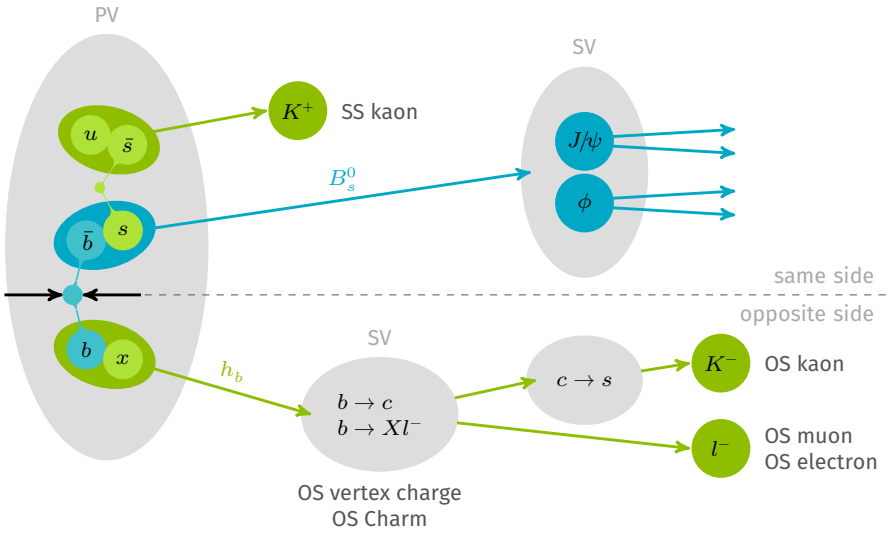


Figure 6.1: Schematic view of the tagging algorithms relevant in this analysis. The top half contains the signal B_s^0 decay and the same-side (SS) tagger. The bottom half shows the opposite-side (OS) taggers. Figure modified from Ref. [44].

or the flavour of the charm meson produced in a $b \to c$ decay. The information of all available OS taggers are centrally combined and transformed to a single tagging decision.

In contrast to the OS algorithms, the SS tagger is specific for the light quark of the signal B meson. In the case of the B_s^0 meson, the aim is to reconstruct the charged kaon that is often produced in association to the hadronization of the b quark with a s quark. In the case of a B^0 meson the kaon has to be replaced by a pion. Again, the charge of the kaon determines the flavour of the B_s^0 meson.

The algorithms of both tagging categories rely on a preselection of particles that aims to suppress contributions from the remaining underlying event, which spoils their performance. Kinematic and geometrical properties of these preselected particles and the respective signal B_s^0 decay are then used as input for multivariate algorithms that aim to determine the flavour of the B meson. These algorithms are trained either on real B^+ meson decays in the case of the OS taggers or on simulated samples in the case of the SS kaon tagger. Finally, the output of these algorithms is transformed into a tagging decision with values 1 (B_s^0), 0 (no tagged), -1($\overline{B}_s^0$) and a mistag probability that is a per-event estimate of the chance of a wrongly tagged B-meson.

6.2 Calibration of the estimated mistag probability

In order to correctly take into account the possibility of wrongly tagged candidates, the previously mentioned estimated mistag probabilities need to be calibrated to be applicable for the selected $B_s^0 \to J/\psi\,\phi$ data sample. Given the estimated mistag probability η for one of the tagging algorithms, the respective calibrated version of it, $\omega(\eta)$, is parametrized by:

$$\begin{aligned}\omega(\eta) &= \left(p_0 + \frac{\Delta p_0}{2}\right) + \left(p_1 + \frac{\Delta p_1}{2}\right)(\eta - \langle\eta\rangle) \quad \text{for initial } B_s^0,\\ \bar{\omega}(\eta) &= \left(p_0 - \frac{\Delta p_0}{2}\right) + \left(p_1 - \frac{\Delta p_1}{2}\right)(\eta - \langle\eta\rangle) \quad \text{for initial } \overline{B}_s^0.\end{aligned} \tag{6.1}$$

Here, $\langle\eta\rangle$ is the average estimated mistag probability of the sample used in the calibration, and p_0 and p_1 are the main calibration parameters. The additional calibration parameters Δp_0 and Δp_1 allow for different calibrated mistag probabilities w and $\bar{\omega}$ for initial B_s^0 and $\overline{B}_s^0$ mesons, respectively. Separate calibration procedures and parameters are applied and obtained for the OS and SS taggers. They rely on standard techniques and decay channels that are commonly used in the LHCb collaboration, but they were performed with a specific selection and weighting of the calibration channels in the scope of the analysis presented here. This ensures the portability of the calibration parameters from the calibration channels to the selected set of $B_s^0 \to J/\psi\,\phi$ candidates.

Opposite-side tagging algorithm

The combined opposite-side (OS) tagger is calibrated using the decay channel $B^\pm \to J/\psi\,K^\pm$ for which the charge of the kaon determines the initial flavour of the $B^\pm$ meson. The rate of wrongly tagged candidates is directly accessible and can be compared to the estimated mistag probability. After a similar selection and background subtraction as for the $B_s^0 \to J/\psi\,\phi$ data sample, an unbinned maximum likelihood fit is performed to determine the calibration parameters introduced in Equation (6.1). Depending on the initial $B^\pm$ flavour, the PDF is given by:

$$\begin{aligned}\mathrm{PDF}(a|\eta) &= (1-a)\omega(\eta) + a(1-\omega(\eta)) \quad \text{for initial } B^+,\\ \mathrm{PDF}(a|\eta) &= (1-a)\bar{\omega}(\eta) + a(1-\bar{\omega}(\eta)) \quad \text{for initial } B^-,\end{aligned} \tag{6.2}$$

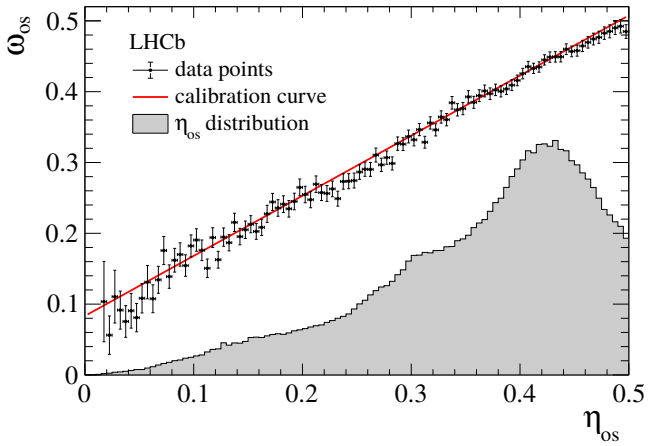


Figure 6.2: Calibration function for the estimated mistag probability η_{os} of the combined OS tagger. The red line shows the fitted calibration function, while the true mistag probability obtained from $B^+ \to J/\psi K^+$ decays in bins of η_{os} is represented by the black points. In addition, the distribution of η_{os} in the background-subtracted $B_s^0 \to J/\psi\,\phi$ data sample is shown in gray. The figure is taken from Ref. [63].

where a is 1 and 0 for correctly and wrongly tagged candidates, respectively. To improve the portability of the calibration parameters from the $B^+ \to J/\psi K^+$ to the $B_s^0 \to J/\psi\,\phi$ data sample, the former is weighted in kinematic and event multiplicity variables, like the number of reconstructed tracks, to match the respective distributions in the $B_s^0 \to J/\psi\,\phi$ data sample.

Figure 6.2 shows the resulting calibration curve for the combined set of B^+ and B^- candidates. In addition, in the same figure, the distribution of the estimated OS mistag probability in the $B_s^0 \to J/\psi\,\phi$ data sample is shown. The portability from the calibration to the signal decay is studied using respective simulated samples. For each of them, the calibration is repeated using the truth information and the difference of the obtained calibration parameters is assigned as a systematic uncertainty. Table 6.1 shows the final values and uncertainties for the calibration parameters.

Same-side tagging algorithm

The calibration of the same-side (SS) kaon tagger is less straight forward than the one of the OS tagger since the calibration channel has to be a B_s^0 decay and is therefore also subject to $B_s^0 - \overline{B}_s^0$ mixing. However, using the flavour specific calibration channel $B_s^0 \to D_s^- \pi^+$ allows at least to determine the flavour at decay. Following the formulas presented in Section 1.4.1, and considering decay-time acceptance and

Table 6.1: Calibration parameters for the OS and SS taggers. Where given, the first uncertainty is statistical and the second is systematic. Numbers are taken from Ref. [63].

Tagger	OS	SS
p_0	$0.3890 \pm 0.0007 \pm 0.0028$	$0.4325 \pm 0.0108 \pm 0.0030$
p_1	$0.8486 \pm 0.0062 \pm 0.0265$	$0.9241 \pm 0.1314 \pm 0.0196$
Δp_0	0.0090 ± 0.0014	0.00 ± 0.03
Δp_1	0.0143 ± 0.0124	0.00 ± 0.03
$\langle\eta\rangle$	0.3602	0.4167

resolution effects, see Chapter 7, the expected decay-time distribution of the decay $B_s^0 \to D_s^- \pi^+$ is given by:

$$\mathrm{PDF}(t|q_{mix}) \propto \varepsilon(t)\,[\Gamma(t|q_{mix}) \otimes G(t)]\,, \tag{6.3}$$

with

$$\Gamma(t|q_{mix}) = e^{-\Gamma_s t}\,[\cosh(\Delta\Gamma_s t/2) + q_{mix}(1-2\omega(\eta))\cos(\Delta m_s s t)]\,. \tag{6.4}$$

The decay-time acceptance and resolution are represented by $\varepsilon(t)$ and $G(t)$ and q_{mix} is either 1 or -1 depending on whether the B_s^0 meson has or has not changed its flavour. Since the mistag probability $\omega(\eta)$ enters here, the calibration parameters of the SS kaon tagger can be determined from a maximum likelihood fit to the decay-time distribution of mixed and not-mixed B_s^0 mesons in this decay channel. While the decay-time resolution is obtained from a prompt sample of $D_s^- \pi^+$ candidates in analogy to the procedure described in Section 7.1, the decay-time acceptance is modeled with the empirical function $\varepsilon(t) = 1 - 1/(1 + (at)^n + b)$ and is determined directly in the fit. The values for the B_s^0 meson parameters Γ_s, $\Delta\Gamma_s$ and Δm_s are constrained to the current world averages [3].

In analogy to the OS tagging calibration, the selected and background-subtracted $B_s^0 \to D_s^- \pi^+$ data is weighted to match the $B_s^0 \to J/\psi\,\phi$ data sample before performing the fit to the decay-time distribution. The projections of this fit are shown in Figure 6.3a. Clear oscillations are visible. Their amplitude is proportional to the dilution factor $(1 - 2\omega(\eta))$ and allows therefore the calibration of the mistag probability. Figure 6.3b shows the obtained calibration curve together with the distribution of the estimated SS mistag probability in the $B_s^0 \to J/\psi\,\phi$ data sample. Since the true initial flavour is not known, the tagging calibration parameters Δp_0 and Δp_1 are fixed to 0. A systematic uncertainty of 0.03 is assigned based

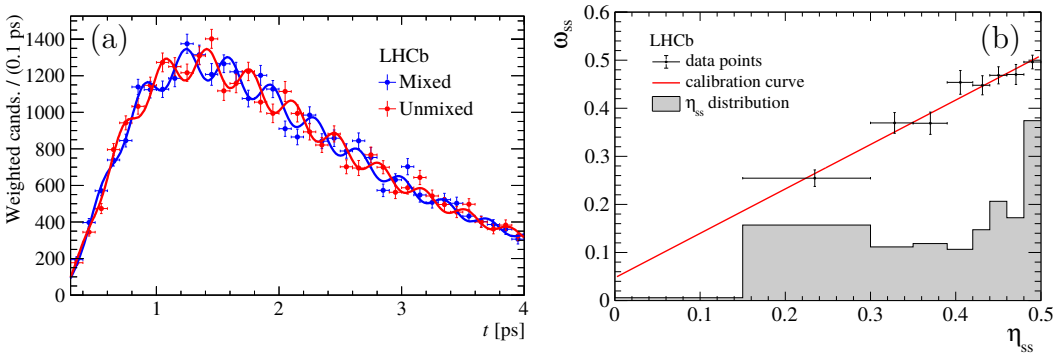


Figure 6.3: Background subtracted decay-time distribution of $B_s^0 \to D_s^- \pi^+$ decays split according to the agreement (unmixed) or disagreement (mixed) of the assigned B_s^0 flavour at decay and production (a). Fit projections of the respective categories are overlaid. Calibration function for the estimated mistag probability η_{ss} of the SS tagger (b). The red line shows the fitted calibration function, while the true mistag probability obtained from a fit to $B_s^0 \to D_s^- \pi^+$ decays in bins of η_{ss} is represented by the black points. In addition, the distribution of η_{ss} in the background-subtracted $B_s^0 \to J/\psi\,\phi$ data sample is shown in gray. The figures are taken from Ref. [63].

on the study presented in [81]. Furthermore, the same procedure as for the OS tagging calibration is performed to estimate the systematic uncertainty due to the portability from calibration to signal sample. Table 6.1 shows the final values and uncertainties for the calibration parameters.

6.3 Performance and combination

When studying the performance of a tagging algorithm, it is useful to understand the effect of wrongly tagged candidates on the measured CP asymmetry. Assuming no direct CP violation, the difference between the decay rates of initial B_s^0 and $\overline{B}_s^0$ mesons, see Equations (1.37) and (1.38), is proportional to $\sin\phi_s \sin(\Delta m_s t)$. Therefore, also the CP asymmetry A_{CP} is to first order given by:

$$A_{CP}(t) \coloneqq \sin\phi_s \sin(\Delta m_s t). \tag{6.5}$$

In the presence of a mistag probability ω, decays can not longer be uniquely assigned to either Equation (1.37) or Equation (1.38), but obtain also contributions from

the respective opposite initial flavour. Thus, the observed CP asymmetry is then given by:

$$A_{CP}(t) \coloneqq D \sin\phi_s \sin(\Delta m_s t), \tag{6.6}$$

where $D = (1 - 2\omega)$ is the dilution factor. The precision on a diluted asymmetry, and therefore in this case on $\sin\phi_s$, is directly proportional to this dilution factor. Alternatively, one can state that the effective size of a tagged sample is reduced by the factor D^2. Since not all candidates get a tag hypothesis assigned, the tagging efficiency ε has to be considered additionally and further reduces the effective sample size. The combined quantity εD^2 is called the effective tagging power and is a measure of the tagging performance.

For a weighted sample with partially tagged events and specific mistag probabilities ω_i for each candidate i, the effective tagging power is then given by:

$$\varepsilon D^2 = \frac{\sum_{\text{tagged}} v_i (1 - 2\omega_i)^2}{\sum_{\text{all}} v_i}, \tag{6.7}$$

where v_i is the weight of the respective candidate. These weights are for example given by the background subtraction with the ${}_s\mathcal{P}lot$ technique, see Section 4.4, and are labeled here as v_i instead of w_i to avoid a confusion with the mistag probability. Equation (6.7) allows to calculate the effective tagging power for candidates that are either exclusively tagged by the OS or the SS tagger. For candidates where both taggers give a decision, their output has to be combined. Given the two tagging decisions of the two tagger categories q^{ss} and q^{os} and the respective uncorrelated mistag probabilities $\overset{\frown}{\omega}{}^{ss}$ and $\overset{\frown}{\omega}{}^{os}$, the probabilities $p(B_s^0)/p(\overline{B}_s^0)$ for an initial $B_s^0/\overline{B}_s^0$ meson are given by:

$$p(B_s^0) = \frac{1}{N}(1 + q^{os}(1 - 2\omega^{os}))(1 + q^{ss}(1 - 2\omega^{ss})), \tag{6.8}$$

$$p(\overline{B}_s^0) = \frac{1}{N}(1 - q^{os}(1 - 2\bar{\omega}^{os}))(1 - q^{ss}(1 - 2\bar{\omega}^{ss})), \tag{6.9}$$

where $N = p(B_s^0) + p(\overline{B}_s^0)$ is a normalization factor. A detailed derivation of these equations can be found in Appendix D. The dilution factor $D = (1 - 2\omega)$ can be expressed in terms of these probabilities as $D = |p(B_s^0) - p(\overline{B}_s^0)|$, which, together with Equation (6.7), allows to calculate the effective tagging power of the combined OS and SS tagged events. The respective numbers for the background-subtracted $B_s^0 \to J/\psi\,\phi$ data sample can be found in Table 6.2. In total, a tagging efficiency of roughly 78% with an average dilution factor of 0.061 is reached. This corresponds

Table 6.2: Tagging performance numbers. The tagging efficiency of the respective category is given with respect to the total number of candidates. Thus, the effective tagging powers and efficiencies can be added linearly. The total dilution factor is then obtained by dividing the total effective tagging power by the total tagging efficiency.

Category	$\varepsilon_{\text{tag}}(\%)$	$\mathcal{D}^2$	$\varepsilon_{\text{tag}}\mathcal{D}^2(\%)$
OS-only	11.349	0.078	0.88 ± 0.04
SSK-only	42.574	0.032	1.38 ± 0.30
OS&SSK	23.837	0.104	2.47 ± 0.15
Total	77.760	0.061	4.73 ± 0.34

to a total effective tagging power of $(4.73 \pm 0.34)\%$ where the uncertainty originates from the total uncertainties on the tagging calibration parameters.

6.4 Embedding in the probability density function

For a given combination of tagging decisions q^{os} and q^{ss} and estimated mistag probabilities η^{os} and η^{ss} of the opposite-side (OS) and same-side (SS) taggers, Equations (6.8) and (6.9) give the probability for an initial B_s^0 and $\overline{B}_s^0$ meson, respectively. Thus, they can be directly plugged into Equation (3.1) to obtain the PDF for a sample with tagging information:

$$\begin{aligned}\text{PDF}(t, \Omega | q^{os}, q^{ss}, \eta^{os}, \eta^{ss}) = \qquad (6.10)\\ \frac{1}{N^{\eta^{os},\eta^{ss}}_{q^{os},q^{ss}}} \sum_{k=1}^{10} A_k \, f_k(\Omega) \\ [(1 + q^{os}(1 - 2\omega^{os}(\eta_{os})))(1 + q^{ss}(1 - 2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t) \\ +(1 - q^{os}(1 - 2\bar{\omega}^{os}(\eta_{os})))(1 - q^{ss}(1 - 2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)].\end{aligned}$$

The calibration relations of the mistag probabilities $\overset{(-)}{\omega}{}^{os/ss}$ are part of the PDF and their parameters have Gaussian constraints, see Chapter 4, to the values presented in this section. The normalization is given by:

$$
\begin{aligned}
N^{\eta^{os},\eta^{ss}}_{q^{os},q^{ss}} \quad = \quad & \\
\int_{t=0.3\,\mathrm{ps}}^{15\,\mathrm{ps}} & \int_{\Omega} \sum_{k=1}^{10} A_k \, f_k(\Omega)\, \mathrm{d}\Omega \\
& [(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t) \\
& +(1-q^{os}(1-2\bar{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)]\, \mathrm{d}t.
\end{aligned}
\tag{6.11}
$$

7 Acceptance and resolution effects

The extraction of the underlying physics parameters is based on a fit to the decay-time and angular distributions in the decay $B_s^0 \to J/\psi\,\phi$. In order to obtain unbiased results, any reconstruction and selection step that changes these distributions has to be taken into account. One can distinguish two categories. The first contains resolution effects that are caused by the finite precision of momentum and position measurements of the particles in the detector. They lead to reconstructed angles and decay times that differ slightly from the underlying true quantities. Reconstruction or selection efficiencies that depend on the decay time or the helicity angles of the respective $B_s^0 \to J/\psi\,\phi$ decay constitute the other category. Such dependencies are called acceptances and alter the shape of the observed distributions.

While the correct determination of the resolution of the measured decay time is crucial to correctly describe the fast $B_s^0 - \overline{B}_s^0$ oscillations caused by the B_s^0 meson mixing, the resolution of the helicity angles is significantly higher than the angular structures that have to be resolved [5]. Therefore, the angular resolution can be neglected.

In the following, the decay-time resolution and the decay-time and angular acceptances are examined. Each section is divided into a short discussion of the origin, a description of the chosen method for the determination and the embedding in the PDF of the respective resolution or acceptance component.

7.1 Decay-time resolution

In order to be able to get the best sensitivity to the CP-violating parameters ϕ_s and λ, it is necessary to resolve the oscillations in the time-dependent decay rate, see Equations (1.55) and (1.56). The relevant frequency is given by the mass splitting $\Delta m_s \approx 17.8\,\mathrm{ps}^{-1}$. This requires a measurement of the decay time of the B_s^0 candidates with a precision that is significantly higher than the according oscillation period. The part of the time-dependent decay rate that gives the highest sensitivity

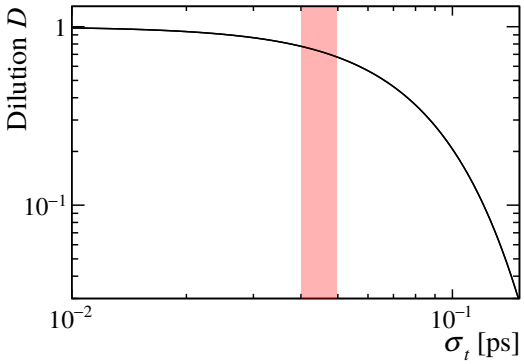


Figure 7.1: Dilution of the CP asymmetry in $B_s^0 \to J/\psi\,\phi$ decays as a function of the decay-time uncertainty σ_t. The area in light red shows the typical range covered by the LHCb experiment.

to ϕ_s is the asymmetry between the rates of B_s^0 and $\overline{B}_s^0$ decays, A_{CP}, which is roughly given by:

$$A_{CP}(t) \coloneqq \sin\phi_s \sin(\Delta m_s t). \tag{7.1}$$

Assuming a decay-time resolution function with a Gaussian shape, the measurable asymmetry $A_{CP}^{meas}(t)$ becomes:

$$A_{CP}^{meas}(t) = \int_{-\infty}^{\infty} \mathrm{d}t'\ \sin\phi_s\ \sin(\Delta m_s t')\ \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t')}{2\sigma_t^2}} \tag{7.2}$$

$$= e^{-\sigma_t^2 \frac{\Delta m_s}{2}}\ \sin\phi_s\ \sin(\Delta m_s t), \tag{7.3}$$

where the width of the resolution function, σ_t, is called the decay-time uncertainty. The dilution D is therefore defined as:

$$D = e^{-\sigma_t^2 \frac{\Delta m_s}{2}} \tag{7.4}$$

and relates the measurable and the underlying CP asymmetry. In analogy to the dilution due to the flavour tagging, see Chapter 6, the precision on ϕ_s is directly proportional to this factor. Figure 7.1 shows the dilution as a function of the decay-time uncertainty σ_t. In the LHCb experiment, typical values for B meson decays lie between 40 and 50 fs, which corresponds to a dilution of about 0.7-0.8. Although, the relative gain in precision by a relative improvement in the decay-time resolution is much higher for larger decay-time uncertainties, there is still a lot to gain in this region. Therefore, it is important to achieve the best resolution possible

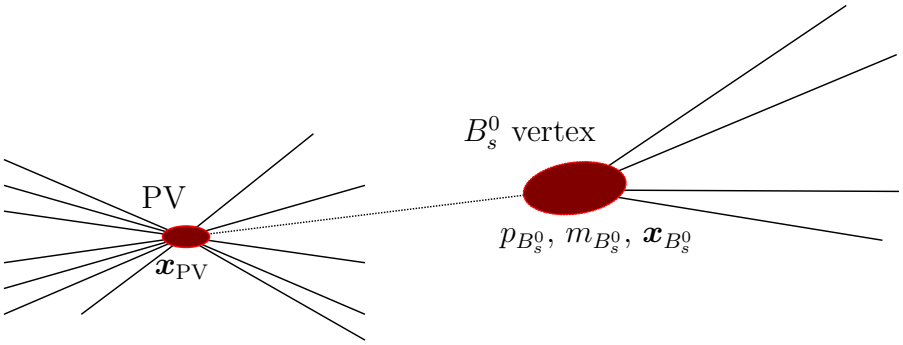


Figure 7.2: Topological illustration of the properties relevant for the decay-time resolution.

by using all available information for every single B_s^0 candidate. This leads to the usage of a decay-time uncertainty that is calculated for every candidate individually, called `per-event decay-time uncertainty`.

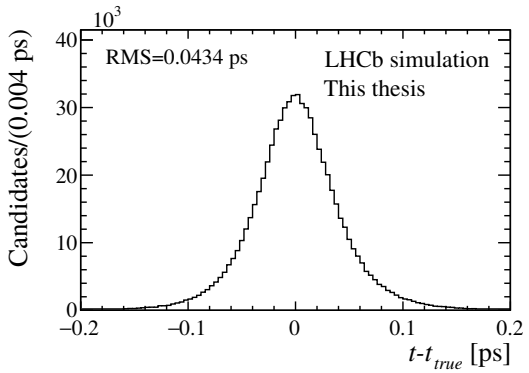
This section will continue with a brief discussion of the origin of the finite decay-time resolution of the LHCb detector, and will then focus on the calculation and calibration of the `per-event decay-time uncertainty`.

7.1.1 Origin of the decay-time resolution

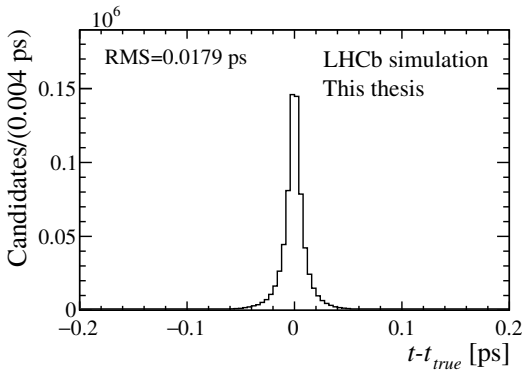
As presented in Chapter 5, the decay time of a $B_s^0 \to J/\psi\,\phi$ candidate is basically determined by the positions of the PV and the B_s^0 decay vertex and the measured momentum and mass of the B_s^0 candidate:

$$t = m_{B_s^0} \frac{|\boldsymbol{x}_{B_s^0} - \boldsymbol{x}_{PV}|}{p_{B_s^0}}. \tag{7.5}$$

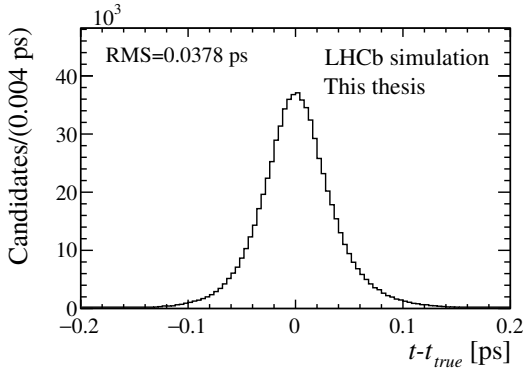
Figure 7.2 illustrates the relevant topology for the decay-time resolution. Due to the limited spacial and momentum resolution of the LHCb detector, the quantities that enter the decay-time calculation come with uncertainties that lead to an uncertainty on the derived decay time. The uncertainty on the measured mass of the B_s^0 candidates is typically negligible or does even cancel part of the uncertainty on the measured momentum. Figure 7.3a shows the decay-time resolution obtained from the simulated $B_s^0 \to J/\psi\,\phi$ sample. The reconstructed decay time in this figure is determined directly from the vertex separation and the measured momentum, without employing the Decay Tree Fitter (DTF). This allows to disentangle the different contributions to the total decay-time uncertainty by using the true information for some of the components. Figures 7.3b to 7.3d show the respective decay-time resolution when using the true information for all components except



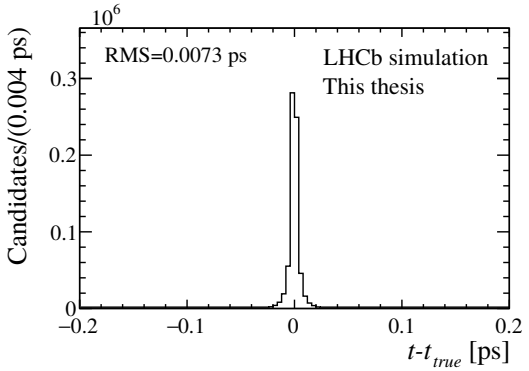
(a) Only reconstructed quantities



(b) Reconstructed PV, true B_s^0 momentum and B_s^0 decay vertex



(c) Reconstructed B_s^0 decay vertex, true B_s^0 momentum and PV

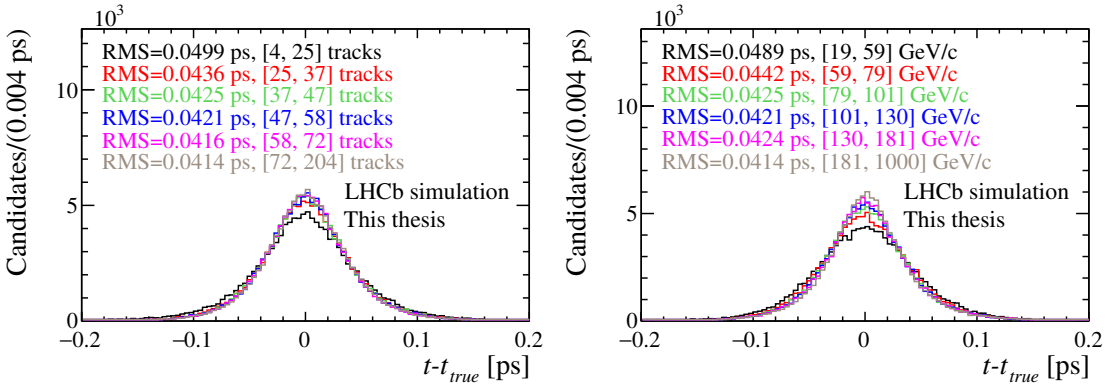


(d) Reconstructed momentum, true B_s^0 vertex and PV

Figure 7.3: Decay-time resolutions obtained from a simulated sample of $B_s^0 \to J/\psi\,\phi$ decays. The reconstructed decay time is calculated using Equation (7.5). The different figures show the resolution when using only reconstructed or reconstructed and true values for the three inputs. The B_s^0 mass is fixed to the nominal value [3].

for one. The dominant contribution to the decay-time uncertainty is the resolution of the B_s^0 decay vertex. A PV is typically reconstructed by much more tracks and has therefore a significantly better resolution.

An important feature of this analysis is the previously mentioned `per-event decay-time uncertainty`. This means that the precision with which the decay time can be measured is not the same for every $B_s^0 \to J/\psi\,\phi$ candidate. There are many different dependencies contributing to the total decay-time uncertainty. Two of them are shown in Figure 7.4. For a given decay time, B_s^0 mesons with a low momentum have a smaller flight distance than the ones with high momentum. Therefore, the same uncertainties on the vertices yields a higher relative uncertainty on the decay time. This can be seen in Figure 7.4a. The effect is not very pronounced



(a) In bins of $\#PV$-tracks

(b) In bins of $p^{B_s^0}$

Figure 7.4: Decay-time resolutions obtained from a simulated sample of $B_s^0 \to J/\psi\,\phi$ decays. The resolution is shown for different equally populated bins of: (a) the number of tracks associated to the PV (a) and (b) the B_s^0 momentum.

since daughters of B_s^0 mesons with higher momentum tend to have a smaller opening angles, which leads to a less precisely reconstructed B_s^0 decay vertex. This partially compensates the dependence on the B_s^0 momentum. Another effect can be seen in Figure 7.4b. Although the impact of the PV resolution is small, there is a small dependence on the number of tracks associated to the PV. A PV reconstructed from a low number of tracks causes a higher decay-time uncertainty.

During the Decay Tree Fit, see Chapter 5, the uncertainties on the reconstructed vertices, and thereby also the previously discussed dependencies, are taken into account, and an estimate of the decay-time uncertainty for the respective candidate is calculated. Figure 7.5 shows the distribution of this estimated decay-time uncertainty δ_t in the simulated $B_s^0 \to J/\psi\,\phi$ sample. In addition, the observed decay-time resolution histogram is shown for different ranges of the estimated decay-time uncertainty. At this point, these graphs are only presented for illustration purposes. The calibration of the `per-event decay-time uncertainty` will follow in one of the next sections.

7.1.2 Strategy to determine the resolution

The estimated decay-time uncertainty allows to optimally exploit the statistical potential of the selected data sample. However, to obtain unbiased results, this estimate of the DTF has to be calibrated. Instead of relying on the simulated sample, a data driven method is employed. This method is based on a sample of prompt fake $B_s^0 \to J/\psi\,\phi$ candidates. Such a sample is dominated by prompt J/ψ

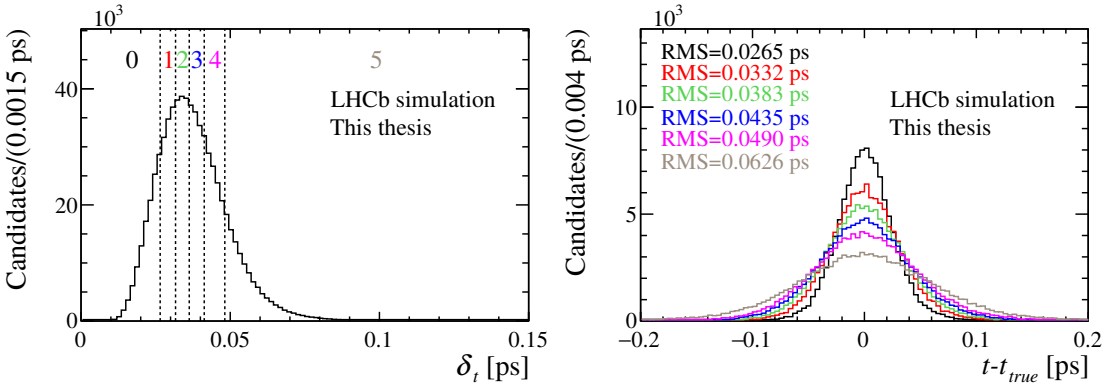


Figure 7.5: Distribution of the estimated decay-time uncertainty δ_t obtained by the DTF for the simulated $B_s^0 \to J/\psi\,\phi$ sample (a). Observed decay-time resolution in bins of δ_t (b). The δ_t bins are indicated by the dotted lines in (a).

mesons that are combined with two random kaons from the PV. By definition, these candidates have a true decay time of 0. Thus, their distribution can be used to determine the decay-time resolution. Splitting the prompt sample in bins of δ_t, allows then to do the desired calibration between the estimated and observed decay-time uncertainty.

The prompt data sample

The sample of prompt fake $B_s^0 \to J/\psi\,\phi$ candidates is triggered and selected in the same way as the sample of signal $B_s^0 \to J/\psi\,\phi$ candidates except that the requirements related to the separation of the PV and B_s^0 decay vertex are omitted. This means that the `Biased` HLT1 trigger configuration is not considered and that the HLT2 trigger line is changed to a version that has no requirement on the flight distance. Since a trigger configuration like this would result in an extremely high rate of events passing the requirements, the HLT2 line is prescaled by a factor of 0.2. This means that only one of five events passing the trigger line is written to disk. During the offline selection, the same cuts and multivariate classifier are applied as for the signal sample, except for the cut on the decay time, which is omitted.

The decay-time distribution of the selected prompt candidates is shown in Figure 7.6. Besides the dominant peak at $t = 0\,\text{ps}$, a tail towards high decay times is visible that originates from decay products of long-lived particles, e.g. from real $B_s^0 \to J/\psi\,\phi$ decays. When determining the resolution, this component has to be taken into account.

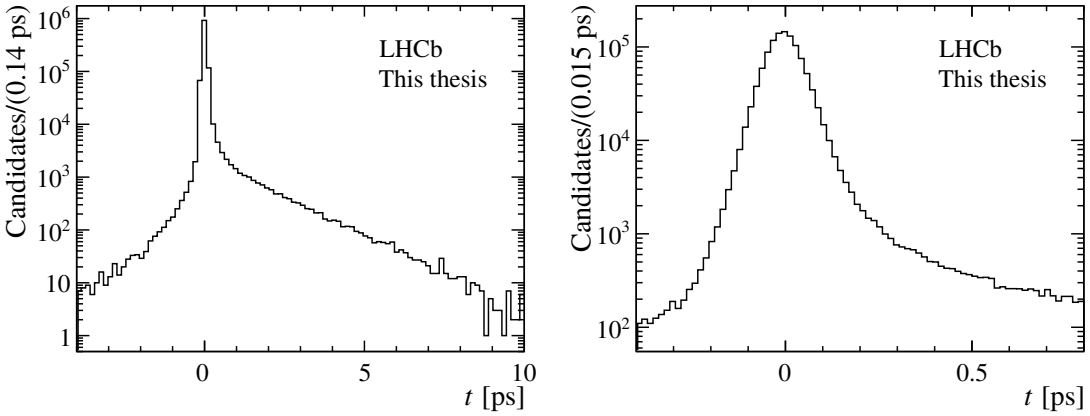


Figure 7.6: Decay-time distribution of the prompt fake $B_s^0 \to J/\psi\,\phi$ candidates. The right plot shows a zoomed view of the region around the peak at 0 ps.

Extracting the resolution from the prompt data sample

The resolution is determined by a maximum likelihood fit to the observed decay-time distribution of the prompt sample. The PDF for this fit consists of the component describing the prompt candidates, $P(t)$, and another component representing the candidates from the decay of long-living particles, $L(t)$. Both are folded with the resolution function $R(t)$. In addition, the *wrong PV* component, $W(t)$, is added to account for prompt $B_s^0 \to J/\psi\,\phi$ candidates that are associated to a wrong PV and have therefore a broad symmetric distribution in the reconstructed decay time. In total, the PDF is then given by:

$$\mathrm{PDF}(t) = (1 - f_{\mathrm{wpv}})\,[f_{\mathrm{prompt}}\,P(t) + (1 - f_{\mathrm{prompt}})\,L(t)] \otimes R(t) + f_{\mathrm{wpv}}\,W(t), \quad (7.6)$$

where f_{prompt} and f_{wpv} define the fractions of the three categories. The *wrong PV* component is modeled by a double-sided double-exponential function whose parameters are determined in a dedicated study that is summarized in Appendix E. The prompt component is given by a delta function and $L(t)$ consists of two exponential functions to account for potentially different lifetimes of the contributing long-living particles:

$$P(t) = \delta(t) \quad (7.7)$$

$$L(t) = \frac{1}{N}\left[f_s\,e^{\left(-\frac{t}{\tau_s}\right)} + (1 - f_s)\,e^{\left(-\frac{t}{\tau_l}\right)}\right]. \quad (7.8)$$

N is a normalization factor and the fractions of the two long-lived components with lifetimes τ_s and τ_l are parametrized via f_s and $(1 - f_s)$.

The resolution function is defined as the sum of three Gaussian distributions:

$$R(t) = \sum_{i=1}^{3} f_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_i}\right)^2}, \tag{7.9}$$

with the common mean μ, the widths σ_i and the relative fractions f_i that sum up to 1. Following previous studies [5], the two components with the smaller widths $(\sigma_1 < \sigma_2 < \sigma_3)$ are parametrized as a function of two parameters σ' and σ'' in order to reduce the correlation of the fit parameters:

$$\sigma' = (1 - f)\sigma_1 + f\sigma_2, \tag{7.10}$$

$$\sigma'' = \sqrt{f(1 - f)}(\sigma_2 - \sigma_1), \tag{7.11}$$

$$\sigma_1 = \sigma' - \sqrt{\frac{f}{1 - f}}\sigma'', \tag{7.12}$$

$$\sigma_2 = \sigma' + \sqrt{\frac{1 - f}{f}}\sigma'', \tag{7.13}$$

with $(1 - f)/f \equiv f_1/f_2$ being the relative fraction between the first and second Gaussian component.

Figure 7.7 shows the fit projections when fitting the decay-time distribution in eleven bins of the `per-event decay-time uncertainty` δ_t. All parameters except the ones describing the shape of the *wrong PV* component are floating in these fits. While the maximum likelihood fit is performed for candidates in the interval $t \in [-4, 10]$ ps, the projections are only shown in a smaller range to better illustrate the different resolution components.

Calibrating the estimated decay-time uncertainty

Instead of the complicated triple Gaussian resolution model defined above, a single Gaussian function is chosen for the implementation in the final time-dependent angular fit. This allows a more reliable and simple evaluation of systematic uncertainties related to the calibration of the `per-event decay-time uncertainty`, see Chapter 10. To make sure that this simple resolution model has the same effect as the more complex one, the dilution defined in Equation (7.4) has to be

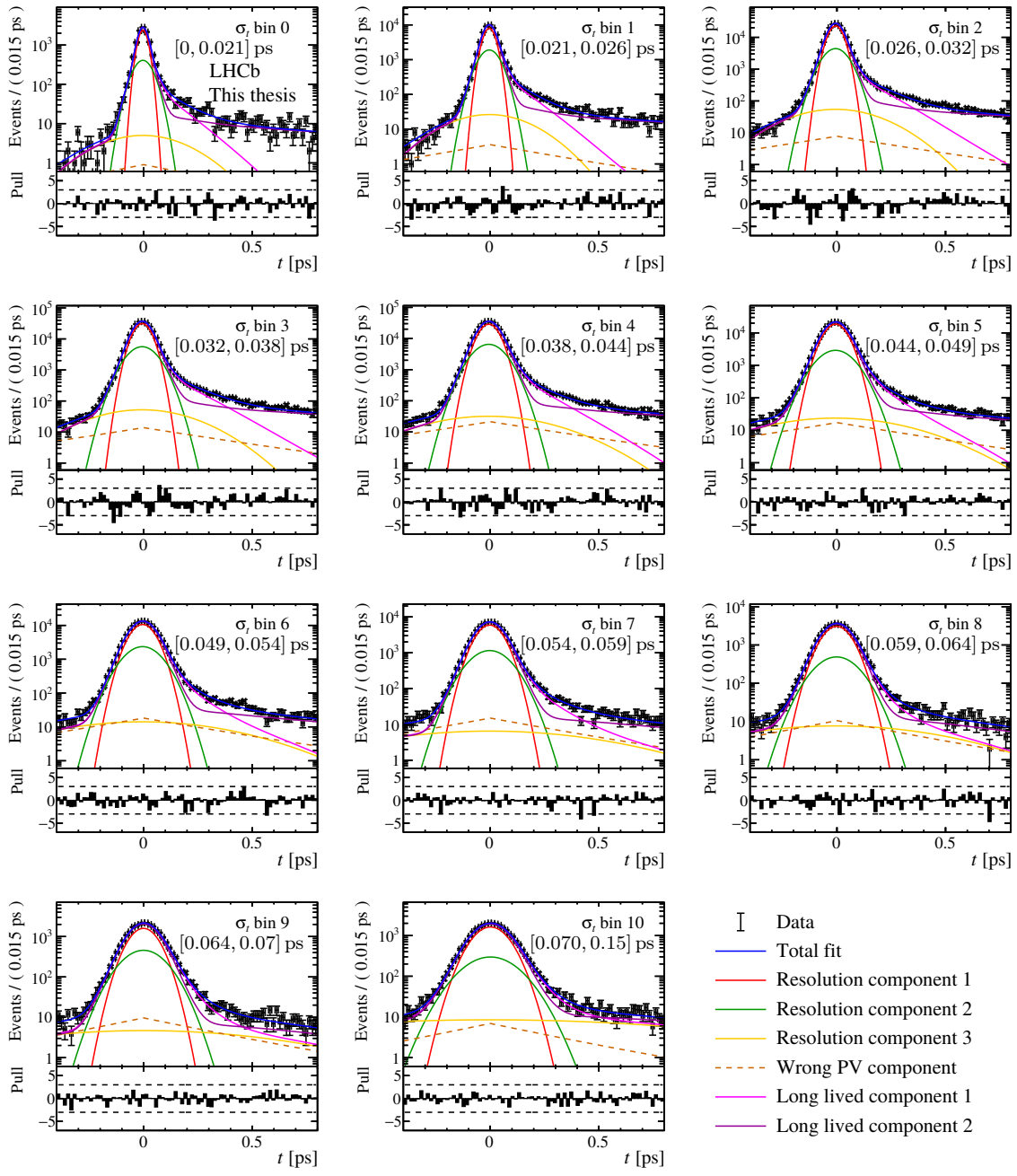


Figure 7.7: Decay-time distributions of the prompt fake $B_s^0 \to J/\psi\,\phi$ data sample in bins of the estimated decay-time uncertainty δ_t. The data histograms are overlaid by projections of the maximum likelihood fit that is used to determine the resolution function.

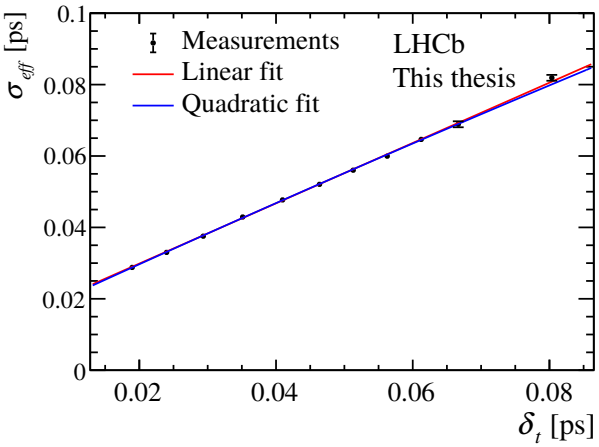


Figure 7.8: Measured effective decay-time resolution as a function of the estimated decay-time uncertainty. The linear and quadratic calibration fits are shown.

the same for both. Replacing the single Gaussian resolution in Equation (7.2) with Equation (7.9) yields the following dilution for the triple Gaussian model:

$$D = \sum_{i=1}^{3} f_i \, e^{-\sigma_i^2 \frac{\Delta m_s^2}{2}}. \tag{7.14}$$

This corresponds to the dilution of a single Gaussian with a width of:

$$\sigma_{eff} = \sqrt{-\frac{2 \ln D}{\Delta m_s^2}}, \tag{7.15}$$

where σ_{eff} is called the effective resolution.

This effective resolution is calculated from the result of the maximum likelihood fit for every bin of the estimated decay-time uncertainty δ_t and added to Figure 7.8. Especially for low values of δ_t, the Decay Tree Fitter underestimates the uncertainty on the measured decay time, which clearly demonstrates the need for a dedicated calibration. This calibration is obtained by describing the effective resolution as a function of δ_t by either a linear or a quadratic function:

$$\sigma_{eff}(\delta_t) = p_0 + p_1 \delta_t + p_2 \delta_t^2 \tag{7.16}$$

with the coefficients p_0, p_1 and p_2. The latter is set to zero in the linear case. Figure 7.8 shows the linear and the quadratic function obtained by a fit to the measured effective resolutions. Since the former yields an already good description of the measurements, the quadratic calibration is only tested as a systematic

study, see Chapter 10. The parameters of the linear calibration are given by $p_0 = (0.01297 \pm 0.00022)$ ps and $p_1 = 0.8446 \pm 0.0057$.

Applying this calibration to the background-subtracted $B_s^0 \to J/\psi\,\phi$ signal data sample yields an average effective resolution of $\langle\sigma_{eff}\rangle = (45.54 \pm 0.05 \pm 0.04)$ fs, where the first uncertainty accounts for the uncertainties on the calibration parameters and the second is due to the limited statistics of the data sample. According to Equation (7.4), this corresponds to an average dilution of $D \approx 0.72$.

7.1.3 Embedding in the probability density function

For the maximum likelihood fit to the background-subtracted sample of $B_s^0 \to J/\psi\,\phi$ decays, the resolution is modeled by a single Gaussian function with a width given by the calibrated effective resolution $\sigma_{eff}(\delta_t)$. Equation (6.10) has to be modified according to:

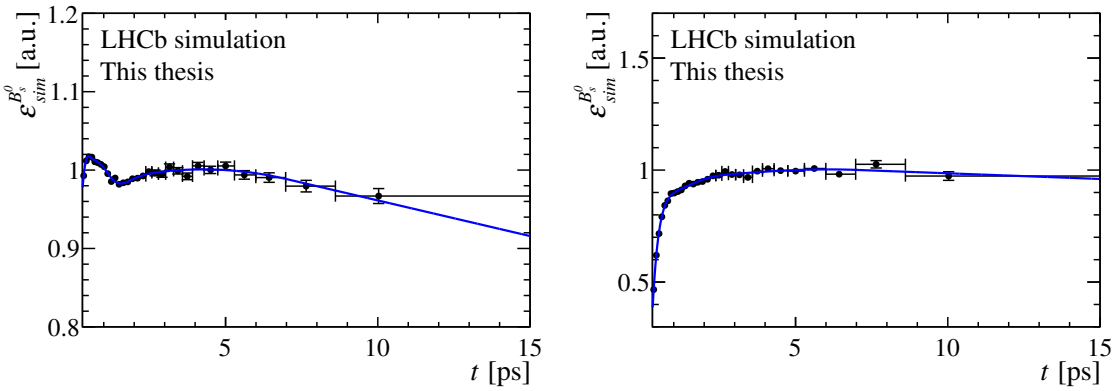
$$\begin{aligned}
&\mathrm{PDF}(t, \Omega | q^{os}, q^{ss}, \eta^{os}, \eta^{ss}, \delta_t) = \qquad (7.17)\\
&\frac{1}{N_{q^{os},q^{ss}}^{\eta^{os},\eta^{ss},\delta_t}} \sum_{k=1}^{10} A_k\, f_k(\Omega)\\
&\quad \{[(1 + q^{os}(1 - 2\omega^{os}(\eta_{os})))(1 + q^{ss}(1 - 2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t)\\
&\quad + (1 - q^{os}(1 - 2\bar{\omega}^{os}(\eta_{os})))(1 - q^{ss}(1 - 2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\},
\end{aligned}$$

where $G(t|\sigma_{eff}(\delta_t))$ is the Gaussian with a width that depends on the estimated decay-time uncertainty δ_t . The normalization is given by:

$$\begin{aligned}
&N_{q^{os},q^{ss}}^{\eta^{os},\eta^{ss},\delta_t} = \qquad (7.18)\\
&\int_{t=0.3\,\mathrm{ps}}^{15\,\mathrm{ps}} \int_{\Omega} \sum_{k=1}^{10} A_k f_k(\Omega)\, \mathrm{d}\Omega\\
&\quad \{[(1 + q^{os}(1 - 2\omega^{os}(\eta_{os})))(1 + q^{ss}(1 - 2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t)\\
&\quad + (1 - q^{os}(1 - 2\bar{\omega}^{os}(\eta_{os})))(1 - q^{ss}(1 - 2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\}\, \mathrm{d}t.
\end{aligned}$$

7.2 Decay-time acceptance

In this section a detailed discussion of the decay-time dependence of the efficiency, abbreviated with decay-time acceptance, is presented. As a first step, a huge sample of simulated $B_s^0 \to J/\psi\,\phi$ decays is used to motivate the shape of the efficiency, which receives contributions from several different effects. After that, a data driven



(a) `Unbiased` (b) `Biased`

Figure 7.9: Decay-time acceptance observed in a simulated $B_s^0 \to J/\psi\,\phi$ sample split by trigger category. The black points represent the ratio of observed and generated decays in the different decay-time bins, and the blue curves are parametrizations of the acceptances by a cubic splines, see Section 7.2.2. The absolute scale is arbitrary.

method is presented that provides analytic acceptance functions for the usage in the extraction of the physics parameters.

7.2.1 Origin of the decay-time acceptance

In Figure 7.9 the decay-time acceptance of simulated $B_s^0 \to J/\psi\,\phi$ decays after all reconstruction and selection steps is shown. The sample is split according to the two previously defined trigger categories, `Unbiased` and `Biased`. This is motivated by the significantly different decay-time acceptance of these two trigger strategies. The samples were simulated with zero decay-width difference $\Delta\Gamma_s$, which results in a generator level decay-time distribution that is completely described by a single exponential. Neglecting the effect of decay-time resolution, this allows a direct extraction of the decay-time acceptance by comparing the number of signal decays in a specific decay-time interval with the number expected from a simple exponential distribution. This ratio is shown as the black points in Figure 7.9. The blue curves in the same figure are obtained by maximum likelihood fits to the respective decay-time distribution taking into account the decay-time resolution. Details on this procedure are given in the next section. For now, these curves are only needed to lead the eye.

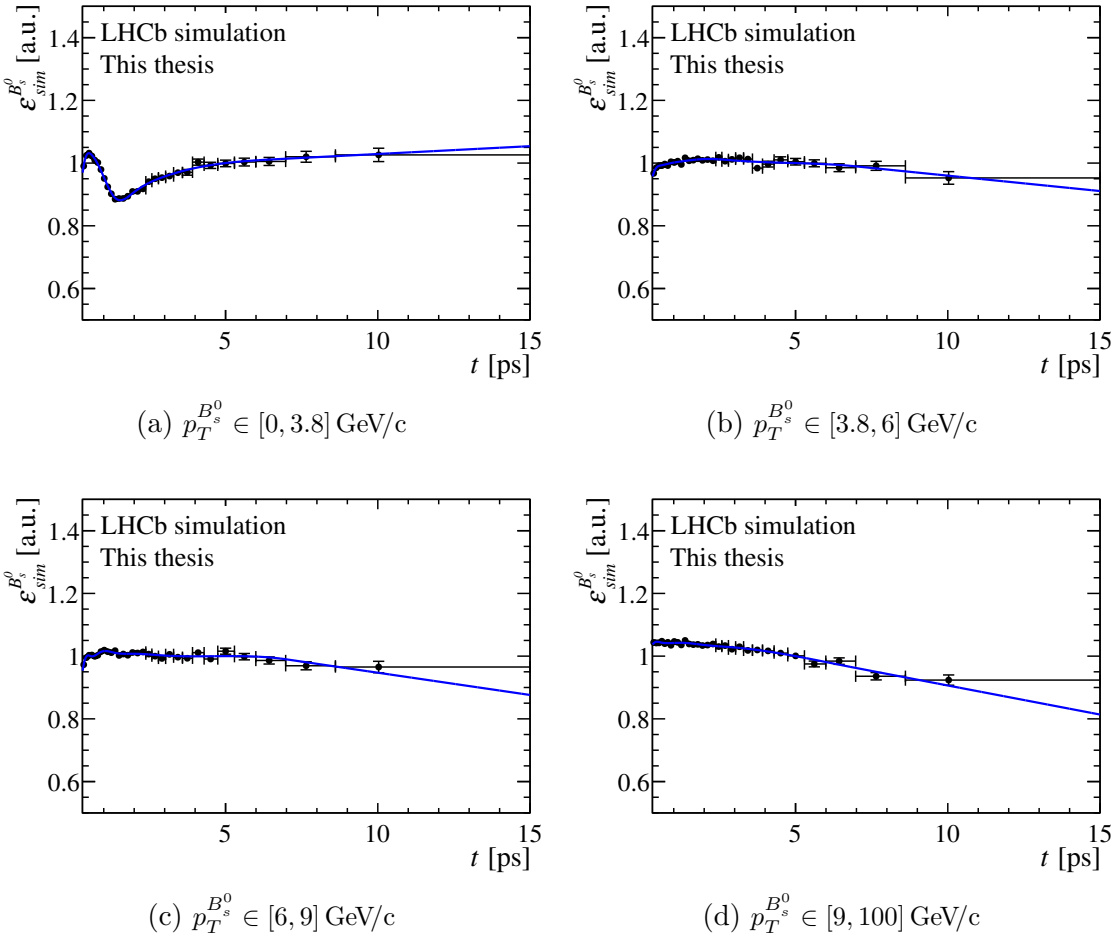
Low decay-time acceptance

When comparing the decay-time acceptance below $t \approx 3\,\text{ps}$, a clear difference between the `Unbiased` and `Biased` trigger category is visible. This difference is the defining feature of these two categories. The low efficiency at low decay times is typically caused by trigger requirements that are based on the separation of the B_s^0 decay vertex from the primary vertex (PV). In the case of the HLT1 lines that define the `Biased` category, these are the cuts on the impact parameter significance of the final state particles, see Chapter 5. A B_s^0 meson with a short decay time tends, depending on its moment, to decay near the PV. Final state tracks of this B_s^0 meson are likely to roughly point to the PV and therefore do not pass this trigger requirement.

Although the HLT1 trigger line of the `Unbiased` category does not contain such cuts, a small effect is visible at very low decay times. This is a result of the chosen HLT2 trigger line that is common for both categories. It contains a requirement of the decay length significance of the J/ψ meson, see Chapter 5. Compared to the effect on the decay-time acceptance due to the HLT1 trigger lines of the `Biased` category, this bias is small, which justifies the name of the `Unbiased` category.

High decay-time acceptance

Consistently for both trigger categories, a decreasing efficiency for very high decay times is observed in Figure 7.9. The main origin of this dependence is the pattern recognition algorithm used to find tracks in the VELO. As briefly described in Section 2.2.1, this algorithm relies on the assumption that the tracks originate from a point on the beam axis. Thus, tracks with a significant separation from this axis are reconstructed less efficiently. Decay products of B_s^0 mesons with a high decay time and therefore also long flight distance are more likely to show such large separations. Since this track reconstruction is already part of the high level trigger, this inefficiency can not be recovered offline and leads to the observed drop of efficiency at high decay times. This effect is nicely visible when determining the decay-time acceptance as a function of the transverse momentum $p_T^{B_s^0}$ of the B_s^0 meson, see Figure 7.10. For a given decay time, the transverse momentum determines the distance of the B_s^0 meson decay vertex to the beam axis. A high p_T B_s^0 meson corresponds to a large such distance, which leads to large separations of the final state tracks from the beam axis. Therefore, it is expected that the drop of efficiency at high decay times is more pronounced for high $p_T^{B_s^0}$ value. This dependence is clearly visible in Figure 7.10.



(a) $p_T^{B_s^0} \in [0, 3.8]$ GeV/c

(b) $p_T^{B_s^0} \in [3.8, 6]$ GeV/c

(c) $p_T^{B_s^0} \in [6, 9]$ GeV/c

(d) $p_T^{B_s^0} \in [9, 100]$ GeV/c

Figure 7.10: Decay-time acceptance observed in a simulated $B_s^0 \to J/\psi\,\phi$ sample for the `Unbiased` trigger category in bins of transverse momentum of the B_s^0 meson. The black points represent the ratio of observed and generated decays in the different decay-time bins, and the blue curves are parametrizations of the acceptances by cubic splines, see Section 7.2.2. The absolute scale is arbitrary.

Other effects

Another prominent feature of the decay-time acceptance is the dip at roughly $1-2$ ps in Figure 7.9. It is mostly visible for the `Unbiased` trigger category, because in the `Biased` sample it is spoiled by the dominant effect of the low decay-time acceptance discussed previously. The dip is caused by events in which the B_s^0 vertex is reconstructed as an additional PV. If this happens, the reconstructed $B_s^0 \to J/\psi\,\phi$ decay does not fulfill the vertex separation criteria and will not pass the trigger. In order to be reconstructed as PV, the B_s^0 vertex has to gather at least one additional

track, since the minimum number of tracks that are needed to form a PV is chosen to be five.

There are counteracting effects contributing to the decay-time dependence of this wrong PV reconstruction. The first is related to the true PV, which has typically more associated tracks and is reconstructed first. For very low B_s^0 decay times, the final state tracks are more likely absorbed into the fit of the first PV and are therefore not longer available to build a second one. Therefore, this wrong PV reconstruction plays no role at very small decay times. At high decay times, which means by trend higher distances from the true PV, it becomes less likely that a random track passes near the B_s^0 vertex and in this way forms an additional PV. Furthermore, there is a maximal distance a vertex is allowed to be separated from the beam axis in order to be classified as a PV. For low multiplicity vertices this distance is 0.2 mm, which is more likely do be exceeded by B_s^0 mesons with high decay time. Especially the latter two effects nicely explains the dependence of this dip as a function of $p_T^{B_s^0}$, which can be seen in Figure 7.10. At low $p_T^{B_s^0}$, many of the B_s^0 vertices fulfill the requirement of the maximal allowed distance to the beam axis and are therefore more likely reconstructed as an additional PV. For higher $p_T^{B_s^0}$ nearly all B_s^0 vertices fall outside this area and are never considered as a PV. In addition, vertices with higher separation from the beam axis do less likely pick up additional random tracks that are necessary to form a PV.

7.2.2 Strategy to determine the acceptance

After the qualitative discussion of the decay-time acceptance above, in this section the method developed to determine the acceptance of the $B_s^0 \to J/\psi\,\phi$ data sample is presented. Although the simulated sample is expected to include all basic acceptance effects, it is better to not rely on the correct description of all these interplaying contributions. Therefore, a data driven method is employed. It is based on the decay $B^0 \to J/\psi\,K^{*0}$, where the K^{*0} is decaying into a K^+ and a π^-. Topologically, this decay is very similar to the decay $B_s^0 \to J/\psi\,\phi$, and since the the decay-width splitting $\Delta\Gamma_d$ in the B^0 system is negligible [82], $B^0 \to J/\psi\,K^{*0}$ decays are produced with a simple exponentially falling decay-time distributing, which allows an easy determination of the decay-time acceptance.

The principle idea is to extract the acceptance $\varepsilon^{B_s^0}(t)$ according to:

$$\varepsilon^{B_s^0}(t) \propto \frac{N^{B^0}(t)}{e^{-\Gamma_d t} \otimes G^{B^0}(t)}, \tag{7.19}$$

where $N^{B^0}(t)$ is the observed decay-time distribution of $B^0 \to J/\psi K^{*0}$ decays and $G^{B^0}(t)$ the decay-time resolution function of this sample. This means that the extracted decay-time acceptance depends on the chosen value of the decay width Γ_d of the B^0 meson. The current world average is given by $\Gamma_d^{\text{w.a.}} = (0.6579 \pm 0.0017)\,\text{ps}^{-1}$ [16], and the decay-time acceptance for any arbitrary value of Γ_d can be parametrized using a deviation $\delta\Gamma_d$ from this value:

$$\Gamma_d = \Gamma_d^{\text{w.a.}} + \delta\Gamma_d. \tag{7.20}$$

The decay-time acceptance for a given value of Γ_d becomes then:

$$\varepsilon^{B_s^0}(t|\Gamma_d) \propto \frac{N^{B^0}(t)}{e^{-(\Gamma_d^{\text{w.a.}}+\delta\Gamma_d)t} \otimes G^{B^0}(t)} \tag{7.21}$$

$$\approx e^{\delta\Gamma_d t} \frac{N^{B^0}(t)}{e^{-\Gamma_d^{\text{w.a.}} t} \otimes G^{B^0}(t)}, \tag{7.22}$$

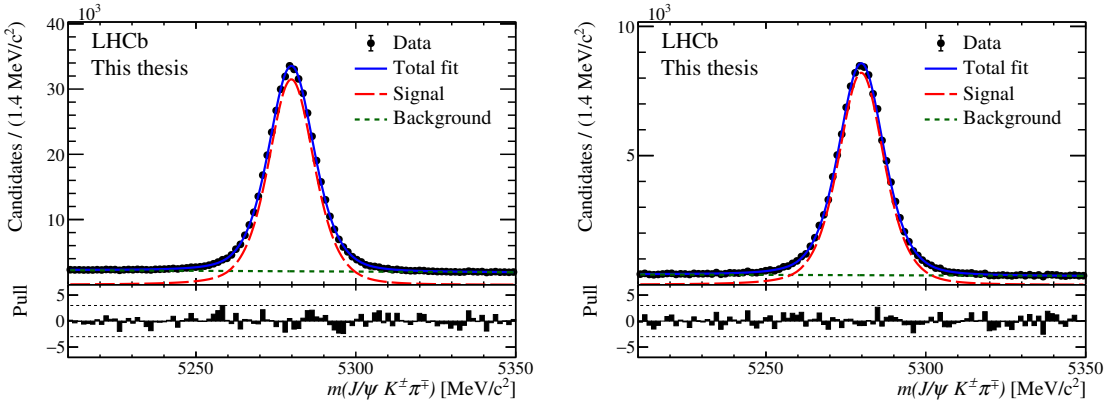
where the last step is valid because the decay-time resolution ($\mathcal{O}(0.045\,\text{ps})$) is significantly smaller than the lower bound on the decay time (0.3 ps). Thus, the decay-time acceptance changes by an exponentially time-dependent factor that is directly given by the change in Γ_d. Thus, by a similar argument, such a change of Γ_d in the acceptance determination would result in the same shift of the measured decay width Γ_s of the B_s^0 system. One way to deal with this is to quote a value of Γ_s that receives a systematic uncertainty arising from the uncertainty of $\Gamma_d^{\text{w.a.}}$. A more elegant solution is to quote the difference between the decay widths, $\Delta\Gamma_d^s = \Gamma_s - \Gamma_d$, which is independent of the exact value of Γ_d used in the acceptance determination and is the actual quantity that is measured when taking the decay-time acceptance from a $B^0 \to J/\psi K^{*0}$ sample. Furthermore, this difference can be easily related to the ratio Γ_s/Γ_d, which is of high theoretical interest, see Section 1.5.8.

Selection of $B^0 \to J/\psi K^{*0}$ decays

The similarity between the decays $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ allows a nearly identical selection strategy for those two channels. In the following, only differences with respect to the selection presented in Chapter 5 are mentioned. While the chosen trigger configuration is not changed, there are a few small differences in the first step of the offline selection. The respective requirements are shown in Table 7.1. The tighter requirements on the transverse momentum of the pion and the K^{*0} resonance account for the larger amount of combinatorial background with respect to the channel $B_s^0 \to J/\psi \phi$ due to the presence of a pion in the final state.

Table 7.1: Selection criteria applied in the first step of the offline selection of $B^0 \to J/\psi K^{*0}$ decays. Only differences with respect to the selection applied to $B_s^0 \to J/\psi \phi$ candidates, see Chapter 5, are given.

Variable	Requirement
$p_T(\pi^-)$	$> 250\,\text{MeV/c}$
$p_T(K^{*0})$	$> 1300\,\text{MeV/c}$
$m(K^+\pi^-)$	$\in [826, 966]\,\text{MeV/c2}$
$\Delta\ln\mathcal{L}_{K\pi}(\pi^-)$	< 0
$m(J/\psi\, K^+\pi^-)$	$\in [5150, 5350]\,\text{MeV/c}^2$



(a) `Unbiased` (b) `Biased`

Figure 7.11: Invariant mass distribution of selected $B^0 \to J/\psi K^{*0}$ candidates. They are split according to the trigger category. Projections of fits are overlaid.

After this first step of the offline selection, the same multivariate classifier as for the B_s^0 decay channel is applied. The transverse momentum of the ϕ meson is replaced by the one of the K^{*0} resonance, but the exact same classifier is used without a dedicated training for the B^0 decay. Although the input variables were chosen such that the impact on the decay-time distribution is small, there are small correlations between the decay time and these variables. Using the identical selection for both channels, ensures that these correlations enter the same way in both cases.

Figure 7.11 shows the invariant mass distribution of the selected $B^0 \to J/\psi K^{*0}$ candidates, split by trigger category. In analogy to the B_s^0 decay, the ${}_s\mathcal{P}lot$ technique is employed to statistically subtract the remaining background contribution. Since no angular analysis of the decay $B^0 \to J/\psi K^{*0}$ is needed, the correlation between the mass resolution and the angular variables is not as crucial as it is in the case of the B_s^0 decay. Therefore, the maximum likelihood fit to the mass distribution can be performed without taking into account the `per-event mass uncertainty`. The

Table 7.2: Approximate number of $B^0 \to J/\psi K^{*0}$ decays after the full selection. The yields are determined by fits to the invariant mass distribution.

Year	Trigger category	Yield
2015	`Unbiased`	56.5k
	`Biased`	18.2k
2016	`Unbiased`	381.5k
	`Biased`	90.0k

background is modeled using an exponential function and the signal is described by a double-sided Ipatia function [76]. The latter allows a correct description of the tails of the signal distribution. As in the case of the B_s^0 decay, the shape parameters are determined from a simulated $B^0 \to J/\psi K^{*0}$ sample. Combined fit projections of the two years of data taking are shown in Figure 7.11. For orientation, Table 7.2 shows the approximate signal yields of the different categories determined by the fit. These yields are significantly larger than the ones of the decay $B_s^0 \to J/\psi \phi$ quoted in Chapter 5 because of the significant larger B^0 hadronization fraction [83].

Correcting differences between $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ decays

Although the decay $B^0 \to J/\psi K^{*0}$ is topologically similar, there are some second order differences with respect to the decay $B_s^0 \to J/\psi \phi$. These might cause differences in the decay-time acceptance between the two channels and have to be taken into account. The main relevant difference is the phase space that is available for the final state particles of the hadronic resonance. While the mass of the two kaons, $m(K^+) + m(K^-) \approx 987\,\mathrm{MeV/c^2}$ [3], is relatively close to the mass of the ϕ meson, $m_\phi \approx 1020\,\mathrm{MeV/c^2}$ [3], the mass of the kaon-pion pair, $m(K^+)+m(\pi^-) \approx 633\,\mathrm{MeV/c^2}$ [3], is significantly lower than the mass of the K^{*0} resonance, $m_{K^{*0}} \approx 892\,\mathrm{MeV/c^2}$ [3]. Thus, the opening angle of the final state mesons tends to be larger in the B^0 decay in comparison to the B_s^0 decay. Larger opening angles can lead to a larger separation of the track from the primary vertex and also from the beam axis. Following the discussion in Section 7.2.1, this can have an effect on the decay-time acceptance.

In order to correct for this and other possible sources of biases, the acceptance determined from the $B^0 \to J/\psi K^{*0}$ data sample, $\varepsilon_{data}^{B^0}$, is corrected using the acceptances observed in simulated $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ decays, $\varepsilon_{sim}^{B^0}$ and $\varepsilon_{sim}^{B_s^0}$:

$$\varepsilon_{data}^{B_s^0}(t) = \frac{\varepsilon_{sim}^{B_s^0}(t)}{\varepsilon_{sim}^{B^0}(t)} \times \varepsilon_{data}^{B^0}(t). \tag{7.23}$$

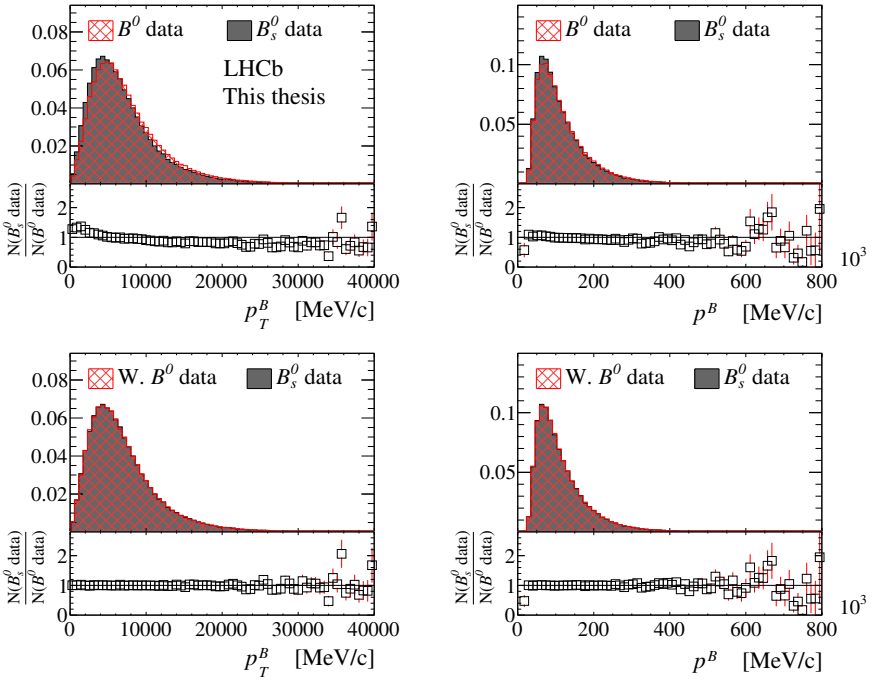


Figure 7.12: Momentum and transverse momentum distributions of the background-subtracted $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ data samples. The top row shows the distributions before and the bottom row the distributions after the weighting of the $B^0 \to J/\psi K^{*0}$ candidates. Only the 2016 data sets are shown.

In addition, the background-subtracted B^0 data sample is weighted[1] to match the respective B_s^0 data distributions of the momentum and transverse momentum of the B mesons. Neglecting the small mass difference between the two B mesons, the weighting in momentum ensures that the relation between decay time and flight distance is the same for both channels. The weighting in the transverse momentum is motivated by the strong dependence of the decay-time acceptance on this quantity, see Figure 7.10. Figure 7.12 shows the momentum and transverse momentum distributions of $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ decays before and after the weighting. A good agreement is achieved.

Besides these corrections of the B^0 data sample, the simulated B_s^0 and B^0 samples are corrected to be more similar to the respective data set. In contrast to the real data sets, the simulated samples do not contain S-wave components, which means that all their final state hadrons originate from an intermediate ϕ or K^{*0} meson. Furthermore, the relative phases and fractions of the different polarization states do not necessarily agree with the ones observed in real data. These two effects

[1] Throughout this thesis, the boosted weighting technique, see Section 4.2, is employed when not stated otherwise.

can cause differences in the angular distributions of the final state particles, which might affect the decay-time acceptance. In order to correct these differences, the simulated samples are weighted to match the S-wave and polarization fractions and phases measured in earlier LHCb analyses [5], [82]. This weighting can be achieved in two steps. In a first step, every simulated event gets a weight that is determined by evaluating the decay-time-dependent angular distribution function (PDF) for this event once with the physics parameters set to the ones used in the simulation (λ^{gen}) and once with the parameters observed in data (λ^{obs}). The ratio of these two values is then used as the weight ω:

$$\omega = \frac{\mathrm{PDF}(t_e, \Omega_e|\lambda^{obs})}{\mathrm{PDF}(t_e, \Omega_e|\lambda^{gen})}. \tag{7.24}$$

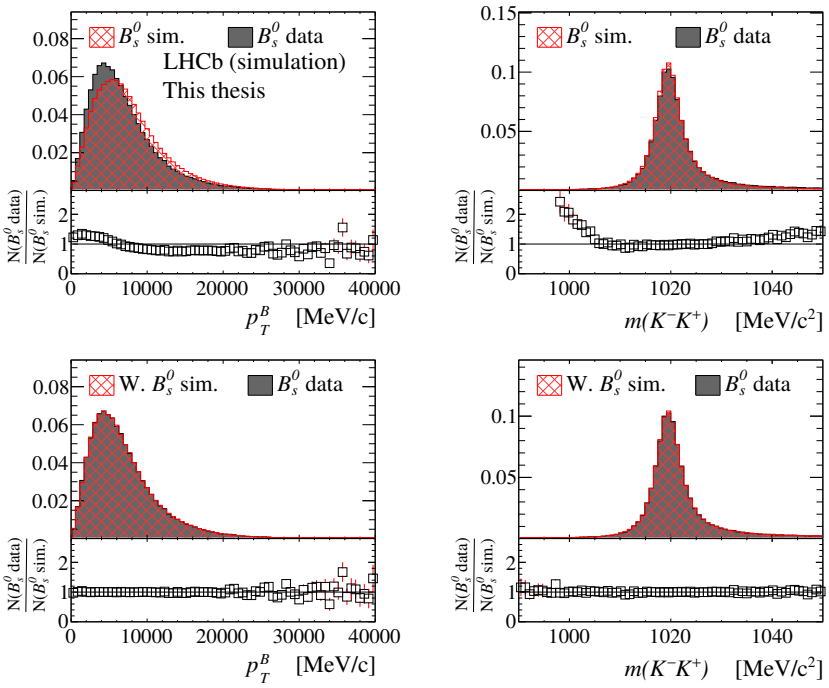
Here, t_e is the decay time and Ω_e represents the angles in the helicity basis of the respective event. The fraction of the S-wave component is part of the PDF and was determined in bins of the mass of the K^+K^- and $K^+\pi^-$ system, respectively. Therefore, a second step is necessary to obtain the correct total fraction of the S-wave: the mass distribution of the hadron system is aligned between simulated and data candidates in terms of a weighting. Together with this last step, again a weighting of the simulated samples according to the differences to data in the transverse momentum distributions of the B_s^0/B^0 meson is performed.

Figure 7.13 shows the distributions of data and simulation before and after this weighting. Again, a good agreement is achieved. The total effect of the corrections applied to the simulated samples can be seen in Figure 7.14. There, the decay-time acceptance of simulated $B_s^0 \to J/\psi\,\phi$ decays is shown with and without them.

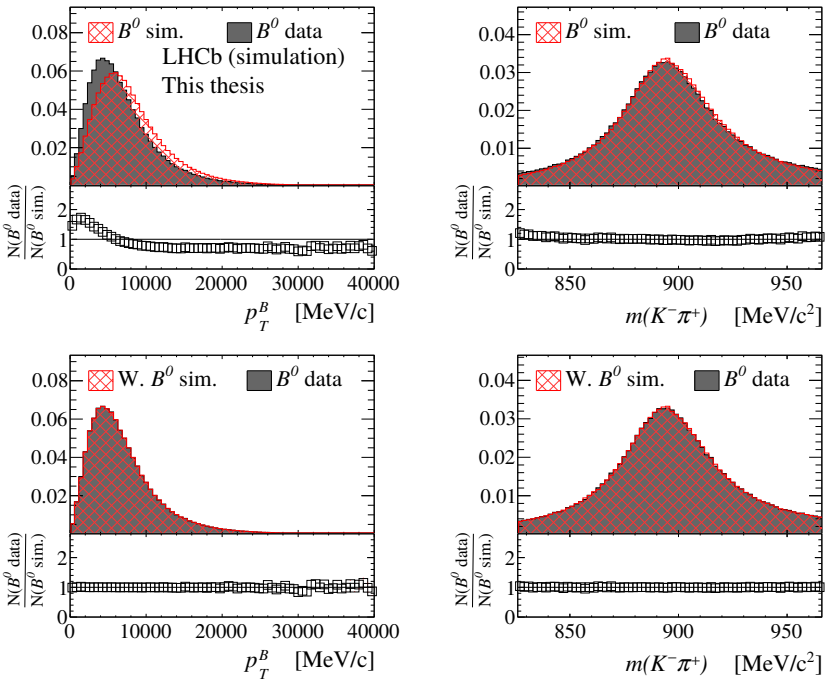
Extracting acceptance parametrizations

In the previous sections, the origin of the decay-time acceptance and the principle strategy of its determination was discussed. Now the technical method to obtain an analytical parametrization of the final acceptances including all corrections is presented. Such an analytical parametrization allows a direct implementation of the decay-time acceptance into the PDF and is furthermore less sensitive to statistical fluctuations that are present when using directly the acceptance histograms.

It was chosen to employ cubic splines for this purpose. A cubic spline is defined on a set of intervals, which are defined by the so-called knot positions or knots [84]. In each of these intervals, the cubic spline is given by a third order polynomial. At every knot position, the spline is required to be continuously differentiable. Given a cubic spline with N knots, this requirement allows for $N + 2$ freely varying

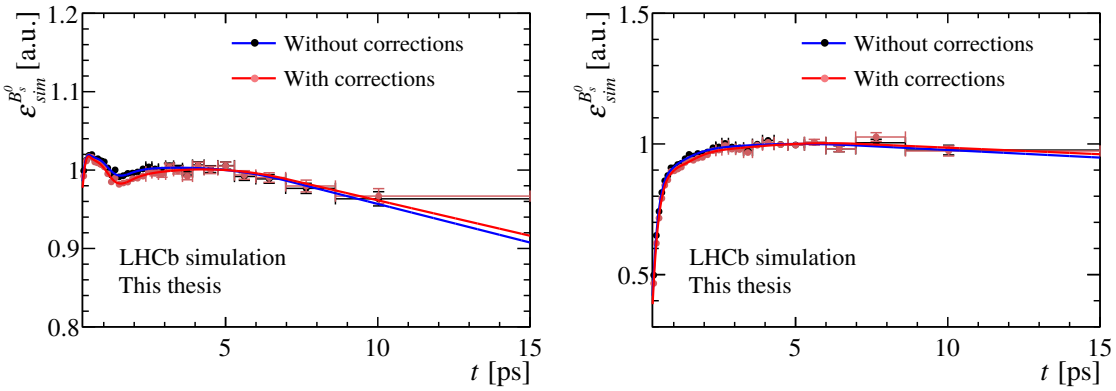


(a) $B_s^0 \to J/\psi\,\phi$



(b) $B^0 \to J/\psi\,K^{*0}$

Figure 7.13: Simulation and background-subtracted data distributions of $B^0 \to J/\psi\,K^{*0}$ and $B_s^0 \to J/\psi\,\phi$ decays. The respective top row shows the distributions before and the bottom row the distributions after the weighting of the simulated candidates. Only the 2016 data sets and corresponding simulated candidates are shown.



(a) `Unbiased` (b) `Biased`

Figure 7.14: Change of the decay-time acceptance of simulated $B_s^0 \to J/\psi\,\phi$ decays due to the corrections applied to math the $B_s^0 \to J/\psi\,\phi$ data sample. The points represent the ratio of observed and generated decays in the different bins of the reconstructed decay time, and the solid curves are parametrizations of the acceptances by a cubic spline, see Section 7.2.2. The absolute scale is arbitrary.

parameters. One such basis of parameters that is completely describing the spline is given by the spline values at the knot positions and the slopes of the spline at the first and last knot. A more common and robust parametrization, which is also chosen in this thesis, is based on the decomposition of the cubic spline in $N+2$ independent B-splines [85]. Every cubic spline $S_3(x)$ can be written as the sum of these B-splines $B_{i,3}(x)$:

$$S_3(x) = c_i \sum_i B_{i,3}(x). \tag{7.25}$$

The coefficients c_i are then the parameters that uniquely define the cubic spline.

Motivated by the exponential distribution of the decay time, the knots are chosen to be distributed in a similar way. Knots at $[0.3, 0.58, 0.91, 1.35, 1.96, 3.01, 7.00]$ ps turned out to be a good choice to describe all features of the acceptance without picking up to many random fluctuations. Since this choice is rather arbitrary, a systematic uncertainty will be assigned by choosing another set of knot positions. After the last knot, the spline is extrapolated with a linear function. This is one further way to get less sensitive to statistical fluctuations, which are more relevant at the sparely populated high decay-time region.

The aim is to determine a single cubic spline that parametrizes the final B_s^0 data decay-time acceptance defined in Equation (7.23). This is achieved by a

simultaneous fit to the decay-time distributions of the $B^0 \to J/\psi K^{*0}$ data set and the simulated samples of $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ decays. Each distribution is described by an exponential function that is folded with a Gaussian resolution function and multiplied with an acceptance. The acceptances, $\varepsilon_{sim}^{B_s^0}(t)$, $\varepsilon_{sim}^{B^0}(t)$ and $\varepsilon_{data}^{B^0}(t)$, of the three samples are parametrized as:

$$\varepsilon_{sim}^{B_s^0}(t) = s_{sim}^{B_s^0}(t), \tag{7.26}$$

$$\varepsilon_{sim}^{B^0}(t) = s_{sim}^{B^0/B_s^0}(t) \times s_{sim}^{B_s^0}(t), \tag{7.27}$$

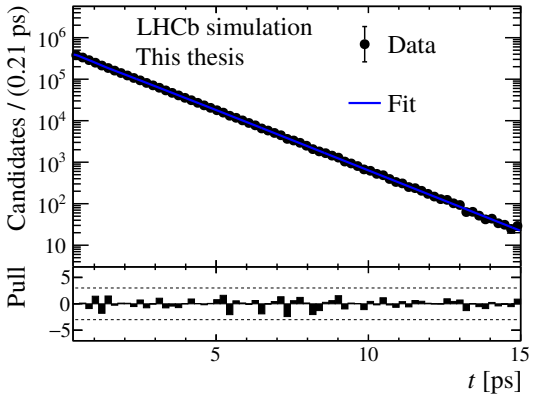
$$\varepsilon_{data}^{B^0}(t) = s_{data}^{B_s^0}(t) \times s_{sim}^{B^0/B_s^0}(t), \tag{7.28}$$

using the three cubic splines $s_{sim}^{B_s^0}(t)$, $s_{sim}^{B^0/B_s^0}(t)$ and $s_{data}^{B_s^0}(t)$. With this definition, the spline $s_{data}^{B_s^0}(t)$ represents exactly the desired combination of acceptances of Equation (7.23): $\varepsilon_{data}^{B_s^0}(t) = s_{data}^{B_s^0}(t)$.

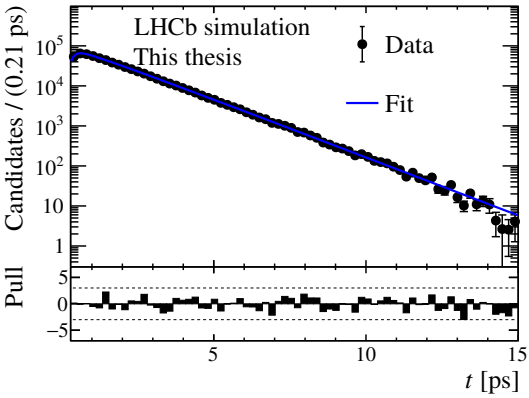
In the fit to the three decay-time distributions, the respective decay widths are fixed to the value used in the respective simulation and to the current world average of Γ_d in the case of the $B^0 \to J/\psi K^{*0}$ data set. The widths of the Gaussian resolution functions are motivated by the study in Section 7.1 and by the ratio of decay-time resolution widths observed in the two simulated samples. The values are $\sigma_{sim}^{B_s^0} = 41.7\,\text{fs}$, $\sigma_{sim}^{B^0} = 38.6\,\text{fs}$ and $\sigma_{data}^{B^0} = 42.4\,\text{fs}$. The free parameters of the fit are the coefficients of the three cubic splines, where the respective first coefficient is fixed to 1 to fix the arbitrary overall scale of the acceptance functions.

For each trigger category and year of data taking, separated acceptances are determined in this way. Figure 7.15 shows the fit projections of the 2016 samples. The respective determined cubic splines are shown in Figure 7.16. While the ratio of acceptances in the two simulated samples is nearly flat for the `Unbiased` trigger category, there is a clear difference visible for the candidates of the `Biased` trigger category. The selection of $B^0 \to J/\psi K^{*0}$ decays is more efficient at low decay times. This might be explained by the difference in the opening angles of the final state particles. Even for `Biased` trigger categories the corrections based on the simulated samples are significantly smaller than the acceptance effects themselves. This justifies the term "data driven method".

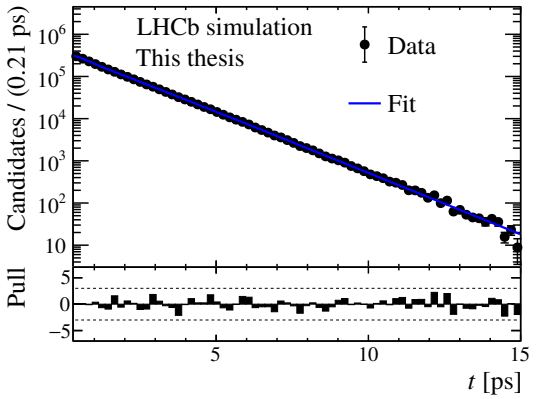
The coefficients of the cubic splines describing the acceptance in the $B_s^0 \to J/\psi \phi$ data sample are given in Table 7.3. Since they are obtained from a simultaneous fit to the three previously mentioned samples, they come with a covariance matrix that represents the uncertainty due to the limited statistics of all these three samples. This will allow an easy determination of a systematic uncertainty in Chapter 10.



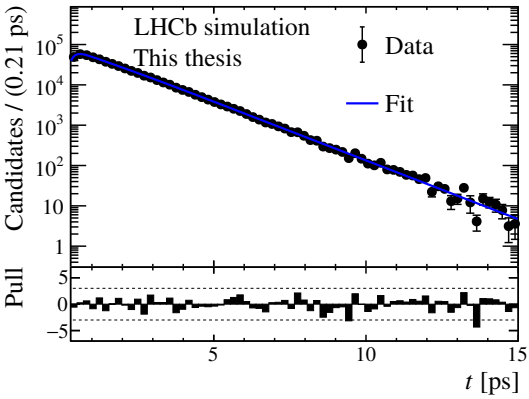
(a) $B_s^0 \to J/\psi\,\phi$ simulation `Unbiased`



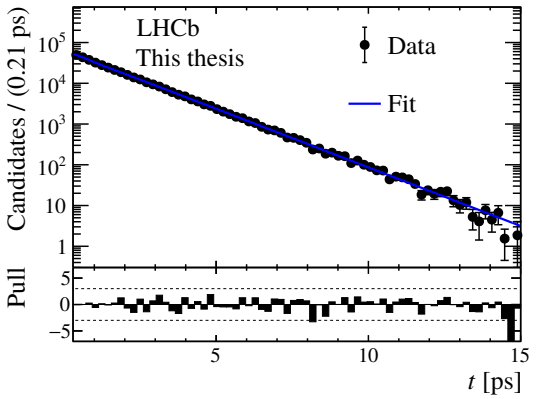
(b) $B_s^0 \to J/\psi\,\phi$ simulation `Biased`



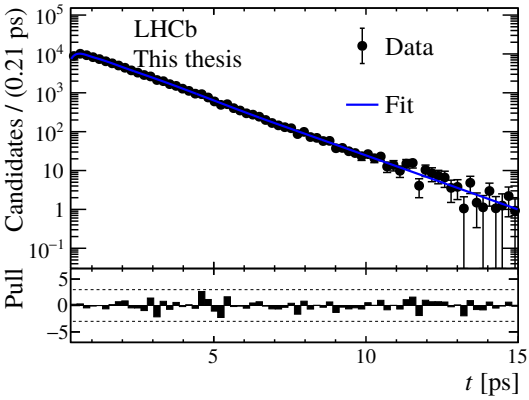
(c) $B^0 \to J/\psi\,K^{*0}$ simulation `Unbiased`



(d) $B^0 \to J/\psi\,K^{*0}$ simulation `Biased`

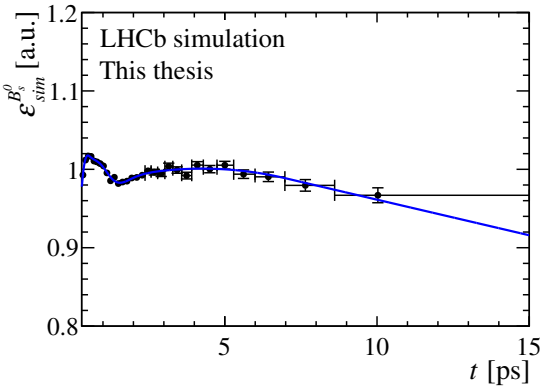


(e) $B^0 \to J/\psi\,K^{*0}$ data `Unbiased`

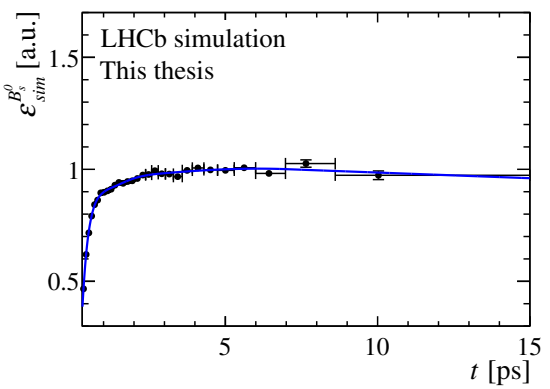


(f) $B^0 \to J/\psi\,K^{*0}$ data `Biased`

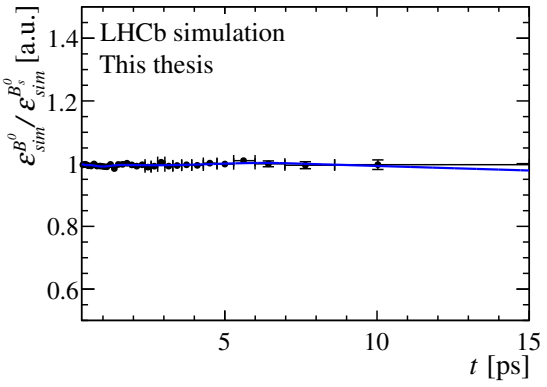
Figure 7.15: Background subtracted decay-time distributions and fit projections of the simultaneous fit to the three samples used to determine the decay-time acceptance. The 2016 samples are shown and split according to the two trigger categories.



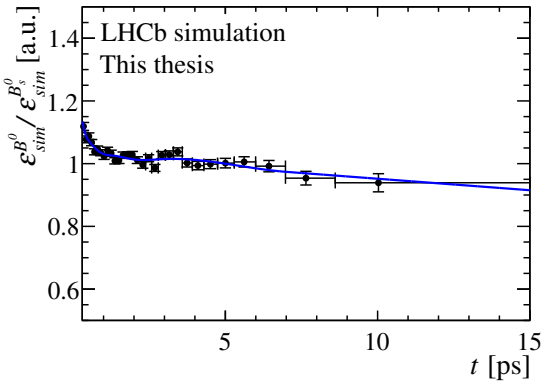
(a) $B_s^0 \to J/\psi\,\phi$ simulation, `Unbiased`



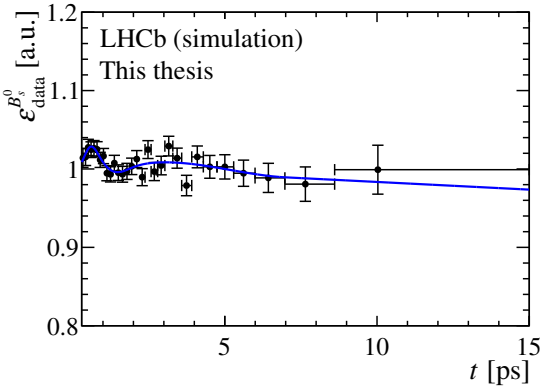
(b) $B_s^0 \to J/\psi\,\phi$ simulation, `Biased`



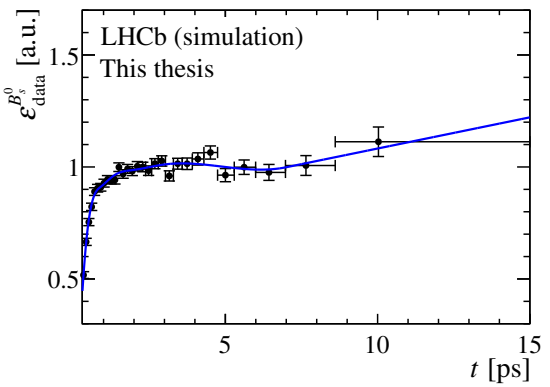
(c) Ratio of simulation acceptances, `Unbiased`



(d) Ratio of simulation acceptances, `Biased`



(e) Final acceptance, `Unbiased`



(f) Final acceptance, `Biased`

Figure 7.16: Acceptance splines obtained by the simultaneous fit to $B^0 \to J/\psi\,K^{*0}$ data and simulation and $B_s^0 \to J/\psi\,\phi$ simulation. The black points are obtained by acceptance histograms, and the blue lines show the cubic splines. Only the 2016 categories are shown.

Table 7.3: Coefficients of the cubic splines that describe the decay-time acceptance of the $B_s^0 \to J/\psi\,\phi$ data sample. They are determined by the simultaneous fit described in the text. The first coefficient is set to 1 in all four data categories to fix the arbitrary absolute scale.

	2015		2016	
	`Unbiased`	`Biased`	`Unbiased`	`Biased`
c_0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0
c_1	1.05 ± 0.07	1.69 ± 0.31	1.008 ± 0.028	1.49 ± 0.16
c_2	1.097 ± 0.049	1.73 ± 0.21	1.031 ± 0.018	2.06 ± 0.13
c_3	0.969 ± 0.049	1.85 ± 0.27	1.001 ± 0.021	2.12 ± 0.17
c_4	1.051 ± 0.051	1.99 ± 0.26	0.984 ± 0.018	2.28 ± 0.16
c_5	1.05 ± 0.047	1.92 ± 0.26	1.0 ± 0.019	2.29 ± 0.17
c_6	1.028 ± 0.064	2.0 ± 0.3	1.009 ± 0.024	2.46 ± 0.19
c_7	1.094 ± 0.06	2.19 ± 0.31	0.989 ± 0.024	2.25 ± 0.18
c_8	1.051 ± 0.05	1.95 ± 0.27	0.987 ± 0.02	2.34 ± 0.17

7.2.3 Validation of the Strategy

Since the strategy for the determination of the decay-time acceptance has changed with respect to previous versions of the analysis, it demands dedicated validation tests. Especially the correction for differences between the decay channels $B_s^0 \to J/\psi\,\phi$ and $B^0 \to J/\psi\,K^{*0}$ via simulation needs to be verified. To do so, two kinds of tests are performed. The first is based on a measurement of the decay-width difference, $\Delta\Gamma_d^u = \Gamma_u - \Gamma_d$, between the channels $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^{*0}$. The decay-time acceptance of the $B^+ \to J/\psi\,K^+$ candidates is determined in full analogy to the method presented above, with replacing the simulated $B_s^0 \to J/\psi\,\phi$ sample with a respective $B^+ \to J/\psi\,K^+$ one. The channel $B^+ \to J/\psi\,K^+$ was triggered in the same way as the other two channels, but the offline selection, including the BDT, has to be changed slightly due to the missing fourth final state track. Details are given in Ref. [65]. Using this decay-time acceptance and a fixed decay-time resolution obtained in the same way as for the $B^0 \to J/\psi\,K^{*0}$ data sample, a maximum likelihood fit to the decay-time distribution of the background-subtracted $B^+ \to J/\psi\,K^+$ data sample is performed.

Figure 7.17 shows the fit projection of this fit for the 2016 data sample. Transforming the measured decay-width difference between the B^+ and B^0 meson, $\Delta\Gamma_d^u$, to the ratio of their lifetimes, the fit returns: $\tau(B^+)/\tau(B^0) = 1.0783 \pm 0.0024$. Only statistical uncertainties are considered. This is in good agreement with the current world average of $\tau(B^+)/\tau(B^0)_{\text{w.a.}} = 1.076 \pm 0.004$ [16]. This gives confidence that

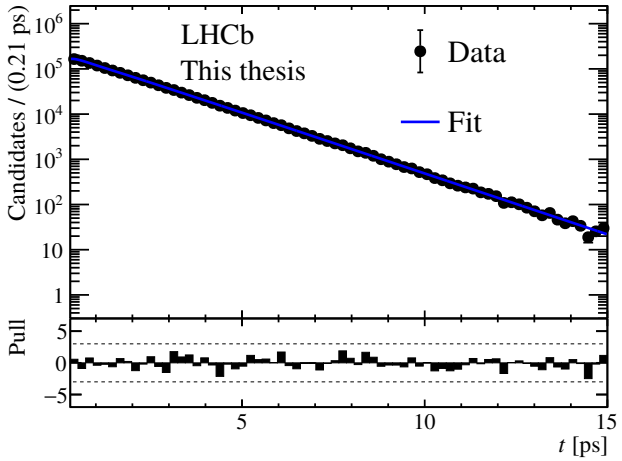


Figure 7.17: Fit projection of the maximum likelihood fit to the background-subtracted $B^+ \to J/\psi K^+$ data sample.

the method does also work for the channel $B_s^0 \to J/\psi\,\phi$ that is even more similar to the channel $B^0 \to J/\psi\, K^{*0}$.

In a second validation study, the $B^0 \to J/\psi\, K^{*0}$ data and simulation samples are split in two. In this way, two independent sets, labeled as A and B, are obtained, which are then used as calibration and testing samples, respectively. The idea is to determine the decay-time acceptance for the data sample A using the data sample B and the simulated samples A and B. This is done in full analogy to the procedure of the $B_s^0 \to J/\psi\,\phi$ decay-time acceptance determination. With this decay-time acceptance at hand, a maximum likelihood fit to the decay-time distribution of the background-subtracted data sample A is performed, and the lifetime is determined. In a first step, the samples are split randomly. By definition this should reproduce the input value for the B^0 lifetime of $\tau_{B^0} = 1.520\,\text{ps}$ and sets the baseline for the further tests. The same random splitting is used, but additional cuts are placed on the samples A:

- A cut on the per-event decay-time uncertainty δ_t: $\delta_t < 0.04\,\text{ps}$.
- A cut on the opening angle of the $K^+\pi^-$ pair α: $\alpha < 0.025\,\text{rad}$.

For each of them, the procedure is repeated. As an additional test, the samples are split according to the mass of the K^* resonance. All these cuts and splittings are chosen to mimic the differences between the channels $B_s^0 \to J/\psi\,\phi$ and $B^0 \to J/\psi\, K^{*0}$, which differ in the decay-time resolution and the hadronic resonance. The same correction procedure as discussed in Section 7.2.2 is applied to the respective samples. However, the tests are also repeated without applying these corrections. This allows

Table 7.4: Values of $\tau(B^0)$, obtained during the validation of the decay-time acceptance method, for the different considered splittings, with (first column) and without applying the corrections discussed in Section 7.2.2.

Splitting	Without corrections	With corrections
Random	1.513 ± 0.006 ps	-
Random + $\delta_t < 0.04$ ps	1.489 ± 0.006 ps	1.507 ± 0.007 ps
Random + $\alpha < 0.025$ rad	1.507 ± 0.007 ps	1.511 ± 0.008 ps
$m(K^{*0})$	1.522 ± 0.006 ps	1.523 ± 0.006 ps

to visualize their effect. Table 7.4 summarizes the measured decay time. While the lifetimes obtained with the random splitting and the splitting in $m(K^{*0})$ have to be compared to the input value of $\tau(B^0) = 1.520$ ps, the values obtained by applying additional cuts on top of the random splitting have to be compared to the value obtained from the random splitting alone, $\tau(B^0) = 1.513$ ps. Of course these values are correlated, but when estimating the uncorrelated statistical uncertainty as $\sqrt{\sigma_2^2 - \sigma_1^2}$, where σ_2 and σ_1 are the uncertainties form the test with and without the additional cut, none of the values show a significant bias.

As in the case of the validation with the channel $B^+ \to J/\psi K^+$, this gives confidence that the decay-time acceptance for $B_s^0 \to J/\psi\,\phi$ decays is correctly determined.

7.2.4 Embedding in the probability density function

The cubic splines describing the decay-time acceptance for the different years of data taking, y, and trigger categories, c, are included in the PDF of the time-dependent angular fit of $B_s^0 \to J/\psi\,\phi$ data. Equation (7.17) has to be modified according to:

$$\begin{aligned}
&\mathrm{PDF}_{y,c}(t,\Omega|q^{os},q^{ss},\eta^{os},\eta^{ss},\delta_t) = \qquad (7.29)\\
&\frac{1}{N^{\eta^{os},\eta^{ss},\delta_t}_{q^{os},q^{ss},y,c}} \sum_{k=1}^{10} A_k\, f_k(\Omega)\, \varepsilon_{y,c}(t)\\
&\quad \{[(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t)\\
&\quad + (1-q^{os}(1-2\overline{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\},
\end{aligned}$$

where $\varepsilon_{y,c}(t)$ is the respective cubic spline. The normalization is given by:

$$N^{\eta^{os},\eta^{ss},\delta_t}_{q^{os},q^{ss},y,c} = \tag{7.30}$$
$$\int_{t=0.3\,\mathrm{ps}}^{15\,\mathrm{ps}} \int_{\Omega} \sum_{k=1}^{10} A_k\, f_k(\Omega)\, \varepsilon_{y,c}(t)\, \mathrm{d}\Omega$$
$$\{[(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t)$$
$$+ (1-q^{os}(1-2\bar{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\}\, \mathrm{d}t.$$

The representation of the acceptance with a single cubic spline allows the analytic integration of the PDF [86].

7.3 Angular acceptance

As discussed in Section 1.5.2, an angular analysis is necessary to disentangle the different CP components of the $J/\psi\,\phi$ final state. Any possible dependence of the reconstruction and selection efficiency on the three helicity angles can bias the determined relative fractions of these components and therefore also the extracted CP-violation and lifetime parameters. Thus, it is crucial to incorporate this angular dependence of the efficiency, called angular acceptance, in the maximum likelihood fit to data. This section will start with a motivation of the angular acceptance shape using large simulated samples. After that, the principle method used for the integration into the PDF is presented. The final determination of the angular acceptance is then based on a sample of simulated $B^0_s \to J/\psi\,\phi$ decays and a dedicated technique developed to correct for possible differences between simulation and data, which is presented at the end of this section.

The angular acceptance, $\varepsilon(\theta_\mu, \theta_K, \varphi_h)$, is defined as follows:

$$\varepsilon(\theta_\mu, \theta_K, \varphi_h) \propto \frac{N^{obs}(\theta_\mu, \theta_K, \varphi_h|t, q)}{N^{exp}(\theta_\mu, \theta_K, \varphi_h|t, q)}, \tag{7.31}$$

where $N^{obs}(\theta_\mu, \theta_K, \varphi_h|t, q)$ is the angular dependent density of the observed data at a give reconstructed decay time t and initial flavour q, while $N^{exp}(\theta_\mu, \theta_K, \varphi_h|t, q)$ is the respective expected density function of the decay at generation. Practically, the shape of the angular acceptance can be obtained from a simulated $B^0_s \to J/\psi\,\phi$ sample by splitting the candidates in three dimensional bins, Ω_i, of the three helicity

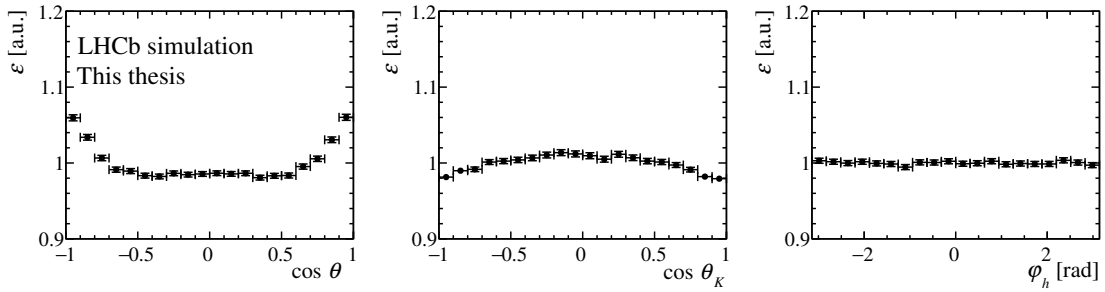


Figure 7.18: Projections of the three dimensional angular acceptance histogram obtained from a simulated sample of $B_s^0 \to J/\psi\,\phi$ decays.

angles and by comparing the observed numbers, $N^{obs}(\Omega_i)$, to the ones expected from the density function at generation:

$$\varepsilon(\Omega_i) \propto \frac{N^{obs}(\Omega_i)}{\sum_q \int_{\Omega_i} \int_t N^{exp}(\theta_\mu, \theta_K, \varphi_h | t, q)\,\mathrm{d}t\,\mathrm{d}\theta_\mu\,\mathrm{d}\theta_K\,\mathrm{d}\varphi_h}. \tag{7.32}$$

This equation is only valid if the angular distribution of the underlying PDF does not depend on the decay time t. As can be seen from Equation (1.37), this is the case if $\Delta\Gamma_s = 0$. Therefore, the sample used to extract the acceptance in this way was generated with this specification. In addition, Equation (7.32) neglects any possible decay-time dependence of the angular acceptance. This assumption is tested in Chapter 10, and an according systematic uncertainty is assigned. Projections of the obtained three dimensional acceptance histogram are shown in Figure 7.18.

7.3.1 Origin of the angular acceptance

The helicity angle of the muons/kaons are defined with respect to B_s^0 flight direction in the center-of-mass system of the $J/\psi\,\phi$ meson. Their relation to momenta distributions in the lab frame is therefore highly unintuitive and always depends on all three defined angles. Thus, it is hardly possible to give a simple explanation of the origin of the shape of the angular acceptance. In order to anyhow understand the different contributions to the shapes observed in Figure 7.18, a large sample of simulated $B_s^0 \to J/\psi\,\phi$ decays with a pure phase space distribution was produced with the RapidSim package [87]. Since the decay products are distributed according to the available phase space, the distributions of $\cos\theta_\mu$, $\cos\theta_K$ and φ_h are expected to be uniformly distributed. This simplifies the determination of the angular acceptance. The sample is a pure generator level simulation, meaning that no

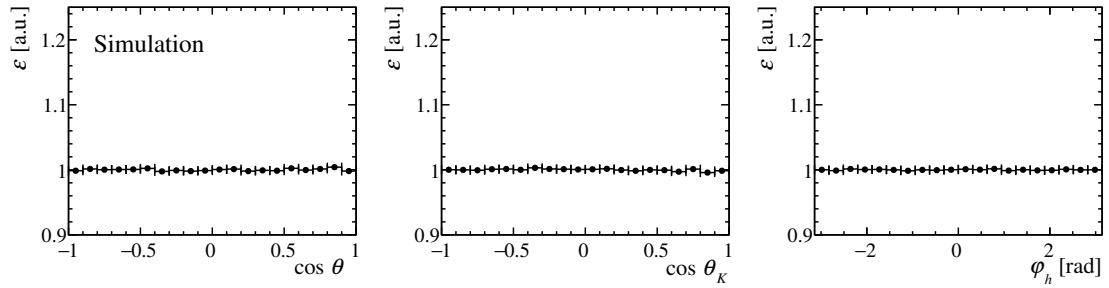


Figure 7.19: Projections of the three dimensional angular acceptance histogram obtained from a generator level phase space simulation of $B_s^0 \to J/\psi\,\phi$ decays.

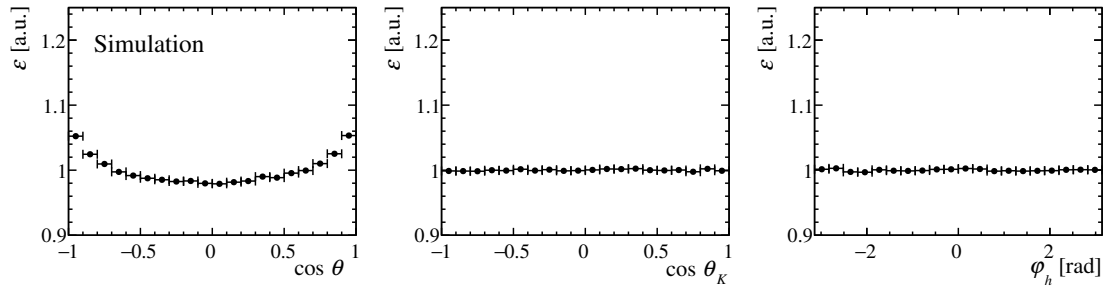


Figure 7.20: Projections of the three dimensional angular acceptance histogram obtained from a generator level phase space simulation of $B_s^0 \to J/\psi\,\phi$ decays. The final state particles are required to lie in the acceptance of the LHCb detector.

detector and therefore no reconstruction and selection is simulated. Thus, the angular acceptance determined from this sample is flat, see Figure 7.19.

As a first step, the final state particles are required to lie inside the acceptance of the LHCb detector. This corresponds to an angle between the beam axis and the final state particle trajectory in the interval $[10, 400]$ mrad. The effect of this requirement is shown in Figure 7.20. While the acceptance projections of $\cos\theta_K$ and φ_h are largely unaffected, there is a significant impact on the acceptance as a function of the helicity angles of the two muons. This is expected since the mass of the J/ψ meson is significantly larger than the mass of two muons, while the kaons originate from the ϕ meson, which is not much heavier than the two kaons combined. This means that for the muons the helicity angle has a much larger impact on their momenta than this is the case for the kaons.

All final state particles of this analysis are required to leave a signal inside the detector after the dipole magnet. This places further momentum dependent cuts on the flight direction of the particles. Choosing only decay candidates where all final state particles fulfill this requirement leads to the angular acceptance shown in Figure 7.21. Again, a similar effect is visible and as previously, the effect is

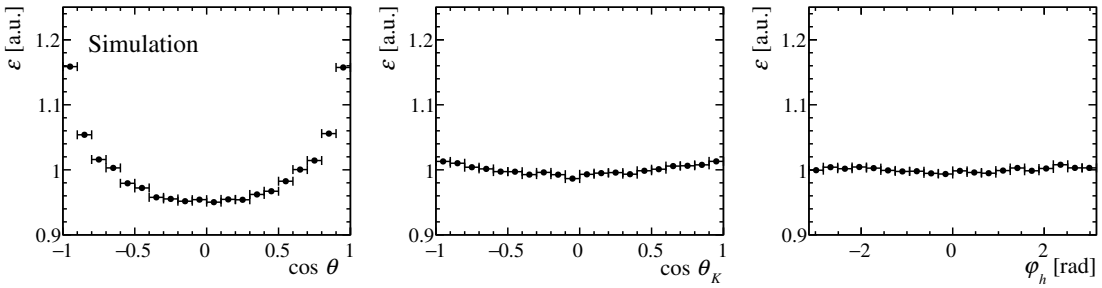


Figure 7.21: Projections of the three dimensional angular acceptance histogram obtained from a generator level phase space simulation of $B_s^0 \to J/\psi\,\phi$ decays. The final state particles are required to lie in the acceptance of the LHCb detector and to not be bent out of the detector by the dipole magnet.

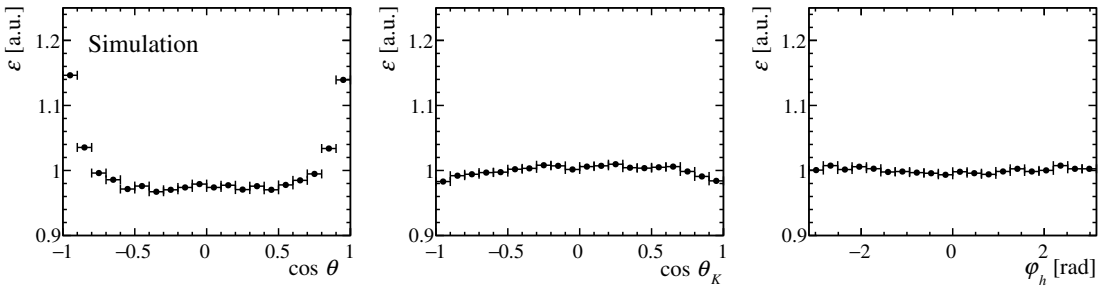


Figure 7.22: Projections of the three dimensional angular acceptance histogram obtained from a generator level phase space simulation of $B_s^0 \to J/\psi\,\phi$ decays. The final state particles are required to lie in the acceptance of the LHCb detector and to not be bent out of the detector by the dipole magnet. In addition the transverse momentum distributions of the final state tracks is weighted to match the fully simulated sample.

much less significant for the kaons. In addition, a very small effect is visible for the angle φ_h.

During the online and offline selection presented in Chapter 5, a range of requirements is placed on the momenta and transverse momenta. Both of these types of criteria are expected to alter the angular acceptance. They are studied separately by weighting the distributions of the generator level simulation to match the respective distribution in the fully simulated sample. Figure 7.22 shows the angular acceptance when applying the weighting in transverse momentum on top of the LHCb acceptance requirements discussed previously. Comparing these acceptances to the ones in Figure 7.21, an opposite effect is visible. While the acceptance requirements especially reject particles with too high rapidity, the selection strategy is typically based on tracks with high transverse momentum, which explains the observed reversed impact on the angular acceptance. Weighting the momentum distributions

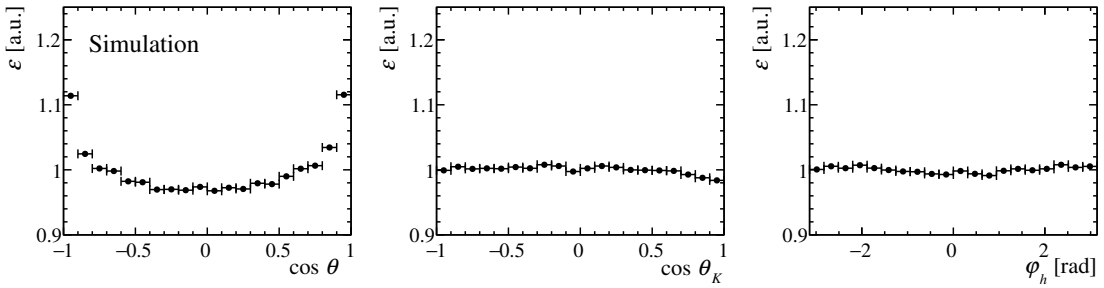


Figure 7.23: Projections of the three dimensional angular acceptance histogram obtained from a generator level simulation of $B_s^0 \to J/\psi\,\phi$ decays. The final state particles are required to lie in the acceptance of the LHCb detector and to not be bent out of the detector by the dipole magnet. In addition the momentum distributions of the final state tracks is weighted to match the fully simulated sample.

instead of the transverse momentum distributions yields the acceptances shown in Figure 7.23. This weighting has a similar effect as the weighting in the transverse momentum.

These different effects on the angular acceptance overlap and give in the end the shape that is observed in the fully simulated $B_s^0 \to J/\psi\,\phi$ sample, see Figure 7.18.

7.3.2 Strategy to determine the acceptance

Normalization weights

In principle, the three dimensional acceptance histogram obtained in the previous section could be directly used for embedding the angular acceptance in the final time-dependent angular PDF. However, there is another way that has a range of benefits with respect to the usage of acceptance histograms. To understand this second strategy, it is useful to write down the logarithm of the likelihood function $\mathcal{L}$ that is later minimized to determine the physics parameters. Showing only the relevant parts and including the angular acceptance $\varepsilon(\Omega)$ it is given by:

$$\begin{aligned}
\log\mathcal{L}(\lambda) &\approx \sum_i \log\left[\mathrm{PDF}(t_i, \Omega_i, q_i|\lambda)\right] \\
&\approx \sum_i \log\left[\frac{\sum_{k=1}^{10} A_k\, h_{k,q_i}(t|\lambda)\, f_k(\Omega)\, \varepsilon(\Omega)}{\int_t \int_\Omega A_k\, h_{k,q_i}(t|\lambda)\, f_k(\Omega)\, \varepsilon(\Omega)\, \mathrm{d}\Omega\, \mathrm{d}t}\right] \\
&= \sum_i \log\left[\frac{\sum_{k=1}^{10} A_k\, h_{k,q_i}(t|\lambda)\, f_k(\Omega)}{\int_t \int_\Omega A_k\, h_{k,q_i}(t|\lambda)\, f_k(\Omega)\, \varepsilon(\Omega)\, \mathrm{d}\Omega\, \mathrm{d}t}\right] + \sum_i \varepsilon(\Omega), \qquad (7.33)
\end{aligned}$$

which is a sum over all the decay candidates i and depends on the physics parameters λ. The last term does not depend on these parameters. It is therefore irrelevant for the shape of the likelihood function and can be neglected. This means that the angular acceptance has only to be taken into account in the normalization of the PDF. Furthermore, the objects $\omega_k = \int_\Omega f_k(\Omega)\varepsilon(\Omega)\,\mathrm{d}\Omega$ are independent of the physics parameters and the properties of the candidate and have therefore to be calculated only once. Since these normalization weights ω_k are the only place where the angular acceptance enters in the maximum likelihood fit, they will be determined in the following and used as a parametrization.

The absolute scale of the efficiency is of no interest, and by using Equation (7.31) the normalization weights can be written as:

$$\begin{aligned}\omega_k &= \int f_k(\Omega)\varepsilon(\Omega)\,\mathrm{d}\Omega \\ &= \int f_k(\Omega)\frac{N^{obs}(\Omega|t,q)}{N^{exp}(\Omega|t,q)}\,\mathrm{d}\Omega. \end{aligned} \tag{7.34}$$

The integration over Ω can be replaced by a Monte Carlo summation integral using fully simulated events:

$$\begin{aligned}\omega_k &= \frac{1}{\tilde{N}_{obs}}\sum_{e\in obs} f_k(\Omega_e)\frac{N^{obs}(\Omega_e|t_e,q_e)}{N^{exp}(\Omega_e|t_e,q_e)}\frac{\int \mathrm{d}\Omega}{N^{obs}(\Omega_e|t_e,q_e)} \\ &= \frac{\int \mathrm{d}\Omega}{\tilde{N}_{obs}}\sum_{e\in obs}\frac{f_k(\Omega_e)}{N^{exp}(\Omega_e|t_e,q_e)}. \end{aligned} \tag{7.35}$$

This is a sum over the simulated decays that are "observed" after all selection steps. $\tilde{N}_{obs}$ is the total number of these "observed" simulated decays. In words, the normalization weights are determined by iterating over the simulated $B_s^0 \to J/\psi\,\phi$ sample and summing the ratio of the respective angular function and the total underlying PDF evaluated for the current candidate. The factor in front of this sum does not depend on k and represents an arbitrary normalization. Typically, the normalization weights are scaled uniformly such that $\omega_1 = 1$.

There are two main advantages of using these normalization weights instead of any analytic description of the angular acceptance based on the three dimensional histogram. First, the normalization weights do not require any specific model for the acceptance and are therefore not subject to any systematic uncertainty arising from the choice of such a model. Furthermore, they do not require a vanishing decay-width splitting $\Delta\Gamma_s$, and also other simulated samples can be added in order to decrease the statistical uncertainty.

Table 7.5: Normalization weights representing the uncorrected angular acceptance for $B_s^0 \to J/\psi\,\phi$ decays.

	2015		2016	
	`Unbiased`	`Biased`	`Unbiased`	`Biased`
ω_1	1 ± 0	1 ± 0	1 ± 0	1 ± 0
ω_2	1.0270 ± 0.0019	1.0291 ± 0.0035	1.02497 ± 0.00065	1.0208 ± 0.0014
ω_3	1.0270 ± 0.0018	1.0283 ± 0.0034	1.02469 ± 0.00064	1.0208 ± 0.0014
ω_4	-0.0020 ± 0.0015	-0.0095 ± 0.0029	-0.00063 ± 0.00052	0.0024 ± 0.0012
ω_5	0.00017 ± 0.00087	0.0039 ± 0.0017	0.00099 ± 0.00031	0.00321 ± 0.00067
ω_6	0.00139 ± 0.00087	0.0020 ± 0.0017	0.00012 ± 0.00031	-0.00018 ± 0.00067
ω_7	1.0082 ± 0.0013	1.0137 ± 0.0024	1.00624 ± 0.00044	1.0113 ± 0.0010
ω_8	-0.0007 ± 0.0011	-0.0033 ± 0.0022	0.00032 ± 0.00040	-0.00003 ± 0.00087
ω_9	0.0007 ± 0.0011	-0.0013 ± 0.0022	0.00024 ± 0.00041	-0.00003 ± 0.00088
ω_{10}	-0.0001 ± 0.0024	-0.0052 ± 0.0045	-0.00104 ± 0.00084	-0.0022 ± 0.0018

As the decay-time acceptance, the angular acceptance is determined separately for the two years of data taking and the two trigger categories. Table 7.5 shows the normalization weights obtained from the available samples of simulated $B_s^0 \to J/\psi\,\phi$ decays.

Correcting for differences between simulation and data

Since the angular acceptance is determined from simulation, it is important to check that this simulation correctly describes the data. As shown in the previous section, there is a direct relation between the angular acceptance and the kinematic distributions of the final state particles. Figure 7.24 compares these distributions for the data and the simulated sample. Clear differences are visible for the muons as well as for the kaons, which could be caused by a wrong simulation of the detector acceptance. Therefore, corrections have to applied to the simulated samples to account for these differences. This is done in terms of a multidimensional weighting[2] according to the differences observed in the final state kinematics, which corrects the differences in the acceptance as a function of the final state kinematics and therefore also the angular acceptance.

However, besides a wrongly simulated angular acceptance, there are other possible contributions to the differences seen in these distributions. The most obvious is a difference in the momentum distribution of the initial B_s^0 mesons. Given a correctly described acceptance, the kinematic distributions would nevertheless differ from data if the B_s^0 mesons have a wrong three dimensional momentum distribution

[2]Throughout this thesis, the boosted weighting technique, see Section 4.2, is employed when not stated otherwise.

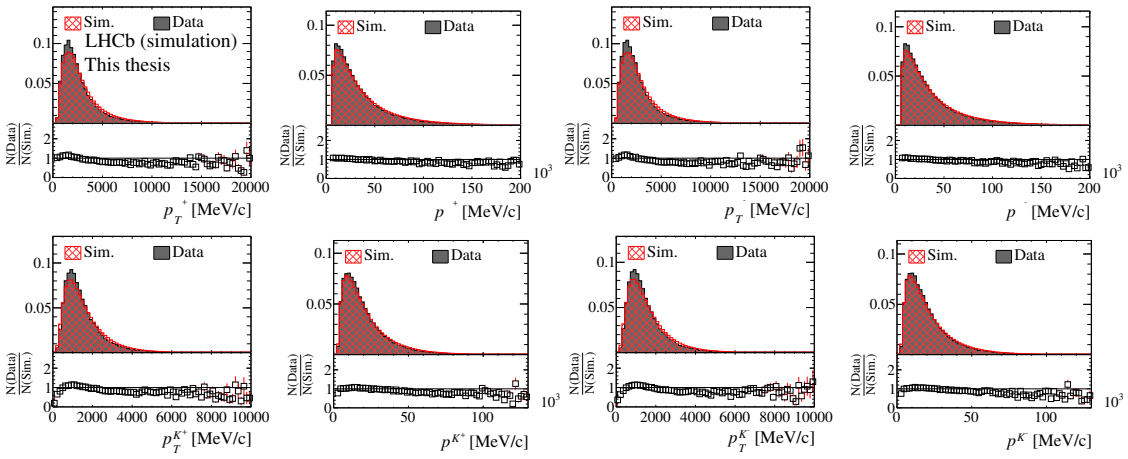


Figure 7.24: Distributions of kinematic variables of the muons and kaons in the background-subtracted $B_s^0 \to J/\psi\,\phi$ data and simulation samples. The 2016 `Unbiased` trigger category is shown.

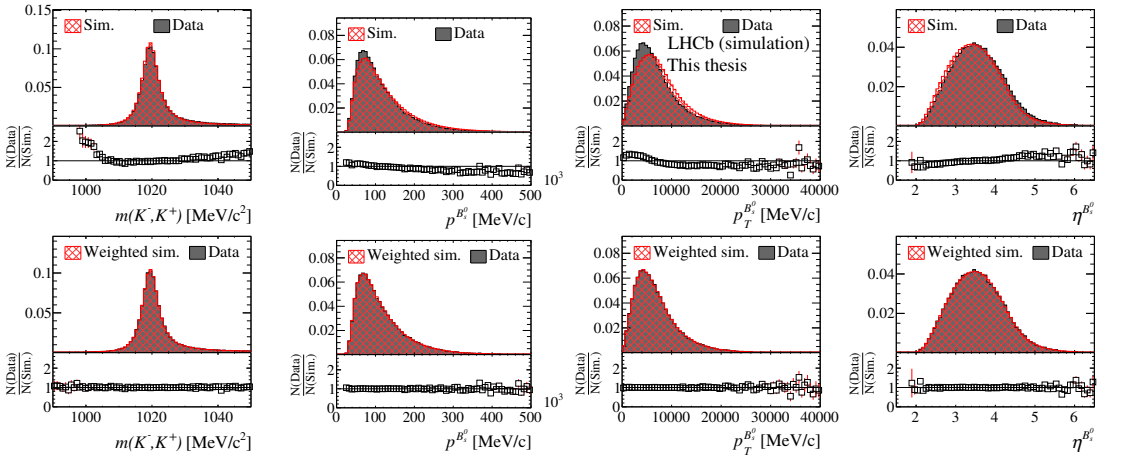


Figure 7.25: Distributions of the three dimensional momentum of B_s^0 mesons and the invariant mass of the two kaons for the background-subtracted $B_s^0 \to J/\psi\,\phi$ data and simulation samples. The bottom row shows the distributions after a weighting of the simulated sample in the variables $p^{B_s^0}$, $p_T^{B_s^0}$ and $m(K^-, K^+)$. The 2016 `Unbiased` trigger category is shown.

in simulation. An additional effect of second order is a possible difference in the simulated two kaon invariant mass, which can also cause differences in the kinematic variables of the two kaons. The top row of Figure 7.25 compares the respective distributions in data and simulation. The differences are corrected by weighting the simulated sample to match data in the momentum of the B_s^0 meson, $p^{B_s^0}$, its transverse momentum, $p_T^{B_s^0}$, and the invariant mass of the two kaons, $m(K^-, K^+)$. The result can be seen in the bottom row of Figure 7.25. All distributions agree.

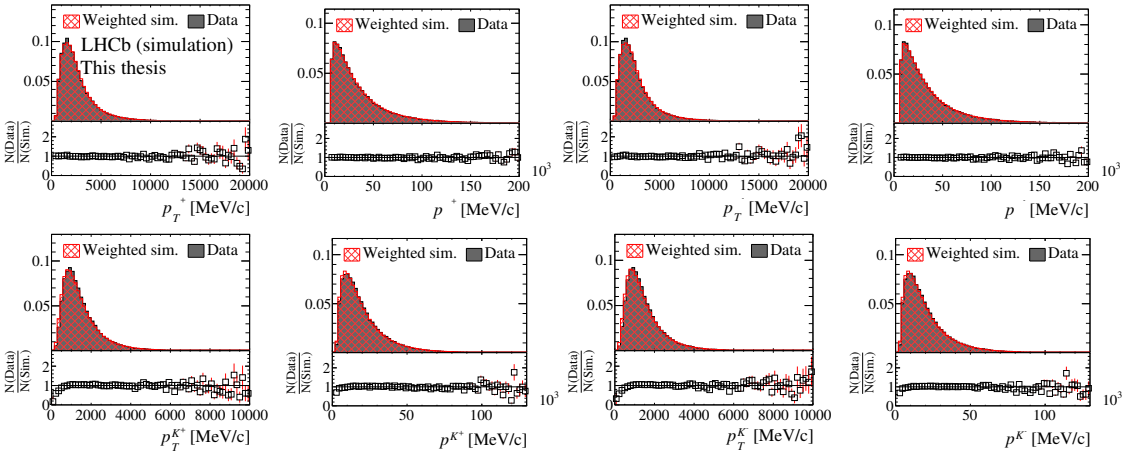


Figure 7.26: Distributions of kinematic variables of the muons and kaons in the background-subtracted $B_s^0 \to J/\psi\,\phi$ data and simulation samples. A weighting of the simulated sample in the variables $p^{B_s^0}$, $p_T^{B_s^0}$ and $m(K^-, K^+)$ is applied. The 2016 `Unbiased` trigger category is shown.

With this correction applied, the final state kinematic distributions, shown in Figure 7.26, agree significantly better. However, there are still small differences visible in the distributions of the momentum and transverse momentum of the kaons. A further effect that can cause these differences is a wrongly modeled angular distribution in the simulation at generator level. Although the samples are generated with physics parameters close to the ones observed in data, there are some second order differences. The most significant one is the absence of the S-wave, i.e. non resonant K^+K^-, component in the simulated samples. To take into account these differences, the simulated sample is weighted to be effectively a sample that is generated with the physics parameters observed in data. This is achieved by giving every simulated decay candidate e a weight ω_e according to:

$$\omega_e = \frac{\mathrm{PDF}(t_e, \Omega_e, q_e|\lambda^{obs})}{\mathrm{PDF}(t_e, \Omega_e, q_e|\lambda^{gen})}, \tag{7.36}$$

where λ^{gen} are the physics parameters with which the simulation was generated and λ^{obs} the respective ones determined from the maximum likelihood fit to data. $\mathrm{PDF}(t, \Omega, q|\lambda)$ is the probability density function at generator level, i.e. without any acceptance and resolution effects. Since these weights depend on the parameters determined in the maximum likelihood fit to data, which depend on the angular normalization weights determined using these weights, the simultaneous determination of the physics parameters from data and of the correct angular normalization weights is done iteratively. Figure 7.27 illustrates this procedure. After the simu-

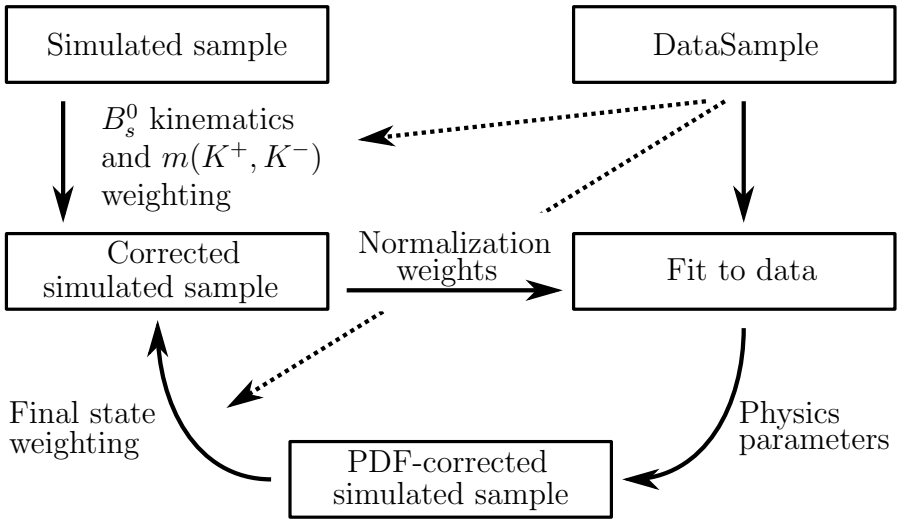


Figure 7.27: Schematic illustration of the iterative procedure used to correct the angular acceptance obtained from simulated $B_s^0 \to J/\psi\,\phi$ samples.

lated sample is corrected for the differences in the B_s^0 meson kinematics and the $m(K^+K^-)$ distribution, a set of normalization weights is determined. These weights are used to represent the angular acceptance in a full time-dependent angular fit to the data sample. The obtained physics parameters allow then a weighting of the simulated sample according to Equation (7.36). Finally, the corrected simulated sample is then weighted to match data in some final state kinematic distributions. A new set of normalization weights is calculated from this weighted simulated sample. This procedure is repeated until the extracted normalization weights do not change anymore. While the weighting of the simulated sample is always done separately for the two years of data taking and trigger categories, the fit to data is performed simultaneously to all samples.

Motivated by the differences between data and simulation observed in Figure 7.26, the weighting of the final state kinematics during the iterative procedure was chosen to cover the four dimensional momentum and transverse momentum distribution of the two kaons. Figure 7.28 shows the kinematic distributions of the two kaons for data and simulation without and with the weights obtained in the last iteration of the correction procedure. A good agreement is achieved. The angular normalization weights for the 2016-`Unbiased` sample, obtained at each step of the correction procedure, are shown in Table 7.6. After five iterations the convergence is achieved. As the decay-time acceptance, the angular acceptance is determined separately for the two years of data taking and the two trigger categories. Table 7.7 shows the

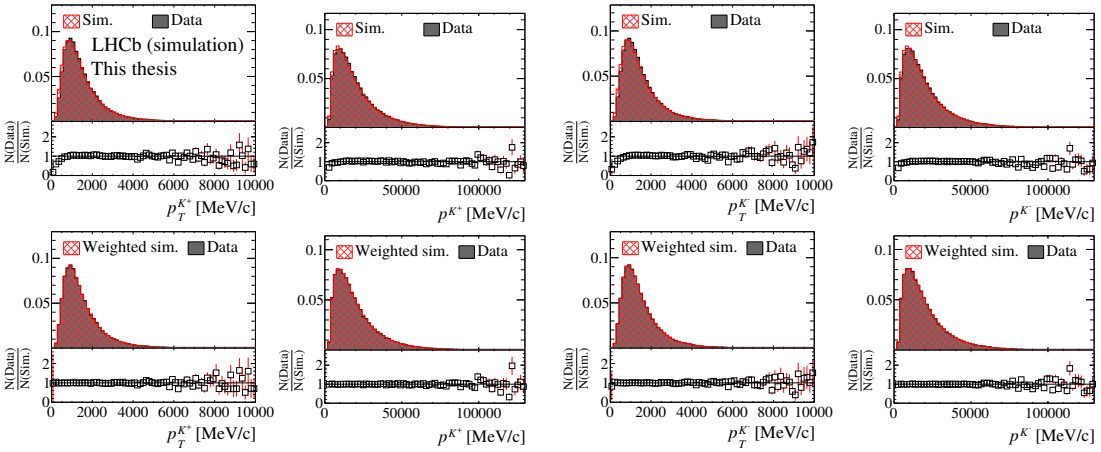


Figure 7.28: Distributions of kinematic variables of the kaons in the background-subtracted $B_s^0 \to J/\psi\,\phi$ data and simulation samples. In addition to the weighting of the simulated sample in the variables $p^{B_s^0}$, $p_T^{B_s^0}$ and $m(K^-, K^+)$, the weights of the last iterative weighting step are applied. The 2016 `Unbiased` trigger category is shown.

Table 7.6: Normalization weights representing the angular acceptance for $B_s^0 \to J/\psi\,\phi$ decays at different steps of the correction procedure. The numbers for the 2016 `Unbiased` category are shown.

Step	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Uncorrected	1.0	1.02497	1.02469	−0.00063	0.00099	0.00012	1.00624	0.00032	0.00024	−0.00104
B_s^0 and m(K^-K^+)	1.0	1.02614	1.02582	−0.00084	0.00078	0.00023	1.00648	0.00045	0.00016	−0.00074
Iteration 1	1.0	1.03660	1.03632	−0.00080	0.00035	0.00025	1.00941	0.00008	0.00010	−0.00384
Iteration 2	1.0	1.03754	1.03729	−0.00078	0.00030	0.00023	1.00999	0.00004	0.00011	−0.00355
Iteration 3	1.0	1.03785	1.03762	−0.00079	0.00027	0.00023	1.01021	0.00005	0.00011	−0.00329
Iteration 4	1.0	1.03787	1.03765	−0.00079	0.00026	0.00023	1.01022	0.00004	0.00010	−0.00362
Iteration 5	1.0	1.03788	1.03765	−0.00079	0.00026	0.00023	1.01022	0.00004	0.00010	−0.00362
Iteration 6	1.0	1.03788	1.03765	−0.00079	0.00026	0.00023	1.01022	0.00004	0.00010	−0.00362

normalization weights obtained from the available samples of simulated $B_s^0 \to J/\psi\,\phi$ decays after the iterative correction procedure.

Table 7.7: Normalization weights representing the angular acceptance for $B_s^0 \to J/\psi\,\phi$ decays.

	2015		2016	
	`Unbiased`	`Biased`	`Unbiased`	`Biased`
ω_1	1 ± 0	1 ± 0	1 ± 0	1 ± 0
ω_2	1.0434 ± 0.0020	1.0463 ± 0.0039	1.03788 ± 0.00070	1.0336 ± 0.0015
ω_3	1.0442 ± 0.0020	1.0445 ± 0.0038	1.03765 ± 0.00069	1.0336 ± 0.0015
ω_4	-0.0026 ± 0.0016	-0.0105 ± 0.0032	-0.00079 ± 0.00054	0.0028 ± 0.0013
ω_5	-0.00142 ± 0.00094	0.0037 ± 0.0018	0.00026 ± 0.00033	0.00298 ± 0.00074
ω_6	0.00139 ± 0.00093	0.0023 ± 0.0018	0.00023 ± 0.00033	-0.00020 ± 0.00072
ω_7	1.0156 ± 0.0014	1.0262 ± 0.0027	1.01022 ± 0.00047	1.0196 ± 0.0011
ω_8	-0.0014 ± 0.0012	-0.0045 ± 0.0024	0.00004 ± 0.00042	0.00019 ± 0.00094
ω_9	0.0006 ± 0.0012	-0.0007 ± 0.0024	0.00010 ± 0.00043	0.00019 ± 0.00094
ω_{10}	-0.0171 ± 0.0026	-0.0348 ± 0.0050	-0.00362 ± 0.00089	0.0057 ± 0.0019

7.3.3 Embedding in the probability density function

As described in Section 7.3.2, the normalization weights model the effect of the angular acceptance on the maximum likelihood fit in its entirety. Therefore, the PDF, Equation (7.29), has to be changed only slightly:

$$\begin{aligned}
&\mathrm{PDF}_{y,c}(t,\Omega|q^{os},q^{ss},\eta^{os},\eta^{ss},\delta_t) = \qquad (7.37)\\
&\frac{1}{N^{\eta^{os},\eta^{ss},\delta_t}_{q^{os},q^{ss},y,c}} \sum_{k=1}^{10} A_k\, f_k(\Omega)\, \varepsilon_{y,c}(t)\\
&\quad \{[(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss})))\cdot h_{k,+1}(t)\\
&\quad + (1-q^{os}(1-2\bar{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss})))\cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\},
\end{aligned}$$

with the normalization:

$$\begin{aligned}
&N^{\eta^{os},\eta^{ss},\delta_t}_{q^{os},q^{ss},y,c} = \qquad (7.38)\\
&\int_{t=0.3\,\mathrm{ps}}^{15\,\mathrm{ps}} \sum_{k=1}^{10} A_k\, \omega_k^{y,c}\, \varepsilon_{y,c}(t)\\
&\quad \{[(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss})))\cdot h_{k,+1}(t)\\
&\quad + (1-q^{os}(1-2\bar{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss})))\cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\}\,\mathrm{d}t,
\end{aligned}$$

where $\omega_k^{y,c}$ is the normalization weight for the respective angular function and data category.

8 The fit procedure

8.1 Including the S-wave component

As described in Section 1.5.2, the presence of the scalar component in the two kaon system leads to the necessity of an additional amplitude and further interference terms in the time- and angular-dependent decay rates of B_s^0 and $\overline{B}_s^0$ mesons, see Equations (1.53) and (1.54). However, this S-wave component furthermore allows to resolve a twofold ambiguity in these equations. The transformation

$$(\phi_s, \Delta\Gamma_s, \delta_\parallel, \delta_\perp, \delta_S, \delta_0) \rightarrow (\pi - \phi_s, -\Delta\Gamma_s, -\delta_\parallel, \pi - \delta_\perp, -\delta_S, -\delta_0) \tag{8.1}$$

does not change the measurable decay rate and would therefore lead to two indistinguishable results for the physics parameters extracted in the maximum likelihood fit.

This is not longer the case if the different two kaon mass (m_{KK}) dependencies of the S-wave and P-wave amplitudes are taken into account. The four amplitudes are given by:

$$\begin{aligned} A_0(m_{KK}) &= A_0\,\rho(m_{KK}) \qquad & A_\parallel(m_{KK}) &= A_\parallel\,\rho(m_{KK}) \\ A_\perp(m_{KK}) &= A_\perp\,\rho(m_{KK}) \qquad & A_S(m_{KK}) &= A_S\,\sigma(m_{KK}), \end{aligned} \tag{8.2}$$

where the line shapes ρ and σ parametrize the mass dependence of the P-wave and S-wave component, respectively.

Instead of including these functions in the final fit, which would require to also model the m_{KK} distribution, the time-dependent angular analysis is simultaneously performed in the six bins of m_{KK} that are defined in Section 5.3. For a given mass of the two kaon system, the amplitudes appear in the differential decay rate as $A_i(m_{KK})A_j^*(m_{KK})$, $A_i(m_{KK})A_S^*(m_{KK})$ and $A_S(m_{KK})A_S^*(m_{KK})$ with $i, j \in [0, \parallel, \perp]$.

Averaging these amplitude combinations over a given m_{KK} range $[m_1, m_2]$ yields then the following contributions:

$$\begin{aligned}
\int_{m_1}^{m_2} A_i\, A_j^*\, \rho(m_{KK})\, \rho^*(m_{KK})\, \mathrm{d}m_{KK} &= |A_i A_j| \\
\int_{m_1}^{m_2} A_S\, A_S^*\, \sigma(m_{KK})\, \sigma^*(m_{KK})\, \mathrm{d}m_{KK} &= |A_S|^2 \\
\int_{m_1}^{m_2} A_i\, A_S^*\, \rho(m_{KK})\, \sigma^*(m_{KK})\, \mathrm{d}m_{KK} &= |A_i A_S|\, C_{SP}\, e^{-i\theta_{SP}}.
\end{aligned} \tag{8.3}$$

Here, the functions ρ and σ are chosen to be normalized in the respective $m_{K^-K^+}$ region, and the factor $C_{SP}e^{-i\theta_{SP}}$ accounts for the interference between S-wave and P-wave. While the phase θ_{SP} gets absorbed by the strong phase δ_S, the real factor C_{SP} has to be calculated in order to correctly determine the fraction of the S-wave component. For this calculation, the P-wave is described by a Breit-Wigner function with the mean and width of the ϕ meson [3], and the S-wave is assumed to be a f_0 resonance, which is described by a Flatté distribution [88]. Both distributions are modified according to Ref. [89]. Alternative descriptions of the S-wave component are considered, and a systematic uncertainty is assigned in Chapter 10.

Given the line shapes of the two components, the C_{SP} factors can be calculated according to:

$$C_{SP}e^{-i\theta_{SP}} = \frac{\int_{m_1}^{m_2} \rho(m_{KK})\, \sigma^*(m_{KK})\, \mathrm{d}m_{KK}}{\sqrt{\int_{m_1}^{m_2} |\rho(m_{KK})|^2\, \mathrm{d}m_{KK} \int_{m_1}^{m_2} |\sigma(m_{KK})|^2\, \mathrm{d}m_{KK}}}. \tag{8.4}$$

Figure 5.10 shows the m_{KK} distribution and the six bins for which separate such correction factors have to be calculated. Since the mass of the two kaon system can only be reconstructed with finite precision, decay candidates can migrate between the bins. This can be taken into account in terms of efficiencies $\varepsilon^j(m_{KK})$, for every m_{KK} bin j. Figure 8.1 shows these efficiencies as a function of the true value of m_{KK}. With this, Equation (8.4) gets:

$$C_{SP}^j e^{-i\theta_{SP}^j} = \frac{\int \rho(m_{KK})\, \sigma^*(m_{KK})\, \varepsilon^j(m_{KK})\, \mathrm{d}m_{KK}}{\sqrt{\int |\rho(m_{KK})|^2\, \varepsilon^j(m_{KK})\, \mathrm{d}m_{KK} \int |\sigma(m_{KK})|^2\, \varepsilon^j(m_{KK})\, \mathrm{d}m_{KK}}}. \tag{8.5}$$

Table 8.1 shows the C_{SP} factors for the six m_{KK} bins. Since no differences between the two years of data taking and the two trigger categories are observed, the same factors are used in all categories. The integration of these factors into the maximum likelihood fit is described in the next section where the final PDF is shown.

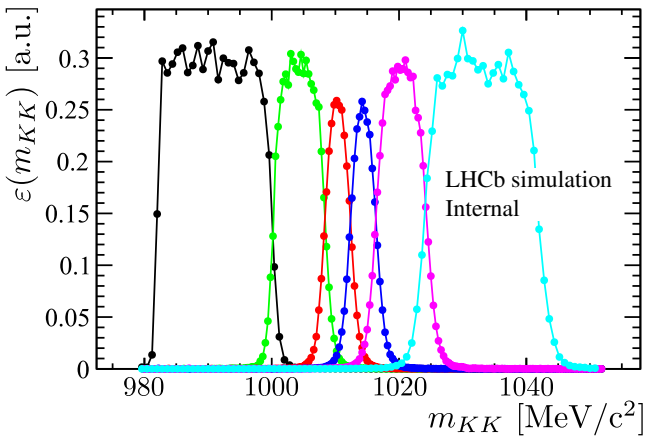


Figure 8.1: Efficiency of the division into the six bins of the reconstructed two kaon mass as a function of the true m_{KK} value. The values are obtained from a simulated $B_s^0 \to J/\psi\,\phi$ sample. Figure taken from Ref. [65].

Table 8.1: C_{SP} factors in the six m_{KK} bins. They account for the interference between S-wave and P-wave line shapes. Numbers are taken from Ref. [63].

	m_{KK} bin					
	1	2	3	4	5	6
C_{SP}	0.8463	0.8756	0.8478	0.8833	0.9415	0.9756

8.2 The complete probability density function

A simultaneous fit to the four data categories created by the splitting in the two trigger categories, $c \in$[`Unbiased`, `Biased`], and in the two years of data taking, $y \in$ [2015, 2016], is preformed. This is necessary, because each of these sample is associated to a different set of decay-time and angular acceptances. In addition, each sample is split in six bins of m_{KK}, $j \in$ [1,2,3,4,5,6], in order to resolve the twofold ambiguity of the differential decay rate. For each of these samples, a time- and angular-dependent PDF is defined that depends on five additional conditional observables to account for the tagging (q^{os}, q^{ss}, η^{os}, η^{ss}) and estimated decay-time resolution (δ_t):

$$\begin{aligned}
&\mathrm{PDF}^{j}_{y,c}(t,\Omega|q^{os},q^{ss},\eta^{os},\eta^{ss},\delta_t) = \qquad (8.6)\\
&\frac{1}{N^{\eta^{os},\eta^{ss},\delta_t,j}_{q^{os},q^{ss},y,c}} \sum_{k=1}^{10} \tilde{A}^{j}_{k}\, f_k(\Omega)\, \varepsilon_{y,c}(t)\\
&\quad \{[(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss}))) \cdot h_{k,+1}(t)\\
&\quad + (1-q^{os}(1-2\bar{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss}))) \cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\}.
\end{aligned}$$

The normalization factor $N^{\eta^{os},\eta^{ss},\delta_t,j}_{q^{os},q^{ss},y,c}$ is given by:

$$N^{\eta^{os},\eta^{ss},\delta_t,j}_{q^{os},q^{ss},y,c} = \tag{8.7}$$
$$\int_{t=0.3\,\mathrm{ps}}^{15\,\mathrm{ps}} \sum_{k=1}^{10} \tilde{A}^j_k\, \varepsilon_{y,c}(t)\, \omega^k_{y,c}$$
$$\{[(1+q^{os}(1-2\omega^{os}(\eta_{os})))(1+q^{ss}(1-2\omega^{ss}(\eta_{ss})))\cdot h_{k,+1}(t)$$
$$+ (1-q^{os}(1-2\bar{\omega}^{os}(\eta_{os})))(1-q^{ss}(1-2\bar{\omega}^{ss}(\eta_{ss})))\cdot h_{k,-1}(t)] \otimes G(t|\sigma_{eff}(\delta_t))\}\,\mathrm{d}t.$$

The amplitudes $\tilde{A}^j_k$ are given by A_k for $k \leq 7$ and by $C^j_{SP}\,A_k$ for the S-wave interference terms, $k \in [8,9,10]$. The amplitudes A_k and the angular and time-dependent functions f_k and $h_{k,\pm1}$ are defined in Tables 1.2 and 1.3. The absolute amplitudes squared $A^2_{\perp,0,\parallel,S}$ are parametrized by a S-wave fraction, $F^j_S = A^2_S$, for every bin in $m(K^-K^+)$ and values for $|A_\perp|^2$ and $|A_0|^2$ representing the respective fraction of the resonant component. The parallel component is fixed by $A^2_\perp + A^2_0 + A^2_\parallel = 1$. As discussed in Chapter 7, the decay-time acceptance and resolution, and the angular acceptance are taken into account by the efficiency functions $\varepsilon_{y,c}(t)$, the calibrated Gaussian decay-time resolution $G(t|\sigma_{eff}(\delta_t))$, and the angular acceptance normalization weights $\omega^k_{y,c}$. The raw mistag probabilities $\eta_{os/ss}$ are transformed by the calibration functions $\overset{(-)}{\omega}{}^{os/ss}$. All physics parameters besides the S-wave fractions and phases are shared between all years of data taking, trigger categories and m_{KK} bins. In case of the S-wave related parameters, different values are allowed for the different m_{KK} bins.

During the simultaneous fit, the tagging calibration parameters are constrained, see Section 4.3, to the values obtained in Chapter 6. Again, these parameters are shared between all data categories.

8.3 Validation of the fit procedure

In order to validate the implementation and general behavior of the PDF and the maximum likelihood fit, toy studies are performed. A sample of $B^0_s \to J/\psi\,\phi$ decays is generated, the maximum likelihood fit is performed, and the results are compared to the values used to generate the decays. The aim is to include all acceptance, resolution and tagging effects as they are present in real data. For every trigger category, year of data taking and m_{KK} bin, separate samples of $B^0_s \to J/\psi\,\phi$ decays are generated with the parameters measured in this analysis, see next chapter. The number of generated decays corresponds to the effective sample size observed

in data for the respective category. For the generation of every single decay the following steps are performed:

1. A random initial flavour of the $B_s^0/\overline{B}_s^0$ meson is generated.

2. Given this initial flavour, the respective PDF without any acceptance and resolution effect is used to randomly generate a value for the decay time and the three angles in the helicity basis.

3. A random value for the estimated decay-time uncertainty δ_t is generated from the distribution observed in background-subtracted data, and a reconstructed decay time is generated by smearing the generated true value with a Gaussian of the width $\sigma_{eff}(\delta_t)$. The same calibration function $\sigma_{eff}(\delta_t)$ as in data is used.

4. The decay is randomly rejected according to its reconstructed decay time and the respective decay-time acceptance function. This includes the lower and upper decay-time cuts at 0.3 ps and 15 ps, respectively.

5. Similarly, the decay is randomly rejected according to its generated helicity angles and an angular acceptance function that is obtained by fitting a multidimensional polynomial to the acceptance histogram obtained from the fully corrected simulated sample, see Section 7.3.2.

6. The conditional tagging variables are randomly generated using the tagging efficiencies and mistag distributions observed in data.

7. A set of tagging calibration parameters is generated from the measured values and their correlations and uncertainties. With these parameters, the generated initial B_s^0/B_s^0 flavour and the generated mistag probabilities, the assigned tagging decisions are randomly changed.

After the generation of this toy sample, a maximum likelihood fit is performed based on the exact same PDFs that are also used in the fit to real data. The only difference is that the angular acceptance normalization weights are analytically calculated for the parametrized acceptance shape. This procedure is repeated 5000 times. Figures 8.2 and 8.3 show the distribution of the observed deviations from the generated value in units of the statistical uncertainty for every parameter. Positive values mean that the result from the fit is larger than the input value. While most of the main physics parameters show an unbiased behavior, especially the S-wave related parameters feature strongly asymmetric pull distributions. This behavior is

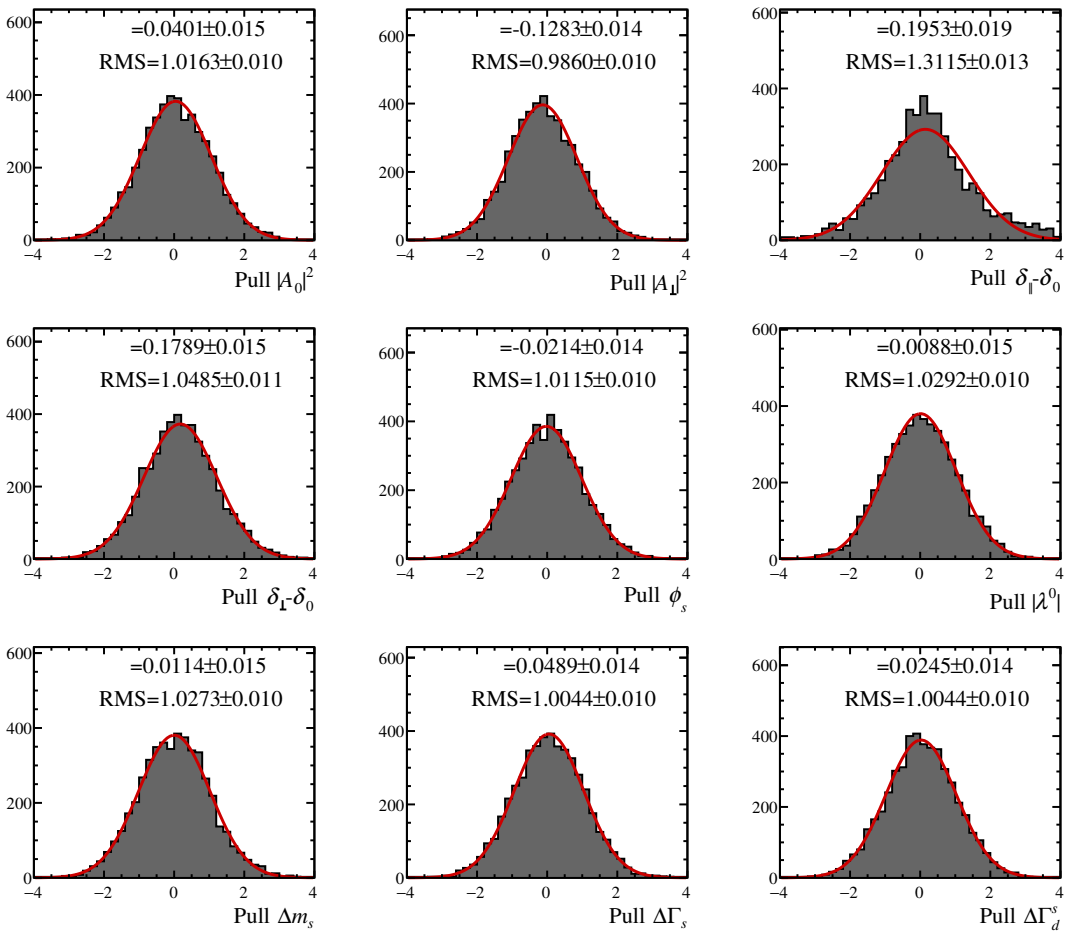


Figure 8.2: Pull distributions of the main physics parameters observed in the toy studies. The mean and the root mean square are given. For illustration, a Gaussian function is fitted to the distribution.

studied later and systematic uncertainties are assigned on all parameters based on the observed biases.

Impact of the effective resolution

As describe in Section 7.1.2, a triple Gaussian resolution function is transformed into a single Gaussian with an effective width. While this simplifies the fitting procedure and the estimation of systematic uncertainties, it might as well bias the extracted physics parameters. A dedicated toy study is performed to access possible systematic uncertainties arising from this strategy. In analogy to the procedure described in the previous section, toy data sets are generated. The only difference is that instead of a single Gaussian, a triple Gaussian resolution function with the same effective dilution is chose. The fractions and width ratios between the three

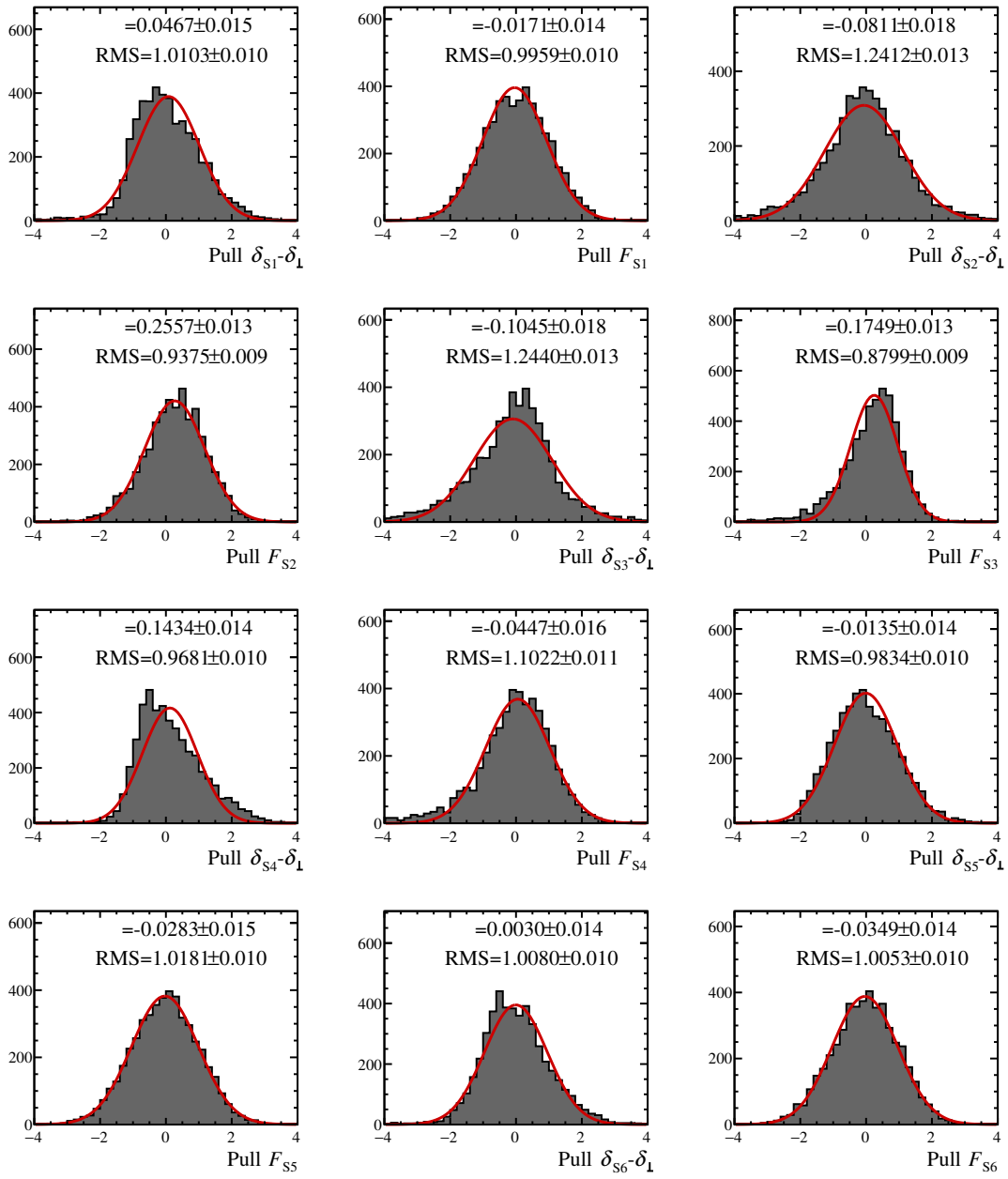


Figure 8.3: Pull distributions of the S-wave parameters observed in the toy studies. The mean and the root mean square are given. For illustration, a Gaussian function is fitted to the distribution.

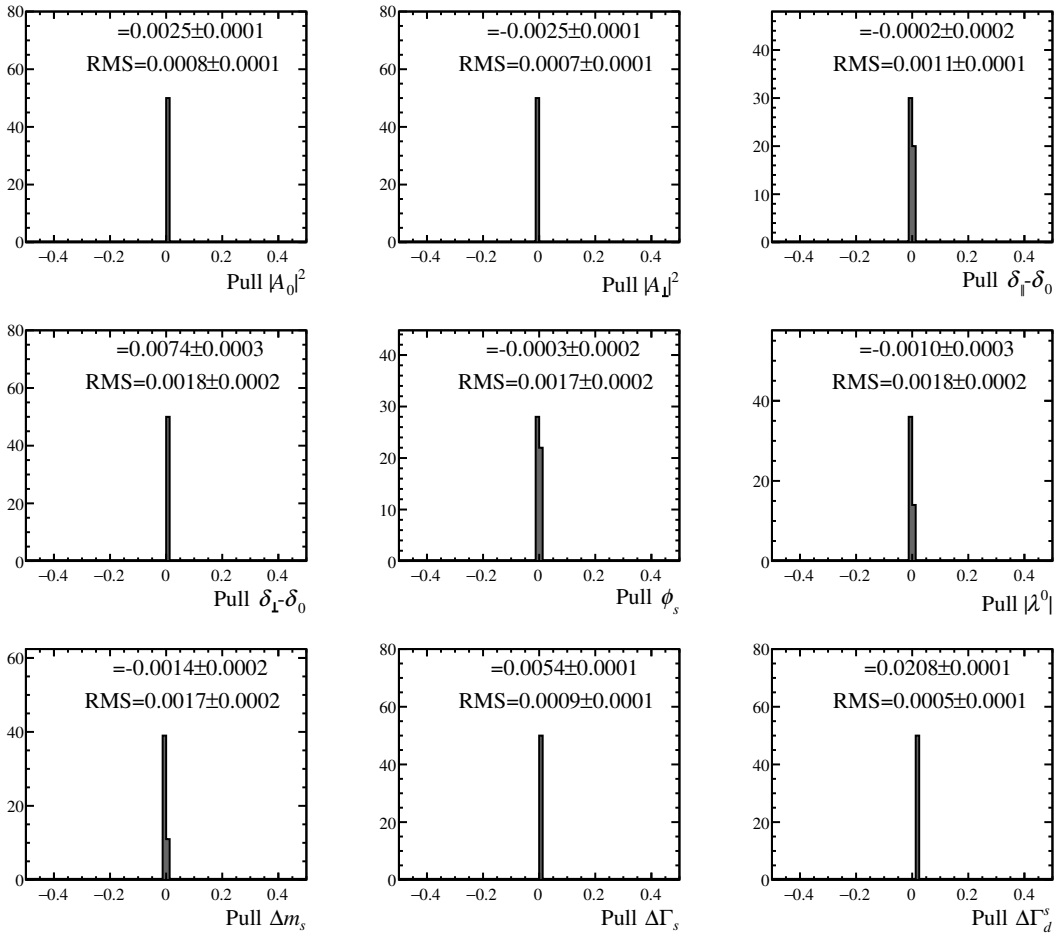


Figure 8.4: Pull distributions of the main physics parameters observed in the toy studies for the single Gaussian resolution. The difference between the fit result obtained with single Gaussian and triple Gaussian resolution model are shown in units of the statistical uncertainty. The mean and the root mean square are given.

Gaussian contributions are fixed to values that are observed in the fit to the prompt data sample during the decay-time resolution determination.

These generated samples are once fitted with PDFs that contain the full triple Gaussian resolution and once with the PDFs that contain only the single Gaussian effective resolution. Figure 8.4 shows the distribution of the observed differences between these two fits for every parameter in units of the statistical uncertainty of the fit with the single Gaussian resolution. Positive values mean that the result from the fit with triple Gaussian resolution is larger. In Appendix F, the respective distributions for the S-wave related parameters are shown. The only parameter with a sizable bias is the decay-width difference $\Delta\Gamma_d^s$. This is caused by the lower decay time cut ($t = 0.3\,\text{ps}$) being relatively close to the physical boundary at $t = 0\,\text{ps}$.

Neglecting the long tails in the decay-time resolution alters the expected decay-time distribution at very low decay times. However, the bias is small in comparison to the statistical uncertainty, which justifies the usage of the single Gaussian resolution.

Besides a possible shift in the central value of the fit results, also the estimated uncertainty could be biased by the choice of a single Gaussian resolution. Using the same toy study, no significant differences in the estimated uncertainties are observed.

9 Results of the fit

In this section, the results of the maximum likelihood fit described in Chapter 8 are presented. Table 9.1 shows the estimated values for all physics parameters together with their statistical uncertainties. The correlations between these parameters are given in Appendix G.2. If appropriate, asymmetric uncertainties are quoted. These asymmetric uncertainties partially account for asymmetric shapes of the likelihood projections along the respective parameters. Especially the strong phases are affected by this. The complete picture is visible when looking at Figures G.1 and G.2 in Appendix G.1. There, the likelihood scans for all parameters are shown.

Table 9.1: Fit results for the freely floating parameters of the maximum likelihood fit to the decay-time and angular distributions.

Parameter	Fit result
$\phi_s\,[\mathrm{rad}]$	-0.083 ± 0.041
$\lvert\lambda\rvert$	1.012 ± 0.016
$\Delta\Gamma_s\,[\mathrm{ps}^{-1}]$	$0.0773\pm^{0.0076}_{0.0077}$
$\Delta m_s\,[\mathrm{ps}^{-1}]$	$17.702\pm^{0.057}_{0.059}$
$\Delta\Gamma^s_d\,[\mathrm{ps}^{-1}]$	-0.0040 ± 0.0023
$\lvert A_\perp\rvert^2$	0.2455 ± 0.0040
$\lvert A_0\rvert^2$	0.5189 ± 0.0029
$\delta_\parallel-\delta_0\,[\mathrm{rad}]$	$3.060\pm^{0.084}_{0.073}$
$\delta_\perp-\delta_0\,[\mathrm{rad}]$	2.64 ± 0.13
F_{S1}	0.491 ± 0.043
F_{S2}	$0.0406\pm^{0.0081}_{0.0075}$
F_{S3}	$0.0044\pm^{0.0029}_{0.0018}$
F_{S4}	$0.0069\pm^{0.0061}_{0.0046}$
F_{S5}	0.073 ± 0.013
F_{S6}	$0.151\pm^{0.019}_{0.018}$
$\delta_{S1}-\delta_\perp\,[\mathrm{rad}]$	$2.21\pm^{0.17}_{0.20}$
$\delta_{S2}-\delta_\perp\,[\mathrm{rad}]$	1.55 ± 0.29
$\delta_{S3}-\delta_\perp\,[\mathrm{rad}]$	$1.07\pm^{0.49}_{0.34}$
$\delta_{S4}-\delta_\perp\,[\mathrm{rad}]$	$-0.28\pm^{0.16}_{0.27}$
$\delta_{S5}-\delta_\perp\,[\mathrm{rad}]$	$-0.536\pm^{0.090}_{0.103}$
$\delta_{S6}-\delta_\perp\,[\mathrm{rad}]$	$-1.10\pm^{0.13}_{0.16}$

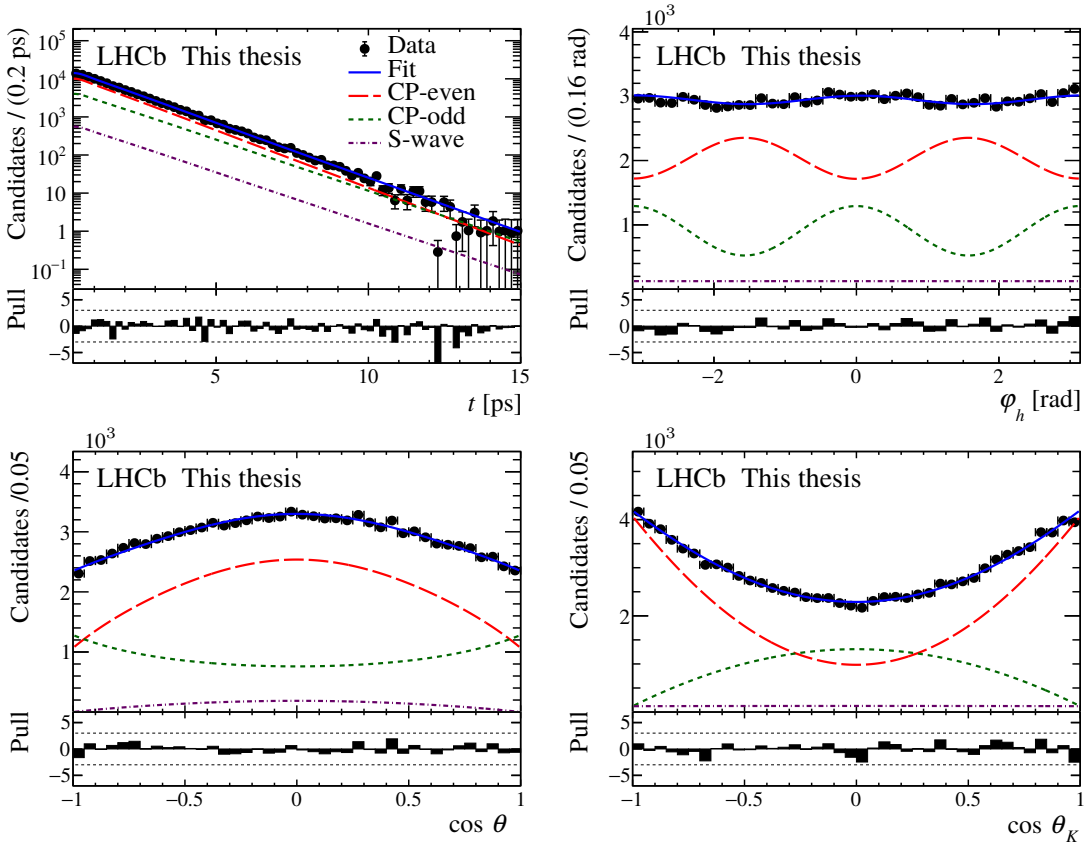


Figure 9.1: Projections of the decay-time and angular fit. Besides the background-subtracted data (black) and the overall fit projections (blue), also the CP-even (red), CP-odd (green) and the S-wave (purple) components are shown.

During a likelihood scan, the maximum likelihood is calculated as a function of the parameter of interest. For every value of this parameter, the likelihood is maximized by varying all other parameters. The main physics parameters of interest, namely the CP and lifetime observables, have a nearly Gaussian shape, which justifies the interpretation of the quoted uncertainties as the 68% confidence interval. In contrast, many of the S-wave related amplitudes and phases show strongly asymmetric shapes, which corresponds to the asymmetric pull distributions observed in the toy study presented in Section 8.3.

For illustration, Figure 9.1 shows the projection of the final fit together with the background-subtracted data distribution for the decay time and the helicity angles. In these plots, neither the data nor the fit projections are split according to the estimated B_s^0 flavour. An oscillation as in Figure 1.4 would be anyhow not visible because of the small value of ϕ_s and the dilution due to the mistag probability, the decay time resolution and the presence of different CP components in the final state. A better way to visualize an oscillating difference between the decay rates of B_s^0

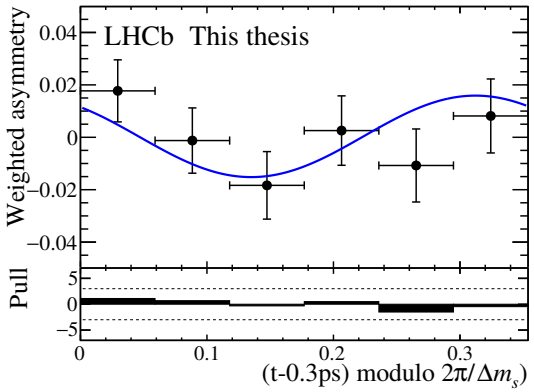


Figure 9.2: Asymmetry between B_s^0 and $\overline{B}_s^0$ tagged decays as a function of the decay time. The full decay-time range is projected to one oscillation period. The decay candidates are weighted to enhance the visible asymmetry. The fit projection is shown in blue, and the weighted data is shown in black.

and $\overline{B}_s^0$ mesons, is to project the asymmetry over the full decay-time range to a single oscillation period, see Figure 9.2. In Appendix G.4, a detailed description of the procedure to create this plot is given.

For completeness, in Appendix G.3, the above presented results are compared to the official LHCb publication presented in Ref. [63].

Polarization-dependent CP violation

As described in Section 1.5.7, the data can also be fitted with a model that allows polarization-dependent CP violation. The according PDF is obtained from the nominal one developed so far by replacing the time-dependent terms $h_k(t)$ in Equation (3.1) with the ones defined in Appendix A. The maximum likelihood fit is repeated, and Wilks' theorem [90] is employed to test the hypothesis of polarization-dependent CP violation. Given that there is no polarity dependent CP violation, the difference in logarithmic likelihood of the two fit results should be distributed according to a χ^2 distribution, with the number of degrees of freedom given by the difference of the number of floating parameters of the two models. The observed difference in likelihood is 6.1, which corresponds, with 6 additional parameters in the case of the polarization-dependent model, to a p-value of 0.41. This means that the data is perfectly consistent with the hypothesis of no polarization-dependent CP violation.

9.1 Consistency checks

A common way to ensure that all effects that could bias the extracted parameters are correctly taken into account, is to repeat the fit in subsets of the data set. Assuming an appropriate description of all detector effects and a working background subtraction, the results obtained from these subsets should be consistent. The data set is split according to a wide range of criteria. For the following splittings only the final fit is repeated:

- A splitting according to the year of data taking: Table 9.2
- A splitting in the two trigger categories `Unbiased` and `Biased`: Table 9.3
- A splitting according to which tagging algorithm provides a decision (SS and OS, only OS, only SS): Table 9.4
- A splitting according to the polarity of the dipole magnet: Table 9.5
- A splitting according to the number of primary vertices in the respective event: Table 9.6

For criteria that are expected to heavily affect the decay-time or angular acceptance, a larger part of the analysis chain is repeated. For the following splittings, a separate background subtraction and acceptance determination is performed:

- A splitting according to the transverse momentum of the B_s^0 candidate: Table 9.8
- A splitting according to the pseudo rapidity of the B_s^0 candidate: Table 9.7
- A splitting according to the estimated decay-time uncertainty of the B_s^0 candidate: Table 9.9

The respective tables show the main parameters obtained in each of the categories, and a p-value for their consistency is quoted. When interpreting these p-values, two aspects have to be taken into account. They are calculated assuming Gaussian uncertainties and can therefore underestimate the consistency for parameters that show large asymmetries in their likelihood profiles, see Appendix G.1. Especially for the relatively small data sample collected in 2015 and the parameter $\delta_{\parallel} - \delta_0$ this is the case. Therefore, the small p-value observed in Table 9.2 is clearly underestimated. Furthermore, the uncertainties of the parameters do not include any systematic uncertainty. While for most parameters and sources of systematic uncertainties these

Table 9.2: Main parameters of the final fit when splitting in the years of data taking. A p-value for consistency is given for each of them.

Parameter	2015	2016	p-value
ϕ_s [rad]	-0.06 ± 0.11	-0.211 ± 0.044	0.202
$\lvert A_0 \rvert^2$	0.5226 ± 0.0083	0.5185 ± 0.0031	0.643
$\lvert A_\perp \rvert^2$	0.242 ± 0.012	0.2449 ± 0.0042	0.842
$\delta_\parallel - \delta_0$ [rad]	3.53 ± 0.12	3.090 ± 0.074	0.002
$\delta_\perp - \delta_0$ [rad]	2.99 ± 0.43	2.63 ± 0.13	0.423
$\lvert\lambda\rvert$	1.003 ± 0.055	1.015 ± 0.016	0.829
$\Delta\Gamma_s$ [ps^{-1}]	0.147 ± 0.023	0.1663 ± 0.0081	0.431
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0003 ± 0.0069	-0.0045 ± 0.0025	0.568
Δm_s [ps^{-1}]	17.57 ± 0.13	17.715 ± 0.057	0.308

Table 9.3: Main parameters of the final fit when splitting in the two trigger categories. A p-value for consistency is given for each of them.

Parameter	`Unbiased`	`Biased`	p-value
ϕ_s [rad]	-0.176 ± 0.046	-0.200 ± 0.089	0.812
$\lvert A_0 \rvert^2$	0.5181 ± 0.0032	0.5223 ± 0.0066	0.566
$\lvert A_\perp \rvert^2$	0.2464 ± 0.0045	0.2424 ± 0.0088	0.686
$\delta_\parallel - \delta_0$ [rad]	3.042 ± 0.096	3.12 ± 0.12	0.612
$\delta_\perp - \delta_0$ [rad]	2.55 ± 0.15	2.87 ± 0.26	0.280
$\lvert\lambda\rvert$	1.009 ± 0.016	1.014 ± 0.048	0.915
$\Delta\Gamma_s$ [ps^{-1}]	0.1643 ± 0.0086	0.158 ± 0.017	0.756
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0039 ± 0.0026	-0.0038 ± 0.0051	0.980
Δm_s [ps^{-1}]	17.649 ± 0.070	17.79 ± 0.10	0.276

are strongly correlated between the subsamples and would therefore not improve the consistency, the uncertainty related to the finite size of the samples used in the decay-time acceptance determination has to be treated differently. In case such acceptances are determined separately for each category, this systematic uncertainty is completely uncorrelated. In Section 10.2.2 the systematic uncertainty is evaluated for the overall samples and turns out to be halve the statistical uncertainty for the decay-width difference $\Delta\Gamma_d^s$. Assuming a similar uncertainty in each of the bins of the transverse momentum, see Table 9.8, the respective p-value increases to roughly 4%.

In total, no significant dependencies are observed. This gives further confidence in the extracted nominal fit result and in the correct treatment of detector acceptance and resolution effects.

Table 9.4: Main parameters of the final fit when splitting according to which tagging algorithm provides a decision. A p-value for consistency is given for each of them.

Parameter	OS and SS	Only OS	Only SS	p-value
ϕ_s [rad]	-0.174 ± 0.057	-0.19 ± 0.11	-0.226 ± 0.096	0.898
$\lvert A_0 \rvert^2$	0.5207 ± 0.0059	0.5028 ± 0.0087	0.5202 ± 0.0045	0.171
$\lvert A_\perp \rvert^2$	0.2465 ± 0.0086	0.261 ± 0.012	0.2414 ± 0.0063	0.330
$\delta_\parallel - \delta_0$ [rad]	2.89 ± 0.11	3.18 ± 0.19	3.05 ± 0.11	0.356
$\delta_\perp - \delta_0$ [rad]	2.38 ± 0.17	2.85 ± 0.26	2.97 ± 0.19	0.056
$\lvert \lambda \rvert$	0.999 ± 0.023	0.934 ± 0.049	1.076 ± 0.031	0.031
$\Delta\Gamma_s$ [ps^{-1}]	0.169 ± 0.015	0.147 ± 0.022	0.156 ± 0.012	0.690
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0033 ± 0.0047	-0.0011 ± 0.0070	-0.0060 ± 0.0036	0.792
Δm_s [ps^{-1}]	17.669 ± 0.081	17.79 ± 0.11	17.74 ± 0.14	0.661

Table 9.5: Main parameters of the final fit when splitting according to the polarity of the dipole magnet. A p-value for consistency is given for each of them.

Parameter	Mag. up	Mag. down	p-value
ϕ_s [rad]	-0.237 ± 0.056	-0.140 ± 0.060	0.238
$\lvert A_0 \rvert^2$	0.5245 ± 0.0040	0.5132 ± 0.0042	0.053
$\lvert A_\perp \rvert^2$	0.2403 ± 0.0054	0.2505 ± 0.0059	0.208
$\delta_\parallel - \delta_0$ [rad]	3.11 ± 0.12	3.02 ± 0.10	0.563
$\delta_\perp - \delta_0$ [rad]	2.53 ± 0.17	2.64 ± 0.18	0.642
$\lvert \lambda \rvert$	1.028 ± 0.019	0.993 ± 0.024	0.257
$\Delta\Gamma_s$ [ps^{-1}]	0.174 ± 0.010	0.153 ± 0.011	0.168
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0076 ± 0.0032	-0.0001 ± 0.0034	0.108
Δm_s [ps^{-1}]	17.605 ± 0.077	17.732 ± 0.080	0.248

Table 9.6: Main parameters of the final fit when splitting according to the number of primary vertices. A p-value for consistency is given for each of them.

Parameter	#PV = 1	#PV = 2	#PV > 2	p-value
ϕ_s [rad]	-0.242 ± 0.057	-0.100 ± 0.071	-0.173 ± 0.096	0.299
$\lvert A_0 \rvert^2$	0.5212 ± 0.0044	0.5154 ± 0.0048	0.5214 ± 0.0068	0.624
$\lvert A_\perp \rvert^2$	0.2462 ± 0.0060	0.2457 ± 0.0067	0.2429 ± 0.0093	0.957
$\delta_\parallel - \delta_0$ [rad]	3.04 ± 0.10	3.10 ± 0.26	3.09 ± 0.13	0.952
$\delta_\perp - \delta_0$ [rad]	2.41 ± 0.19	2.70 ± 0.21	2.95 ± 0.26	0.239
$\lvert \lambda \rvert$	1.016 ± 0.019	1.031 ± 0.027	0.977 ± 0.034	0.449
$\Delta\Gamma_s$ [ps^{-1}]	0.172 ± 0.011	0.156 ± 0.013	0.159 ± 0.018	0.639
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0046 ± 0.0035	0.0022 ± 0.0039	-0.0126 ± 0.0053	0.073
Δm_s [ps^{-1}]	17.558 ± 0.079	17.798 ± 0.079	17.88 ± 0.13	0.035

Table 9.7: Main parameters of the final fit when splitting according to the pseudo rapidity of the B_s^0 candidate. A p-value for consistency is given for each of them.

Parameter	$\eta < 3.3$	$\eta \in [3.3, 3.9]$	$\eta > 3.9$	p-value
ϕ_s [rad]	-0.170 ± 0.059	-0.135 ± 0.070	-0.317 ± 0.091	0.260
$\lvert A_0 \rvert^2$	0.5133 ± 0.0046	0.5214 ± 0.0051	0.5290 ± 0.0056	0.091
$\lvert A_\perp \rvert^2$	0.2400 ± 0.0064	0.2466 ± 0.0070	0.2480 ± 0.0075	0.672
$\delta_\parallel - \delta_0$ [rad]	3.11 ± 0.11	3.02 ± 0.12	3.13 ± 0.17	0.835
$\delta_\perp - \delta_0$ [rad]	2.52 ± 0.17	2.61 ± 0.21	2.69 ± 0.27	0.843
$\lvert \lambda \rvert$	1.012 ± 0.020	1.026 ± 0.028	0.982 ± 0.026	0.500
$\Delta\Gamma_s$ [ps^{-1}]	0.171 ± 0.012	0.156 ± 0.013	0.165 ± 0.014	0.729
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0048 ± 0.0037	-0.0006 ± 0.0041	-0.0058 ± 0.0042	0.644
Δm_s [ps^{-1}]	17.653 ± 0.070	17.734 ± 0.093	17.40 ± 0.12	0.086

Table 9.8: Main parameters of the final fit when splitting according to the transverse momentum of the of the B_s^0 candidate. A p-value for consistency is given for each of them.

Parameter	$p_T < 4\,\text{GeV/c}$	$p_T \in [4, 7]\,\text{GeV/c}$	$p_T > 7\,\text{GeV/c}$	p-value
ϕ_s [rad]	-0.16 ± 0.10	-0.205 ± 0.076	-0.157 ± 0.054	0.876
$\lvert A_0 \rvert^2$	0.5278 ± 0.0057	0.5160 ± 0.0051	0.5143 ± 0.0045	0.147
$\lvert A_\perp \rvert^2$	0.2506 ± 0.0076	0.2490 ± 0.0071	0.2396 ± 0.0065	0.470
$\delta_\parallel - \delta_0$ [rad]	3.12 ± 0.13	3.08 ± 0.12	2.909 ± 0.091	0.320
$\delta_\perp - \delta_0$ [rad]	3.00 ± 0.26	2.49 ± 0.23	2.72 ± 0.16	0.340
$\lvert \lambda \rvert$	1.045 ± 0.045	0.988 ± 0.025	1.024 ± 0.023	0.432
$\Delta\Gamma_s$ [ps^{-1}]	0.169 ± 0.014	0.154 ± 0.014	0.163 ± 0.012	0.755
$\Delta\Gamma_d^s$ [ps^{-1}]	0.0061 ± 0.0044	-0.0108 ± 0.0041	-0.0010 ± 0.0037	0.018
Δm_s [ps^{-1}]	18.10 ± 0.19	17.67 ± 0.11	17.740 ± 0.064	0.133

Table 9.9: Main parameters of the final fit when splitting according to the estimated decay-time uncertainty of the B_s^0 candidate. A p-value for consistency is given for each of them.

Parameter	$\delta_t < 4\,\text{GeV/c}$	$\delta_t \in [4, 7]\,\text{GeV/c}$	$\delta_t > 7\,\text{GeV/c}$	p-value
ϕ_s [rad]	-0.084 ± 0.059	-0.211 ± 0.063	-0.312 ± 0.091	0.080
$\lvert A_0 \rvert^2$	0.5167 ± 0.0056	0.5266 ± 0.0048	0.5157 ± 0.0050	0.218
$\lvert A_\perp \rvert^2$	0.2409 ± 0.0080	0.2428 ± 0.0065	0.2510 ± 0.0069	0.570
$\delta_\parallel - \delta_0$ [rad]	2.876 ± 0.092	3.11 ± 0.10	3.15 ± 0.19	0.158
$\delta_\perp - \delta_0$ [rad]	2.54 ± 0.19	2.46 ± 0.18	2.39 ± 0.25	0.878
$\lvert \lambda \rvert$	0.997 ± 0.024	0.997 ± 0.023	1.031 ± 0.019	0.417
$\Delta\Gamma_s$ [ps^{-1}]	0.176 ± 0.017	0.165 ± 0.013	0.146 ± 0.012	0.302
$\Delta\Gamma_d^s$ [ps^{-1}]	-0.0079 ± 0.0053	-0.0065 ± 0.0040	-0.0003 ± 0.0035	0.359
Δm_s [ps^{-1}]	17.641 ± 0.083	17.686 ± 0.075	17.49 ± 0.12	0.353

10 Systematic uncertainties

In this chapter, for each of the components of the analysis systematic studies are presented. If necessary, systematic uncertainties are estimated. The exact numbers assigned as uncertainties can be found in Tables 10.1 and 10.2 at the end of this chapter.

10.1 Selection and background subtraction

This section shortly describes the systematic uncertainties assigned to take into account possible biases arising from certain assumptions or choices during the selection and background subtraction procedure.

Peaking backgrounds

As presented in Section 5.2, two peaking background components are considered. Dedicated vetos are applied and the number of remaining misidentified decays is estimated. While this number is negligible in the case of $B^0 \to J/\psi\, K^{*0}$ decays, the $\Lambda_b^0 \to J/\psi\, pK$ component has to be further suppressed by injecting simulated Λ_b^0 decays with negative weights. The yield of these injected simulated candidates is varied by $\pm 1\sigma$ of the estimated $\Lambda_b^0 \to J/\psi\, pK$ yield in data, and the final fit is repeated. Since no significant changes in the extracted parameters are observed, no systematic uncertainty is assigned.

Mass shape

For the correct subtraction of the combinatorial background with the ${}_s\mathcal{P}lot$ technique, a correct description of the mass distribution observed for the $B_s^0 \to J/\psi\, \phi$ candidates is needed, see Section 5.3. Two systematic studies are performed to test the impact of a potentially wrong model. The first study tests the intrinsic uncertainty of the mass shape parameters on the background subtraction procedure. Before applying the ${}_s\mathcal{P}lot$ technique, these parameters are randomly varied according to their covariance matrix, and the background subtraction and the

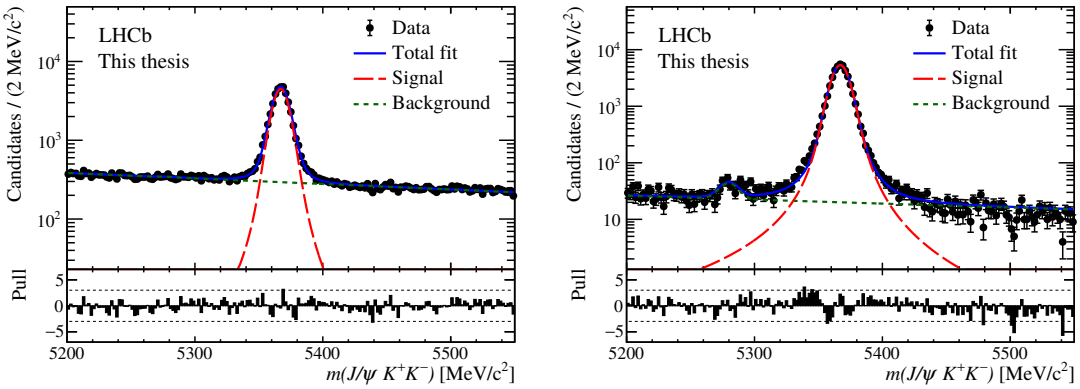


Figure 10.1: Mass distribution of the $B_s^0 \to J/\psi\,\phi$ candidates in two bins of the decay time. On the left and the right only candidates with a decay time smaller or larger than 1.2 ps are used, respectively. The overlaid fit projections corresponds to fits where all parameters except the relative fractions are fixed to those of the fit to the overall data set. Only candidates collected in 2016 and falling in the `Unbiased` trigger category are shown.

final fit are repeated. No significant changes in the physics parameters are observed, and therefore no systematic uncertainty is assigned. The second systematic study is the usage of an alternative scaling function of the estimated `per-event mass uncertainty`. Instead of a second order polynomial scaling function, a liner transformation, $\sigma = s_1\,\delta_M$, between the estimated and real resolution is employed. The background subtraction and the final fit are repeated and the observed differences in the physics parameters are assigned as systematic uncertainties, see Tables 10.1 and 10.2 at the end of this chapter.

Mass factorization

The suppression of the combinatorial background with the ${}_s\mathcal{P}lot$ technique, see Section 5.3, relies on the assumption of no correlation between the mass of the background or signal candidates and the variables that are used in the final fit. The correlation between the signal mass shape and $\cos\theta_\mu$ is already taken into account by the usage of the estimated mass uncertainty. To test for other correlations, the data sample is split in two bins of the respective variable. Each sample is fit with a mass model with all parameters except the yields fixed to the values obtained by the nominal fit. Only for the decay time a small dependence of the slope of the combinatorial background component is found. This is shown in Figure 10.1. To estimate the bias due to this correlation, the mass fit is repeated in bins of the decay time. Three, four and five equally populated bins are tested. For each of these binning schemes, new weights for the background subtraction are calculated

separately in every bin. The final fit to the whole sample with these new weights is repeated. For every parameter, the maximal observed deviation is taken as systematic uncertainty. These numbers are shown in Tables 10.1 and 10.2 at the end of this chapter.

Multiple candidates

The chance to have two $B_s^0 \to J/\psi\,\phi$ decays within one event is extremely small. However, in the data set of this analysis roughly 1.5% of the events contain more than one candidate. In most cases these candidates share the same J/ψ candidate but differ at least in one of the kaons. In principle these fake multiple candidates are distributed like combinatorial background, but they can also peak below the true B_s^0 mass if the wrongly chosen kaon is built from a clone track of the true kaon track. These clone tracks share a large fraction of the detector hits with the underlying true track and are therefore likely to be reconstructed with the same momentum.

In order to estimate the effect of these sort of multiple candidates, two candidates within one event are removed from the selection if it is likely that they are clones. Candidates for which all final state tracks form a respective opening angle of less than 5 mrad fall in this category. The background removal with the ${}_s\mathcal{P}lot$ technique, the acceptance determination and the final fit is repeated. Any difference with respect to the nominal result is assigned as systematic uncertainty for the respective parameter. These numbers are shown in Tables 10.1 and 10.2 at the end of this chapter.

10.2 Acceptance and resolution effects

10.2.1 Decay-time resolution

The calibration parameters of the estimated `per-event decay-time uncertainty` δ_t, determined in Section 7.1.2, come with uncertainties reflecting the limited size of the sample of prompt fake $B_s^0 \to J/\psi\,\phi$ candidates. The final fit is repeated several times with varying these parameters according to their uncertainties and correlations. No effect on the fit parameters is observed. The same is true when using a quadratic instead of a linear calibration function.

However, there is a significantly more severe possible source of systematic uncertainty that has to be considered. The strategy presented in Section 7.1.2 relies on the portability of the calibration from the prompt sample to the sample of

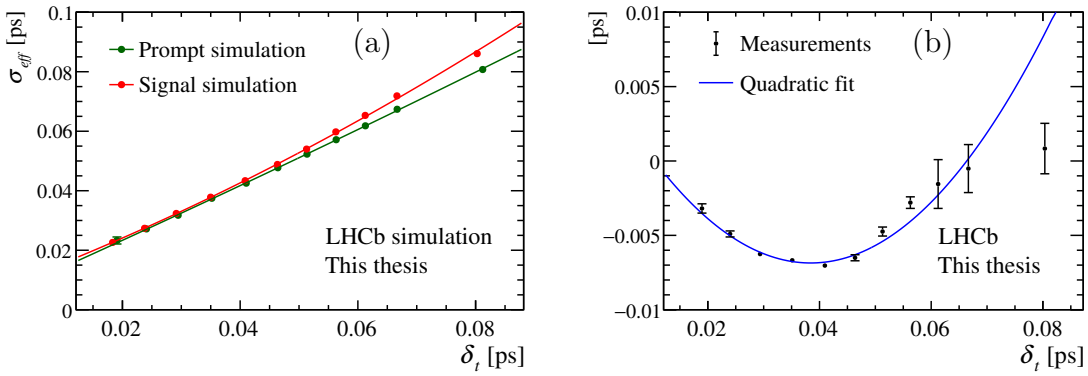


Figure 10.2: Calibration curves of the estimated `per-event decay-time uncertainty` for the simulated prompt fake and the simulated signal $B_s^0 \to J/\psi\,\phi$ samples (a). Mean parameter of the decay time resolution function as a function of the estimated `per-event decay-time uncertainty` as it is observed in the prompt fake $B_s^0 \to J/\psi\,\phi$ data sample (b).

true $B_s^0 \to J/\psi\,\phi$ decays. This assumption is studied on simulation. Figure 10.2a shows the calibration obtained with the same strategy for simulated samples of prompt fake and signal $B_s^0 \to J/\psi\,\phi$ decays. A clear difference is visible between the calibration curves. The signal sample calibration clearly deviates from the linear behavior observed in the prompt data and simulation sample. The ratio of the two calibration curves in simulation is used to scale the effective resolutions measured in the prompt fake $B_s^0 \to J/\psi\,\phi$ data. Then, a quadratic calibration function is fit to these new values. The shifts in the parameters of the final fit when using this new decay-time resolution calibration is assigned as systematic uncertainty.

In the nominal fit, the decay-time resolution is assumed to be centered around zero. However, in the sample of fake prompt $B_s^0 \to J/\psi\,\phi$ candidates, a significant bias is observed. Figure 10.2b shows the mean parameter of the resolution model in bins of δ_t. A clear dependence is visible, which is not present in simulation and was understood[1] to be an effect of misalignments between different regions of the VLEO. It is included in the final fit by a quadratic dependence of the mean value of the resolution function on the estimated decay time uncertainty δ_t. Again, the difference of this fit result with respect to the nominal one is assigned as systematic uncertainty.

An additional systematic uncertainty is assigned to account for the unknown fraction of wrongly assigned primary vertices in the prompt and signal data sample.

[1] The discussion of this is out of the scope of this thesis and was part of a study performed within the LHCb collaboration.

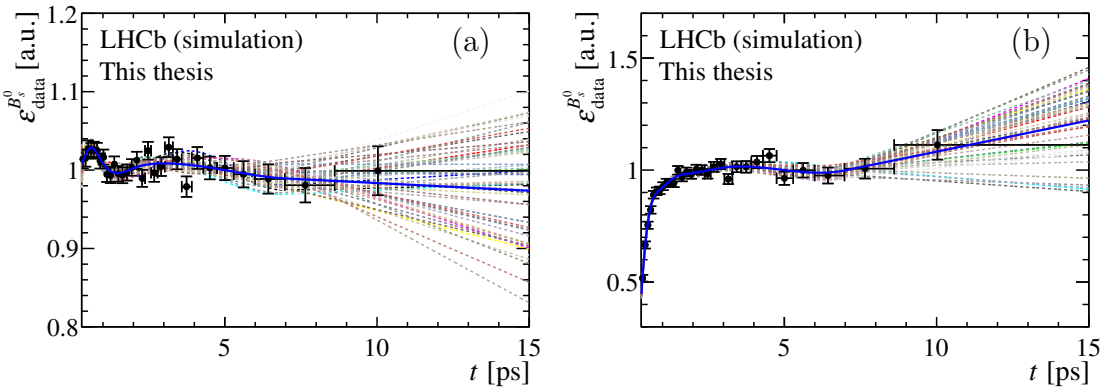


Figure 10.3: Decay-time acceptance variations for the 2016 `Unbiased` (a) and `Biased` (b) categories. The solid blue line indicates the nominal acceptance, while the other curves are obtained by randomly varying the spline coefficients according to their covariance matrix.

The fraction is fixed to either 0 or 1.5% in the calibration fits to the prompt sample. Depending on the δ_t bin, this corresponds to an up to ten times increased fraction. The two obtained calibrations replace the nominal one in the final fit and the maximal observed difference is assigned as systematic uncertainty.

The numbers for all these systematic uncertainties are shown in Tables 10.1 and 10.2 at the end of this chapter.

10.2.2 Decay-time acceptance

An obvious source of systematic uncertainty related to the decay-time acceptance is the limited size of the $B^0 \to J/\psi K^{*0}$ data sample. Although its size is significantly larger than the one of the $B_s^0 \to J/\psi \phi$ data sample, a sizable effect is expected. In addition, also the simulation samples that are used to correct for the difference between the two decay channels contribute in the same way. The combined uncertainty of all three samples is represented by the covariance matrices of the spline coefficients that are determined in the decay-time acceptance fit, see Section 7.2. To translate this uncertainty to a systematic uncertainty on the parameters of the time-dependent angular fit, the spline coefficients are randomly varied according to their covariance matrices and the fit is repeated. Figure 10.3 shows the nominal acceptance together with a range of such random variations for the data collected in 2016.

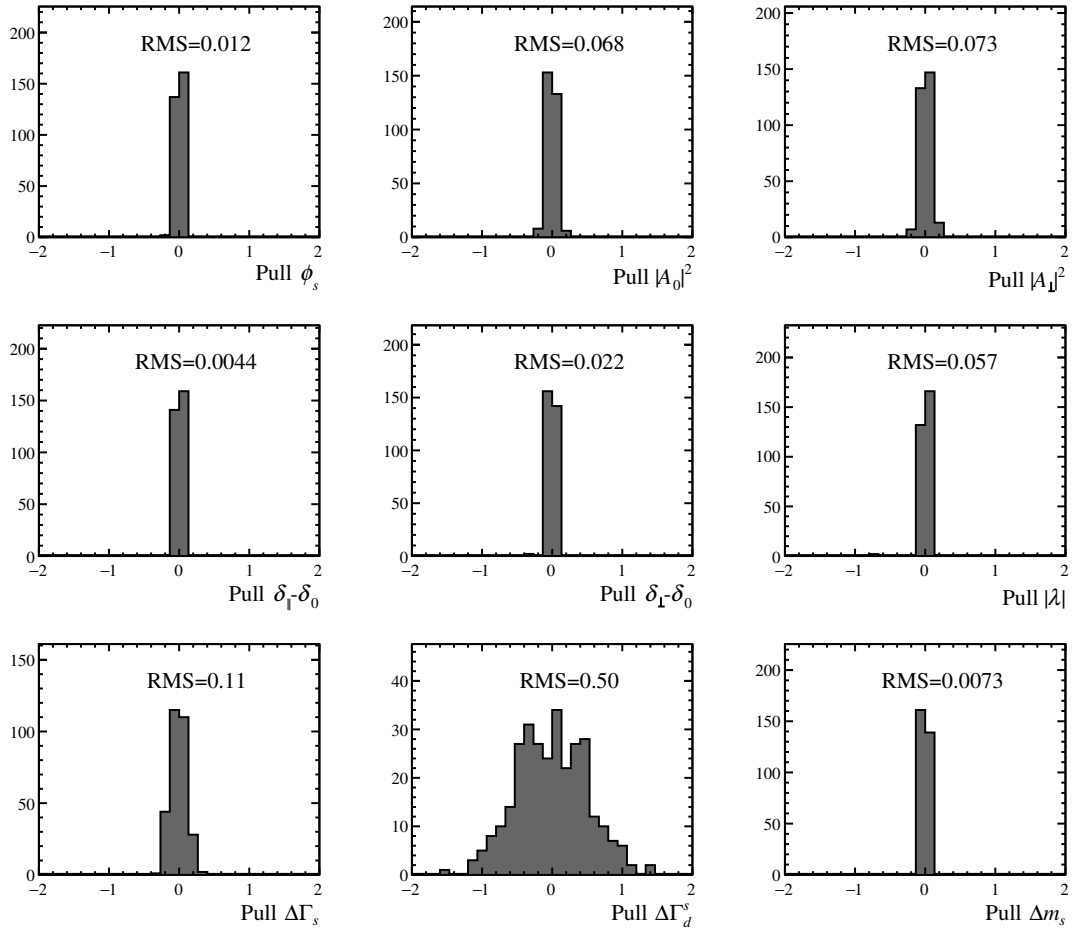


Figure 10.4: Spread of the main physics parameters in units of their statistical uncertainty when varying the decay-time acceptance within its statistical uncertainty.

In total, 300 fits with randomly generated splines are performed. The distributions of the deviations of the main final parameters from the nominal result are shown in Figure 10.4 in units of the statistical uncertainty of the nominal fit. As expected, the decay-width difference $\Delta\Gamma_d^s$ is most sensitive to the spline variations, but also the decay-width splitting $\Delta\Gamma^s$ shows a sizable spread. For all parameters that are subject to a sizable variation, the root mean square (RMS) of the respective distribution is assigned as a systematic uncertainty.

A further systematic uncertainty is assigned by increasing the number of knots of the splines. Thereby, any structures that are not correctly modeled with the nominal spline configuration should be picked up. Figure 10.5 shows the nominal acceptance and the acceptance with knots at [0.3, 0.43, 0.58, 0.74, 0.91, 1.12, 1.35, 1.63, 1.96, 2.40, 3.01, 4.06, 9] ps for the 2016-`Unbiased` data category. Using these

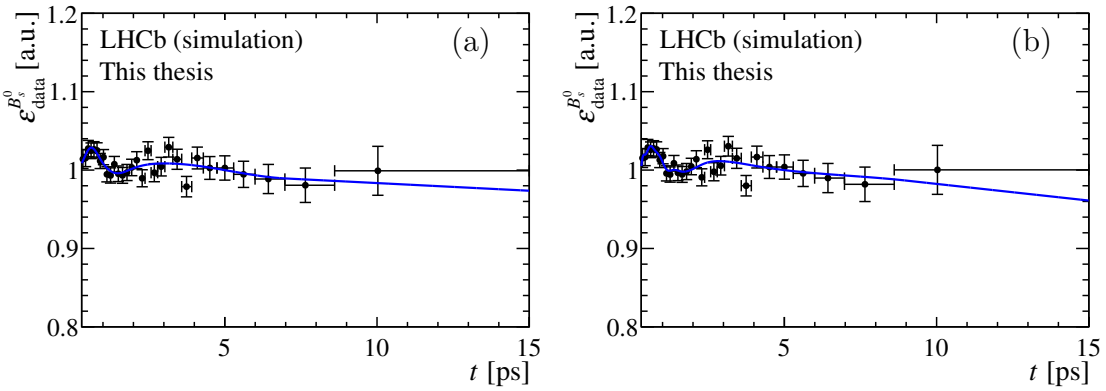


Figure 10.5: Decay-time acceptance of the 2016 `Unbiased` category. In (a) the nominal acceptance and in (b) the acceptance with an increased number of spline knots is shown.

new splines in the final fit results in a minor shift on the measured value of the decay-width difference $\Delta\Gamma_d^s$, which is assigned as a systematic uncertainty.

During the decay-time acceptance determination, the calibration data and simulation samples are corrected in two ways. The first is a weighting of the simulated samples to ensure the correct fraction of the S-wave component. In addition, the simulated samples and the $B^0 \to J/\psi K^{*0}$ data sample are weighted in kinematic variables to match the respective data sample and the $B_s^0 \to J/\psi \phi$ data sample, respectively. To get a conservative estimate of a possible systematic effect arising from the first mentioned "PDF" weighting, it is completely omitted, and new acceptances are determined. Since only very minor changes in the decay-width related parameters are observed, the differences are assigned as conservative systematic uncertainties.

In the case of the kinematic weighting, an alternative, less powerful, configuration of the boosted weighting algorithm, see Section 4.2, is tested. The resulting differences in the final fit are reasonable small and are therefore assigned as systematic uncertainty. In addition, a range of other configurations with similar or higher weighting power are tested, but the resulting differences in the final fit parameters are always smaller than the already assigned uncertainty.

The impact of the chosen values for the decay-time resolution of the different samples on the extracted decay-time acceptance is studied. In a conservative approach, these values are varied by $\pm 10\%$ and new decay-time acceptances are obtained. Repeating the final fit with these acceptances yields no significant shift of the final parameters, and therefore no systematic uncertainty is assigned.

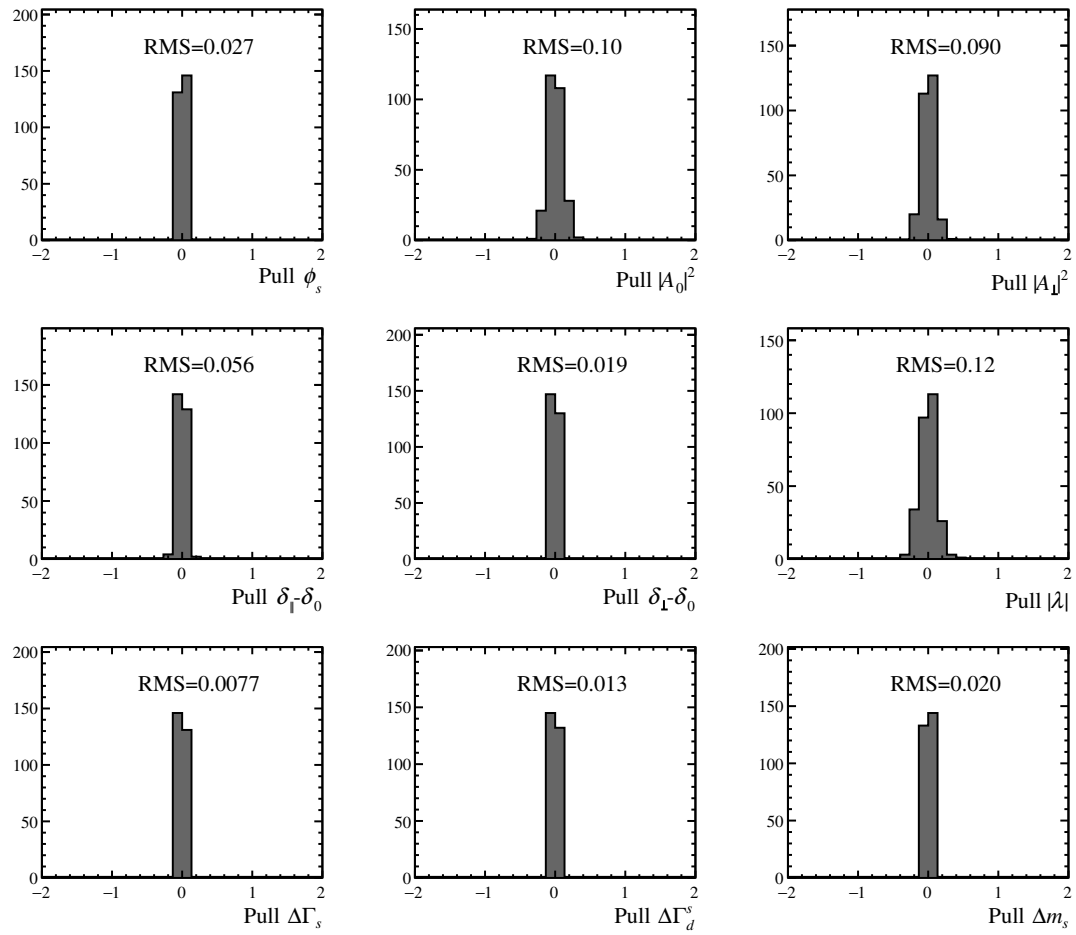


Figure 10.6: Spread of the main physics parameters in units of their statistical uncertainty when varying the angular acceptance within its statistical uncertainty.

All systematic uncertainties related to the decay-time acceptance are summarized in Tables 10.1 and 10.2 at the end of this chapter.

10.2.3 Angular acceptance

As for the decay-time acceptance, the angular acceptance, parametrized with the normalization weights, is subject to statistical fluctuations of the sample used to determine it. The related systematic uncertainty on the final fit parameters is accessed by repeating the decay-time and angular fit 300 times with normalization weights that are randomly varied according to their covariance matrix. Figure 10.6 shows the obtained deviations from the nominal fit results in units of the statistical uncertainty for the main parameters. The root mean square of these distributions is assigned as a systematic uncertainty. None of the parameters is effected by more than 12% of the respective statistical uncertainty.

Since the angular acceptance is determined from simulation, is has to be corrected for any differences between simulation and data, see Section 7.3.2. In this correction procedure, a range of choices of variables to correct for and of weighting configurations are made. To study the effect of these choices and possibly not corrected differences, a variety of alternatives of this procedure is tested:

- The iterative weighting procedure for the angular acceptance determinations starts with a weighting of the simulated sample to match the $p^{B_s^0}$, $p_T^{B_s^0}$ and $m(K^+, K^-)$ distributions in data. To ensure that these distributions also agree at later stages of the procedure, an alternative iterative procedure is tested where these variables are included in addition in all following weighting steps.

- Although the simulated sample reproduce the muon transverse momentum distributions of data quite nicely after the initial weighting step in $p^{B_s^0}$, $p_T^{B_s^0}$ and $m(K^+, K^-)$, an alternative iterative procedure is performed that includes the transverse momenta of both muons in addition to the kinematic variables of the kaons.

- A wide range ($\mathcal{O}(100)$) of different configurations of the boosted weighting algorithm used during the correction procedure is tested. The number and depth of the trees and the shrinkage factor are varied.

For all these alternative angular acceptance determinations, the final time-dependent angular fit is performed and for every parameter the maximal observed difference is considered as a systematic uncertainty. The numbers can be found in Tables 10.1 and 10.2 at the end of this chapter. For all parameters, the variation of the configuration of the boosted weighting algorithm results in the largest shift.

As described in Section 7.3.2, the determination and implementation of the angular acceptance requires that the angular acceptance is independent of the decay time. This assumption is tested by calculating the normalization weights in bins of the decay time. The result is shown in Figure 10.7 for the 2016-`Biased` data category. For some of the normalization weights, a significant inconsistency is observed. This can originate from the impact parameter related requirements in the trigger, see Section 5.1, which place geometrical requirements on the muon tracks that depend on the separation of the B_s^0 decay vertex from the primary vertex and therefore also on the decay time. To study the effect of this dependence on the final fit parameters, fits to the simulated samples of $B_s^0 \to J/\psi\,\phi$ decays are performed. First, the available simulated samples are split in subsamples of the

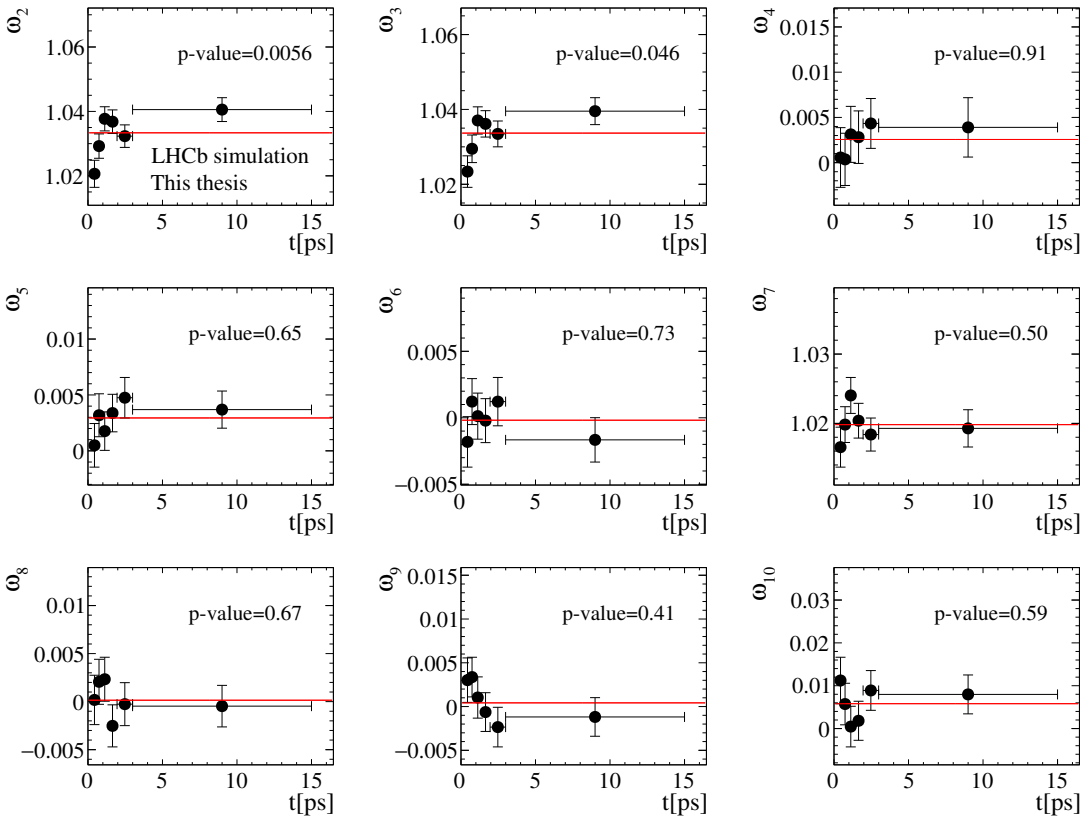


Figure 10.7: Angular acceptance normalization weights in regions of the reconstructed decay time. The values for the 2016 `Biased` category are shown. In addition, p-values for their consistency and a constant fit are shown.

same size as the background-subtracted $B_s^0 \to J/\psi\,\phi$ data sample. The simulated samples are weighted as discussed before, such that they are effectively generated with the parameters observed in the fit to data. From each sample, the decay-time and angular acceptance is determined individually and a time-dependent angular fit it performed. Using the acceptances determined from the same sample reduces the expected statistical spread of the fit parameters and increases the sensitivity to any biases. The decay-time resolution is determined from the full simulated sample by comparing reconstructed and true decay times. As the bias due to the neglected time dependence of the angular acceptance is not expected to depend on the tagging performance, the true initial B_s^0 flavour is used in the fits. In total, $\mathcal{O}(80)$ independent simulated samples and fits are performed and the deviation of the most important final parameters from the generated values is shown in Figure 10.8 in units of the statistical uncertainty observed in data. Especially the distributions of the two polarization amplitudes show a significant bias. A toy study, similar to the one presented in Section 8.3, is performed using the true initial B_s^0

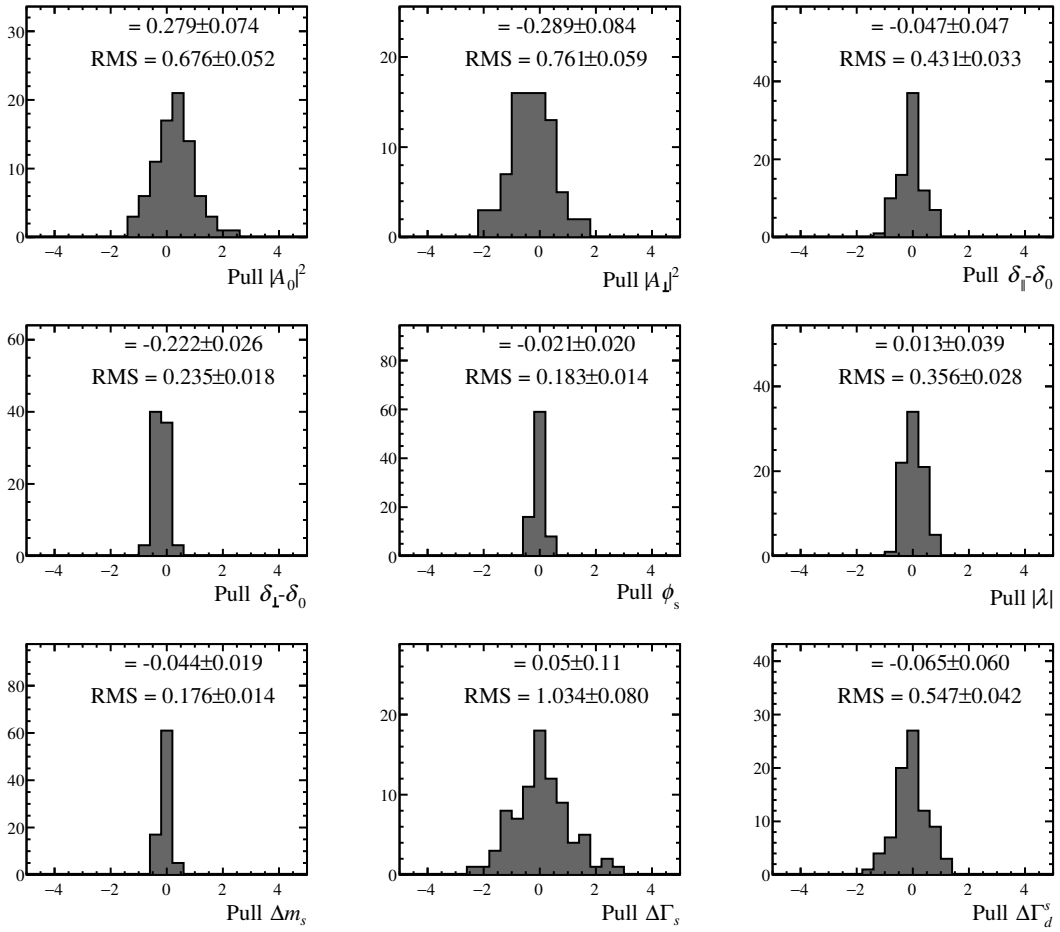


Figure 10.8: Spread around the true values of the main physics parameters observed in fits to simulated samples. The deviations are given in units of the respective statistical uncertainties observed in data.

flavour. No significant biases are found for the main parameters, indicating that the observed biases in the fits to the fully simulated samples are real effects and not related to the intrinsic fit bias. Therefore, the full observed biases, added in quadrature with their respective uncertainty, are taken as systematic uncertainty. One has to note that this systematic uncertainty might also cover other effects than the neglected dependencies of the angular acceptance. However, the main other systematic uncertainties related to the decay-time and angular acceptance are at least diminished by determining these acceptances from the exact same sample. Therefore, any double counting of systematic effects is expected to be small.

Again, all systematic uncertainties related to the angular acceptance are summarized in Tables 10.1 and 10.2 at the end of this chapter.

10.3 Other sources of systematic uncertainties

S-wave line shape

In Section 8.1 the C_{SP} factors are introduced, which are necessary to correctly model the interference between the resonant (P-wave) and S-wave K^+K^- component. Their calculation relies on a model of the P-wave and S-wave line shapes. While the P-wave is known to be a ϕ resonance, which has small model uncertainties, the S-wave line shape is not known to such high precision. Therefore, a systematic uncertainty is assigned by varying the S-wave model in several ways:

- The parameters for the nominal S-wave model, a Flatté distribution describing the $f_0(980)$ resonance, are taken from Ref. [91] and come with uncertainties. They are varied within 1σ in different combinations. In addition, the second solution for the $f_0(980)$ parameters found in Ref. [91] is tested.
- A model-independent parametrization with a cubic spline of the S-wave line shape is also tested as an alternative. This spline has 8 knots in the range $m(K^+K^-) \in [990, 1060]\,\mathrm{MeV/c^2}$, and the complex amplitude at each knot was measured in the analysis presented in Ref. [89].

New C_{SP} factors are calculated and the final fit is repeated for every alternative S-wave model. The largest observed deviation from the nominal result is assigned as systematic uncertainty. As expected, the S-wave amplitudes and phases are affected most. The main parameters show only small deviations, see Tables 10.1 and 10.2 at the end of this chapter.

Tagging related systematic uncertainties

As discussed in Section 6.4, the statistical and systematic uncertainties of the tagging calibration parameters are directly included in the final time- and angular-dependent fit by means of Gaussian constraints. The only additional studied systematic uncertainty originates from using a quadratic opposite side tagging calibration function instead of a linear one. However, the effect is negligible for all relevant parameters, and no systematic uncertainty is assigned.

Fit bias

In Section 8.3 a dedicated toy study was performed to estimate possible biases arising from the maximum likelihood fit itself. For every parameter, the bias observed in the residual distributions is assigned as a systematic uncertainty. It is

important to note that for strongly asymmetric likelihood shapes, see Appendix G.1, this is only a rough estimate. For those parameters the likelihood shapes themselves provide a much better representation of the best value and its uncertainties than a central value with possibly asymmetric statistical and systematic uncertainties. Anyhow, for all parameters the respective systematic uncertainty can be found in Tables 10.1 and 10.2.

Length and momentum scale

The determination of the decay time of a B_s^0 candidate basically only relies on its measured flight distance, momentum and mass. In the LHCb experiment, the length scale uncertainty is estimated to be 0.022% [17], which directly translates to the same relative uncertainty on the parameters $\Delta\Gamma_d^s$, $\Delta\Gamma_s$ and Δm_s. Given the statistical precision of these parameters, only Δm_s is significantly affected and an according systematic uncertainty is assigned, see Tables 10.1 and 10.2 at the end of this chapter.

The uncertainty on the momentum scale is with 0.03% slightly larger [46]. However, since the momentum, as well as the measured mass of the B_s^0 are affected by this uncertainty, the effect is canceled to large extend for the measured decay time and therefore also for $\Delta\Gamma_d^s$, $\Delta\Gamma_s$ and Δm_s. No systematic uncertainty needs to be assigned.

Truth matching of the simulated samples

The analysis presented in this thesis relies in many places on simulated samples of the decays $B_s^0 \to J/\psi\,\phi$ and $B^0 \to J/\psi\,K^{*0}$. As described in Section 2.2.3, these simulated decays are reconstructed in the same way as the real data and therefore have to be matched to the underlying generated decay afterwards. This so-called truth matching is not fully efficient and the decay candidates in the simulated sample that are categorized as background, i.e. not truth matched, can still contain true decays. In the nominal analysis they are included by applying the ${}_s\mathcal{P}lot$ technique to the simulated samples in the same way as it is done for data. A mass fit is performed to the full simulated sample with the tail parameters of the signal fixed to the ones obtained from the fit to the truth matched simulated sample. As in data, an exponential function models the background component. With the ${}_s\mathcal{P}lot$ technique the true background candidates are statistically subtracted from the full simulated samples, which can than be used in the following.

Table 10.1: Systematic uncertainties for some of the final parameters. For comparison, the respective statistical uncertainty is given in the first row.

Source	ϕ_s[rad]	$\lvert\lambda\rvert$	$\Delta\Gamma_s$[ps^{-1}]	Δm_s[ps^{-1}]	$\Delta\Gamma_d^s$[ps^{-1}]
Statistical uncertainty	0.041	0.016	$^{+0.0076}_{-0.0077}$	$^{+0.057}_{-0.059}$	0.0023
Mass parametrization	-	-	0.0002	0.001	-
Mass factorization	0.004	0.004	0.0022	0.016	0.0007
Multiple candidates	0.001	0.001	0.0001	0.001	0.0003
Time res.: prompt	-	-	-	0.001	-
Time res.: mean offset	0.003	0.001	0.0003	0.005	0.0002
Time res.: wrong PV	-	-	-	0.001	-
Dec. time acc.: statistical	-	-	0.0008	-	0.0012
Dec. time acc.: knot position	-	-	-	-	0.0002
Dec. time acc.: PDF weighting	-	-	0.0001	-	0.0001
Dec. time acc.: kinematic weighting	-	-	-	-	0.0002
Ang. acc.: statistical	0.001	0.002	-	0.001	-
Ang. acc.: sim. correction	0.002	0.004	0.0002	0.001	0.0001
Ang. acc.: Neglected dependencies	0.001	0.001	0.0010	0.003	0.0002
C_{SP} factors	0.001	0.001	0.0001	0.002	-
Fit bias	0.001	-	0.0003	0.001	0.0001
Length scale	-	-	-	0.004	-
Simulation truth matching	0.001	0.001	0.0001	-	0.0002
Quadratic sum of syst. unc.	0.006	0.006	0.0026	0.018	0.0015

A systematic is assigned by performing the background subtraction in simulation in an alternative way. Instead of applying the ${}_s\mathcal{P}lot$ technique to the whole simulated sample, it is only applied to the sample of candidates classified as background. The truth matched simulated candidates are kept without applying any weight for the background subtraction. The full analysis is repeated using these alternative simulated samples and the difference in the final parameters is assigned as a systematic uncertainty. Since the fraction of these wrongly classified background candidates in simulation is anyhow quite small (O(2%)), the effect on the final result is small compared to the statistical uncertainties, see Tables 10.1 and 10.2.

10.4 Summary of systematic uncertainties

Tables 10.1 and 10.2 show all considered systematic uncertainties for the main fit parameters. The respective tables for S-wave fractions and phases are shown in Appendix H. For all parameters, the quadratic sum of all systematic uncertainties is smaller than the respective statistical one. Especially the parameter of most interest, ϕ_s, is clearly statistically limited, and also the parameters λ, $\Delta\Gamma_s$ and Δm_s

Table 10.2: Systematic uncertainties for some of the final parameters. For comparison, the respective statistical uncertainty is given in the first row.

Source	$\lvert A_0\rvert^2$	$\lvert A_\perp\rvert^2$	$\delta_\parallel - \delta_0$[rad]	$\delta_\perp - \delta_0$[rad]
Statistical uncertainty	0.0029	0.0040	$^{+0.084}_{-0.073}$	0.13
Mass parametrization	0.0006	0.0005	0.009	0.05
Mass factorization	0.0002	0.0004	0.004	0.01
Multiple candidates	0.0006	0.0001	0.002	0.01
Time res.: prompt	-	-	0.001	-
Time res.: mean offset	-	-	0.001	0.08
Time res.: wrong PV	-	-	0.001	-
Dec. time acc.: statistical	0.0002	0.0003	-	-
Dec. time acc.: knot position	-	-	-	-
Dec. time acc.: PDF weighting	-	-	-	-
Dec. time acc.: kinematic weighting	-	-	-	-
Ang. acc.: statistical	0.0003	0.0004	0.004	-
Ang. acc.: sim. correction	0.0020	0.0011	0.008	0.01
Ang. acc.: Neglected dependencies	0.0008	0.0012	0.006	0.03
C_{SP} factors	-	0.0001	0.005	0.01
Fit bias	0.0001	0.0006	0.037	0.02
Length scale	-	-	-	-
Simulation truth matching	0.0002	0.0001	0.002	-
Quadratic sum of syst. unc.	0.0024	0.0019	0.040	0.10

have a statistical uncertainty, at least twice as large as the systematic one. They all are affected most by the systematic uncertainty assigned for the factorization of the mass and the decay time in the background subtraction procedure.

The lifetime difference $\Delta\Gamma_d^s$ shows a sizable systematic uncertainty. However, its largest contribution comes from the statistical uncertainty related to the decay-time acceptance and therefore directly reflects the statistical uncertainty of the $B^0 \to J/\psi\, K^{*0}$ data sample. Since the decay-width difference between the B_s^0 meson and the B^0 meson is measured, this systematic uncertainty is irreducible, given a $B^0 \to J/\psi\, K^{*0}$ data set, and could also be considered as a statistical one.

As expected, the polarization amplitudes obtain their largest systematic uncertainties from the angular acceptance. Especially the differences between data and simulation and the neglected dependence between angular acceptance and decay time contribute.

11 Summary

This thesis presented a flavour-tagged time- and angular-dependent analysis of the decay rate of $B_s^0 \to J/\psi\,\phi$ decays using proton-proton collision data collected by the LHCb experiment and corresponding to an integrated luminosity of $1.9\,\text{fb}^{-1}$. A complete list of all measured parameters with their statistical and systematic uncertainties is given in Table I.1 in Appendix I. In the following, only the most relevant parameters are discussed.

The weak phase difference ϕ_s, responsible for CP violation in the interference between B_s^0 meson mixing and the decay, and the parameter $|\lambda|$, parametrizing the amount of direct CP violation in this decay, have been measured to be:

$$\begin{aligned}\phi_s &= (-0.083 \pm 0.041_{\text{stat}} \pm 0.006_{\text{syst}})\,\text{rad},\\ |\lambda| &= 1.012 \pm 0.016_{\text{stat}} \pm 0.006_{\text{syst}},\end{aligned}$$

where the first and second uncertainty are statistical and systematic, respectively. Thanks to an improved selection and tagging performance, and the higher center-of-mass energy of 13 TeV, these results supersede the precision reached in the previous LHCb analysis [5]. Furthermore, the mass and decay-width splitting in the B_s^0 meson system, Δm_s and $\Delta\Gamma_s$, were determined to be:

$$\begin{aligned}\Delta\Gamma_s &= (0.0773 \pm^{0.0076}_{0.0077\ \text{stat}} \pm 0.0026_{\text{syst}})\,\text{ps}^{-1},\\ \Delta m_s &= (17.702 \pm^{0.057}_{0.059\ \text{stat}} \pm 0.018_{\text{syst}})\,\text{ps}^{-1}.\end{aligned}$$

These measurements build the basis for a combination of a range of LHCb analysis measuring ϕ_s, $|\lambda|$ and $\Delta\Gamma_s$. This list further contains studies of the decay $B_s^0 \to J/\psi\,\phi$ using LHCb Run I data [5], $B_s^0 \to J/\psi\,\pi^+\pi^-$ decays using Run I and Run II data [22,92], $B_s^0 \to J/\psi\,K^+K^+$ decays with a two kaon invariant mass above $1.05\,\text{GeV/c}^2$ [89], $B_s^0 \to \psi(2S)\phi$ decays [93] and $B_s^0 \to D_s^+ D_s^-$ [94]. The combination

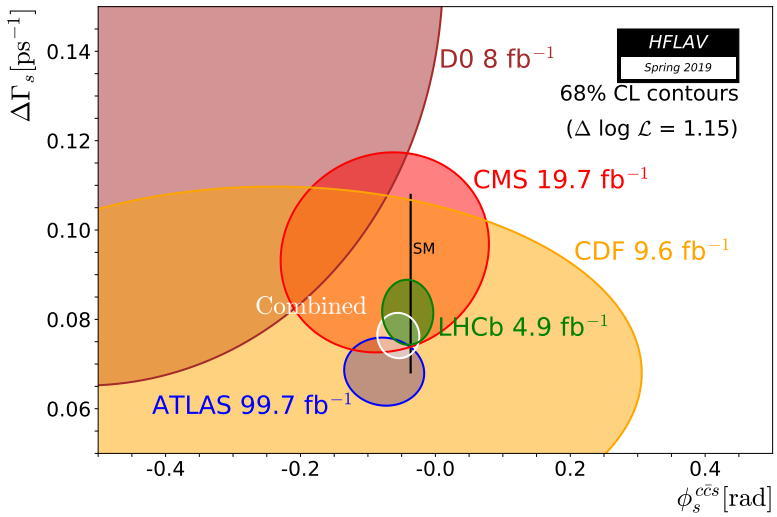


Figure 11.1: The experimental status of ϕ_s and $\Delta\Gamma_s$ measurements shortly after the analysis presented in this thesis was published. The black bar indicates the SM calculation and constraint of $\Delta\Gamma_s$ and ϕ_s, respectively. Preliminary figure taken from Ref. [16].

is presented in [63] and yields the following values for the CP violation parameters and $\Delta\Gamma_s$:

$$\begin{aligned}\phi_s &= (-0.041 \pm 0.025)\,\text{rad},\\ |\lambda| &= 0.993 \pm 0.010,\\ \Delta\Gamma_s &= (0.0816 \pm 0.0048)\,\text{ps}^{-1}.\end{aligned}$$

Figure 11.1 shows the preliminary updated current experimental status. Besides the new LHCb result presented above, a new ATLAS measurement [95], presented at the same day as the LHCb result, strongly contributes to the world average. The corresponding world average values for the parameters:

$$\begin{aligned}\Delta\Gamma_s^{\text{w.a.}} &= (0.0764 \pm 0.0034)\,\text{ps}^{-1},\\ \phi_s^{\text{w.a.}} &= (-0.055 \pm 0.021)\,\text{rad}\end{aligned}$$

are consistent with the prediction within the SM, and therefore no sign for contributions from beyond the SM is visible. However, the experiments are reaching now a level of precision on ϕ_s that clearly underlines the importance of more precise estimates of the impact of higher order penguin diagrams, see Section 1.5.5. Without those, possible deviations from the SM of future measurements, consistent with the current world average, might be explained by these penguin contributions and not

by any effect beyond the SM. The LHCb collaboration has the potential and also the responsibility to provide not only the most precise measurement of ϕ_s from a single experiment, but also to estimate the penguin pollution to a level that allows a clean interpretation of the results. With the strategy followed in Ref. [27], updated measurements of these pollutions will be possible.

Besides these CP violation and mixing related quantities, the decay-width difference between the B_s^0 and B^0 mesons was directly measured. Using the current world average for the B^0 decay width of $(\Gamma_d = 0.6579 \pm 0.0017)\,\mathrm{ps}^{-1}$ [16], this difference can be transformed to the theoretically more interesting ratio of decay widths:

$$\begin{aligned}\Gamma_s - \Gamma_d &= (-0.0040 \pm 0.0023_{\mathrm{stat}} \pm 0.0015_{\mathrm{syst}})\,\mathrm{ps}^{-1},\\ \Gamma_s/\Gamma_d &= 0.9938 \pm 0.0035_{\mathrm{stat}} \pm 0.0023_{\mathrm{syst}}.\end{aligned}$$

Here, the uncertainty on Γ_d is negligible, and the uncertainty on Γ_s/Γ_d is purely determined from the uncertainty on $\Gamma_s - \Gamma_d$. The value for the ratio is consistent within 2.3σ with the current world average, $\Gamma_s/\Gamma_d = 1.0070 \pm 0.0040$, and is reaching the same precision. It is also perfectly compatible with the theory prediction of a value very close to unity [32].

Outlook

Although the current experimental value of the CP violating phase difference ϕ_s is in good agreement with the SM expectation, the interest in more precise measurements remains high. Assuming future updated estimates of the impact of higher order penguin contributions, the SM expectation is still significantly more precise. The LHCb experiment will also in future be the main contributor to the experimental value. In the years 2017 and 2018, it collected additional proton-proton data, corresponding to $4\,\mathrm{fb}^{-1}$ of integrated luminosity, which is now analyzed. With this and further potential improvements in especially the flavor tagging, another doubling of the effective sample size seems to be realistic. Given the small current systematic uncertainty, also these future measurements will be statistically limited. In addition, the LHCb detector is currently upgraded and is planned to be run at a significantly higher instantaneous luminosity. Within the next decade, an integrated luminosity of up to $50\,\mathrm{fb}^{-1}$ will be reached. With this, the precision on ϕ_s should clearly fall below 10 mrad and will either put stringent constraints on non-SM contributions or will establish these.

A Polarization dependent CP violation

As mentioned in Section 1.5.7, the time- and angular-dependent decay rate of $B_s^0 \to J/\psi\,\phi$ decays can be parametrized allowing polarization dependent CP violation. It is then given by:

$$\frac{\mathrm{d}\Gamma(B_s^0 \to J/\psi\,\phi)}{\mathrm{d}t\mathrm{d}\theta_\mu\mathrm{d}\theta_K\mathrm{d}\varphi_h} \propto \sum_{k=1}^{10} A_k \tilde{h}_{k,+1}(t) f_k(\theta_\mu, \theta_K, \varphi_h), \tag{A.1}$$

$$\frac{\mathrm{d}\Gamma(\overline{B}_s^0 \to J/\psi\phi)}{\mathrm{d}t\mathrm{d}\theta_\mu\mathrm{d}\theta_K\mathrm{d}\varphi_h} \propto \sum_{k=1}^{10} A_k \tilde{h}_{k,-1}(t) f_k(\theta_\mu, \theta_K, \varphi_h), \tag{A.2}$$

where the time-dependent terms are given by:

$$\begin{aligned}\tilde{h}_{k,+1} = e^{-\Gamma_s t}\Big[&\tilde{a}_k \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + \tilde{b}_k \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) \\ &+ \tilde{c}_k \cos\left(\Delta m_s t\right) + \tilde{d}_k \sin\left(\Delta m_s t\right)\Big],\end{aligned} \tag{A.3}$$

$$\begin{aligned}\tilde{h}_{k,-1} = e^{-\Gamma_s t}\Big[&\tilde{a}_k \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + \tilde{b}_k \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) \\ &- \tilde{c}_k \cos\left(\Delta m_s t\right) - \tilde{d}_k \sin\left(\Delta m_s t\right)\Big].\end{aligned} \tag{A.4}$$

The angular functions $f_k(\theta_\mu, \theta_K, \varphi_h)$ and amplitudes A_k are the same as in the polarization independent case, and the coefficients $\tilde{a}_k$, $\tilde{b}_k$, $\tilde{c}_k$ and $\tilde{d}_k$ are defined in Table A.1, following Ref. [30].

Table A.1: Coefficients of the time-dependent terms in the decay rate of $B_s^0 \to J/\psi\,\phi$ decays with polarization dependent CP violation.

k	a_k	b_k	c_k	d_k
1	$\frac{1}{2}(1+\lvert\lambda_0\rvert^2)$	$-\lvert\lambda_0\rvert\cos(\phi_0)$	$\frac{1}{2}(1-\lvert\lambda_0\rvert^2)$	$\lvert\lambda_0\rvert\sin(\phi_0)$
2	$\frac{1}{2}(1+\lvert\lambda_\parallel\rvert^2)$	$-\lvert\lambda_\parallel\rvert\cos(\phi_\parallel)$	$\frac{1}{2}(1-\lvert\lambda_\parallel\rvert^2)$	$\lvert\lambda_\parallel\rvert\sin(\phi_\parallel)$
3	$\frac{1}{2}(1+\lvert\lambda_\perp\rvert^2)$	$\lvert\lambda_\perp\rvert\cos(\phi_\perp)$	$\frac{1}{2}(1-\lvert\lambda_\perp\rvert^2)$	$-\lvert\lambda_\perp\rvert\sin(\phi_\perp)$
4	$\frac{1}{2}\Big[\sin(\delta_\perp-\delta_\parallel)-\lvert\lambda_\perp\lambda_\parallel\rvert$ $\sin(\delta_\perp-\delta_\parallel-\phi_\perp+\phi_\parallel)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_\perp\rvert\sin(\delta_\perp-\delta_\parallel-\phi_\perp)$ $+\lvert\lambda_\parallel\rvert\sin(\delta_\parallel-\delta_\perp-\phi_\parallel)\Big]$	$\frac{1}{2}\Big[\sin(\delta_\perp-\delta_\parallel)+\lvert\lambda_\perp\lambda_\parallel\rvert$ $\sin(\delta_\perp-\delta_\parallel-\phi_\perp+\phi_\parallel)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_\perp\rvert\cos(\delta_\perp-\delta_\parallel-\phi_\perp)$ $+\lvert\lambda_\parallel\rvert\cos(\delta_\parallel-\delta_\perp-\phi_\parallel)\Big]$
5	$\frac{1}{2}\Big[\cos(\delta_0-\delta_\parallel)+\lvert\lambda_0\lambda_\parallel\rvert$ $\cos(\delta_0-\delta_\parallel-\phi_0+\phi_\parallel)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_0\rvert\cos(\delta_0-\delta_\parallel-\phi_0)$ $+\lvert\lambda_\parallel\rvert\cos(\delta_\parallel-\delta_0-\phi_\parallel)\Big]$	$\frac{1}{2}\Big[\cos(\delta_0-\delta_\parallel)-\lvert\lambda_0\lambda_\parallel\rvert$ $\cos(\delta_0-\delta_\parallel-\phi_0+\phi_\parallel)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_0\rvert\sin(\delta_0-\delta_\parallel-\phi_0)$ $+\lvert\lambda_\parallel\rvert\sin(\delta_\parallel-\delta_0-\phi_\parallel)\Big]$
6	$-\frac{1}{2}\Big[\sin(\delta_0-\delta_\perp)-\lvert\lambda_0\lambda_\perp\rvert$ $\sin(\delta_0-\delta_\perp-\phi_0+\phi_\perp)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_0\rvert\sin(\delta_0-\delta_\perp-\phi_0)$ $+\lvert\lambda_\perp\rvert\sin(\delta_\perp-\delta_0-\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\sin(\delta_0-\delta_\perp)+\lvert\lambda_0\lambda_\perp\rvert$ $\sin(\delta_0-\delta_\perp-\phi_0+\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_0\rvert\cos(\delta_0-\delta_\perp-\phi_0)$ $+\lvert\lambda_\perp\rvert\cos(\delta_\perp-\delta_0-\phi_\perp)\Big]$
7	$\frac{1}{2}(1+\lvert\lambda_S\rvert^2)$	$\lvert\lambda_S\rvert\cos(\phi_S)$	$\frac{1}{2}(1-\lvert\lambda_S\rvert^2)$	$-\lvert\lambda_S\rvert\sin(\phi_S)$
8	$\frac{1}{2}\Big[\cos(\delta_S-\delta_\parallel)-\lvert\lambda_S\lambda_\parallel\rvert$ $\cos(\delta_S-\delta_\parallel-\phi_S+\phi_\parallel)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\cos(\delta_S-\delta_\parallel-\phi_S)$ $-\lvert\lambda_\parallel\rvert\cos(\delta_\parallel-\delta_S-\phi_\parallel)\Big]$	$\frac{1}{2}\Big[\cos(\delta_S-\delta_\parallel)+\lvert\lambda_S\lambda_\parallel\rvert$ $\cos(\delta_S-\delta_\parallel-\phi_S+\phi_\parallel)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\sin(\delta_S-\delta_\parallel-\phi_S)$ $-\lvert\lambda_\parallel\rvert\sin(\delta_\parallel-\delta_S-\phi_\parallel)\Big]$
9	$-\frac{1}{2}\Big[\sin(\delta_S-\delta_\perp)+\lvert\lambda_S\lambda_\perp\rvert$ $\sin(\delta_S-\delta_\perp-\phi_S+\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_S\rvert\sin(\delta_S-\delta_\perp-\phi_S)$ $-\lvert\lambda_\perp\rvert\sin(\delta_\perp-\delta_S-\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\sin(\delta_S-\delta_\perp)-\lvert\lambda_S\lambda_\perp\rvert$ $\sin(\delta_S-\delta_\perp-\phi_S+\phi_\perp)\Big]$	$-\frac{1}{2}\Big[-\lvert\lambda_S\rvert\cos(\delta_S-\delta_\perp-\phi_S)$ $+\lvert\lambda_\perp\rvert\cos(\delta_\perp-\delta_S-\phi_\perp)\Big]$
10	$\frac{1}{2}\Big[\cos(\delta_S-\delta_0)-\lvert\lambda_S\lambda_0\rvert$ $\cos(\delta_S-\delta_0-\phi_S+\phi_0)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\cos(\delta_S-\delta_0-\phi_S)$ $-\lvert\lambda_0\rvert\cos(\delta_0-\delta_S-\phi_0)\Big]$	$\frac{1}{2}\Big[\cos(\delta_S-\delta_0)+\lvert\lambda_S\lambda_0\rvert$ $\cos(\delta_S-\delta_0-\phi_S+\phi_0)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\sin(\delta_S-\delta_0-\phi_S)$ $-\lvert\lambda_0\rvert\sin(\delta_0-\delta_S-\phi_0)\Big]$

B Trigger criteria

In the following, a more detailed summary of the trigger requirements mentioned in Section 5.1 is given.

B.1 First stage of the software trigger

Table B.1 shows the requirements of the selected trigger lines in the first stage of the software trigger (HLT1). The variables not mentioned in Section 5.1 are defined as following:

- `IsMuon` encodes whether a track has associated hits in the muon stations
- $p_{\text{track}}^{\text{ghost}}$ is the output of a multivariate classifier that was trained to reject ghost tracks
- χ^2_{FD} is the significance of the distance between the PV and the respective vertex.

In addition, the trigger line `Hlt1TwoTrackMVA` contains a cut on a multivariate classifier (MVA) that uses the variables given in the respective row as input.

B.2 Second stage of the software trigger

Table B.2 shows the requirements of the selected trigger line in the second stage of the software trigger (HLT2).

Table B.1: Configuration of the HLT1 trigger lines used in this analysis. The numbers changed over the years 2015 and 2016, and only the configuration with with the largest fraction of data was recorded is shown.

(a) `Hlt1DiMuonHighMass`

Object	Variable	Cut
μ^+, μ^-	p_T	$> 300\,\mathrm{MeV/c}$
	p	$> 6000\,\mathrm{MeV/c}$
	$\chi^2_{\mathrm{track}}/\mathrm{nDoF}$	< 4
	`IsMuon`	`true`
$\mu^+ + \mu^-$	DOCA	$< 0.2\,\mathrm{mm}$
	χ^2_{vtx}	< 25
	$m(\mu^+\mu^-)$	$> 2700\,\mathrm{MeV/c^2}$

(b) `Hlt1TrackMuon`

Object	Variable	Cut
μ^+/μ^-	p_T	$> 1100\,\mathrm{MeV/c}$
	p	$> 6000\,\mathrm{MeV/c}$
	χ^2_{IP}	> 35
	$\chi^2_{\mathrm{track}}/\mathrm{nDoF}$	< 3
	$p^{\mathrm{ghost}}_{\mathrm{track}}$	< 0.2
	`IsMuon`	`true`

(c) `Hlt1TwoTrackMVA`

Object	Variable	Cut
track 1, track 2	p_T	$> 500\,\mathrm{MeV/c}$
	p	$> 5000\,\mathrm{MeV/c}$
	$\chi^2_{\mathrm{track}}/\mathrm{nDoF}$	< 2.5
	$p^{\mathrm{ghost}}_{\mathrm{track}}$	< 0.2
	χ^2_{IP}	> 4
	`IsMuon`	`true`
$\mu^+ + \mu^-$	χ^2_{vtx}	< 10
	$m(\mu^+\mu^-)$	$> 1000\,\mathrm{MeV/c^2}$
	MVA(χ^2_{vtx}, χ^2_{FD}, $\sum p_T$, #tracks with $\chi^2_{\mathrm{IP}} < 16$)	

Table B.2: Configuration of the HLT2 trigger line used in this analysis.

Object	Variable	Cut
μ^+, μ^-	$p^{\mathrm{ghost}}_{\mathrm{track}}$	< 0.4
	$\chi^2_{\mathrm{track}}/\mathrm{nDoF}$	< 4
	`IsMuon`	`true`
$\mu^+ + \mu^-$	χ^2_{FD}	$> 3\,\mathrm{mm}$
	χ^2_{vtx}	< 25
	$\lvert m(\mu^+\mu^-) - m(J/\psi)\rvert$	$< 120\,\mathrm{MeV/c^2}$

C Mass fit parameters

In Section 5.3, a fit to the mass distribution of the selected $B_s^0 \to J/\psi\,\phi$ candidates is performed. Tables C.1 to C.4 show the fit parameters for the different data categories. The tail parameters are fixed from fits to the simulated sample and are therefore quoted without an uncertainty.

Table C.1: Parameters of the fit to the mass of $B_s^0 \to J/\psi\,\phi$ candidates in the 2015 `Unbiased` data category. The values without uncertainty are fixed from a fit to the simulated sample.

	$m(KK)$ bin 1	$m(KK)$ bin 2	$m(KK)$ bin 3	$m(KK)$ bin 4	$m(KK)$ bin 5	$m(KK)$ bin 6
s_1	0.84 ± 0.57	1.02 ± 0.18	0.915 ± 0.089	1.120 ± 0.098	1.16 ± 0.15	1.05 ± 0.24
s_2	0.05 ± 0.11	0.034 ± 0.030	0.043 ± 0.015	0.009 ± 0.017	0.004 ± 0.026	0.001 ± 0.039
μ	5366.66 ± 0.54	5366.46 ± 0.22	5366.75 ± 0.11	5367.48 ± 0.12	5367.38 ± 0.19	5367.07 ± 0.28
γ	0.00193 ± 0.00032	0.00130 ± 0.00035	0.00107 ± 0.00036	0.00185 ± 0.00036	0.00113 ± 0.00033	0.00142 ± 0.00022
f_{sig}	0.176 ± 0.014	0.591 ± 0.013	0.8297 ± 0.0060	0.8212 ± 0.0067	0.614 ± 0.011	0.271 ± 0.010
f_{B^0}	0.021 ± 0.011	0.012 ± 0.011	0.030 ± 0.012	0.0000 ± 0.0047	0.011 ± 0.011	0.0000 ± 0.0048
α_1	2.77	2.09	2.10	2.17	2.22	2.47
α_2	-2.06	-1.92	-2.04	-2.01	-2.12	-2.28
n_1	0.332	1.65	3.09	3.20	1.96	1.11
n_2	1.83	3.00	3.78	3.09	1.69	1.30

Table C.2: Parameters of the fit to the mass of $B_s^0 \to J/\psi\,\phi$ candidates in the 2015 `Biased` data category. The values without uncertainty are fixed from a fit to the simulated sample.

	$m(KK)$ bin 1	$m(KK)$ bin 2	$m(KK)$ bin 3	$m(KK)$ bin 4	$m(KK)$ bin 5	$m(KK)$ bin 6
s_1	0.71 ± 0.70	1.23 ± 0.33	0.92 ± 0.20	0.82 ± 0.20	1.79 ± 0.30	1.11 ± 0.39
s_2	0.09 ± 0.13	0.003 ± 0.054	0.052 ± 0.034	0.065 ± 0.034	-0.084 ± 0.049	0.016 ± 0.066
μ	5364.5 ± 1.0	5366.15 ± 0.42	5366.69 ± 0.20	5367.21 ± 0.23	5367.41 ± 0.40	5367.21 ± 0.54
γ	0.00116 ± 0.00069	0.00090 ± 0.00070	0.00289 ± 0.00066	0.00090 ± 0.00075	0.00061 ± 0.00075	0.00167 ± 0.00049
f_{sig}	0.245 ± 0.034	0.658 ± 0.023	0.850 ± 0.010	0.845 ± 0.012	0.685 ± 0.021	0.348 ± 0.022
f_{B^0}	0.000 ± 0.011	0.000 ± 0.032	0.0000 ± 0.0080	0.033 ± 0.026	0.061 ± 0.029	0.024 ± 0.016
α_1	2.89	2.08	1.98	2.06	2.19	2.59
α_2	-2.23	-2.09	-2.13	-1.77	-2.04	-2.35
n_1	0.133	1.62	3.21	3.33	1.88	1.02
n_2	1.64	2.35	3.54	4.14	2.09	1.22

Table C.3: Parameters of the fit to the mass of $B_s^0 \to J/\psi\,\phi$ candidates in the 2016 `Unbiased` data category. The values without uncertainty are fixed from a fit to the simulated sample.

	$m(KK)$ bin 1	$m(KK)$ bin 2	$m(KK)$ bin 3	$m(KK)$ bin 4	$m(KK)$ bin 5	$m(KK)$ bin 6
s_1	1.09 ± 0.18	1.085 ± 0.074	1.016 ± 0.033	0.992 ± 0.038	0.988 ± 0.064	0.993 ± 0.093
s_2	0.014 ± 0.031	0.019 ± 0.013	0.0243 ± 0.0056	0.0317 ± 0.0065	0.034 ± 0.011	0.026 ± 0.016
μ	5367.26 ± 0.22	5366.669 ± 0.083	5366.824 ± 0.040	5367.401 ± 0.045	5367.430 ± 0.073	5367.02 ± 0.11
γ	0.00193 ± 0.00011	0.00134 ± 0.00012	0.00141 ± 0.00012	0.00126 ± 0.00013	0.00171 ± 0.00012	0.001693 ± 0.000080
f_{sig}	0.1584 ± 0.0048	0.5599 ± 0.0047	0.8113 ± 0.0023	0.7955 ± 0.0026	0.5694 ± 0.0043	0.2626 ± 0.0037
f_{B^0}	0.0059 ± 0.0033	0.0086 ± 0.0039	0.0004 ± 0.0080	0.0070 ± 0.0040	0.0034 ± 0.0036	0.0012 ± 0.0024
α_1	2.63	2.09	2.03	2.12	2.24	2.37
α_2	-2.22	-2.08	-2.02	-1.95	-2.05	-2.30
n_1	0.470	1.68	3.60	3.33	1.97	1.37
n_2	1.71	2.57	4.05	3.64	1.98	1.37

Table C.4: Parameters of the fit to the mass of $B_s^0 \to J/\psi\,\phi$ candidates in the 2016 `Biased` data category. The values without uncertainty are fixed from a fit to the simulated sample.

	$m(KK)$ bin 1	$m(KK)$ bin 2	$m(KK)$ bin 3	$m(KK)$ bin 4	$m(KK)$ bin 5	$m(KK)$ bin 6
s_1	0.94 ± 0.14	1.23 ± 0.15	1.194 ± 0.071	1.009 ± 0.082	0.98 ± 0.14	0.63 ± 0.18
s_2	0.048 ± 0.027	-0.002 ± 0.027	0.000 ± 0.012	0.036 ± 0.014	0.044 ± 0.024	0.100 ± 0.032
μ	5366.88 ± 0.43	5366.65 ± 0.17	5366.761 ± 0.084	5367.402 ± 0.096	5367.15 ± 0.16	5366.81 ± 0.22
γ	0.00153 ± 0.00028	0.00107 ± 0.00033	0.00150 ± 0.00030	0.00104 ± 0.00033	0.00158 ± 0.00030	0.00152 ± 0.00021
f_{sig}	0.233 ± 0.013	0.6712 ± 0.0099	0.8643 ± 0.0043	0.8543 ± 0.0049	0.6688 ± 0.0092	0.3566 ± 0.0095
f_{B^0}	0.0057 ± 0.0089	0.014 ± 0.011	0.000 ± 0.012	0.018 ± 0.010	0.0098 ± 0.0093	0.0049 ± 0.0063
α_1	2.66	2.03	2.02	2.11	2.17	2.52
α_2	-2.23	-2.05	-1.97	-2.01	-2.06	-2.33
n_1	0.396	1.74	3.37	3.28	2.14	1.12
n_2	1.57	2.50	4.17	3.28	1.87	1.38

D Tagging

In this section the general formalism to add the information from two independent tagging algorithms to a PDF that depends on the initial B flavour is derived. Given two tagging algorithms, with tagging decisions q_1, $q_2 \in \{-1, 0, 1\}$ and mistag probabilities $\omega_1^{(-)}$, $\omega_2^{(-)}$, and a PDF that depends on the initial flavour of the B meson, $\mathrm{PDF}(...|\overline{B}^{(-)})$, the PDF describing the observed distributions correctly is given by:

$$\begin{aligned}\mathrm{PDF}(...|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)}) =& p(B|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)})\,\mathrm{PDF}(...|B) \\ &+ p(\bar{B}|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)})\,\mathrm{PDF}(...|\bar{B}),\end{aligned} \tag{D.1}$$

where $p(B|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)})$ and $p(\bar{B}|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)})$ are the probabilities that a certain candidate is an initial B or $\bar{B}$ meson, given its tagging algorithm outputs.

Following Bayes theorem, these probabilities can be expressed as:

$$\begin{aligned}p(B|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)}) =& \frac{p(q_1, q_2|B, \omega_1^{(-)}, \omega_2^{(-)})\, p(B|\omega_1^{(-)}, \omega_2^{(-)})}{p(q_1, q_2|\omega_1^{(-)}, \omega_2^{(-)})} \\ =& \frac{p(q_1, q_2|B, \omega_1^{(-)}, \omega_2^{(-)})\, \frac{p(\omega_1^{(-)}, \omega_2^{(-)}|B)\, p(B)}{p(\omega_1^{(-)}, \omega_2^{(-)})}}{\frac{p(\omega_1^{(-)}, \omega_2^{(-)}|q_1, q_2)\, p(q_1, q_2)}{p(\omega_1^{(-)}, \omega_2^{(-)})}} \\ =& p(q_1, q_2|B, \omega_1^{(-)}, \omega_2^{(-)})\, p(B)\, \frac{p(\omega_1^{(-)}, \omega_2^{(-)}|B)}{p(\omega_1^{(-)}, \omega_2^{(-)})\, p(q_1, q_2)} \\ :=& p(q_1, q_2|B, \omega_1^{(-)}, \omega_2^{(-)}) p(B)\, X^B(q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)}),\end{aligned} \tag{D.2}$$

with treating $p(\bar{B}|q_1, q_2, \omega_1^{(-)}, \omega_2^{(-)})$ in full analogy. Here, X^B is defined as shown above and will be discussed later.

Each tagging algorithm comes with a tagging efficiency $\varepsilon(\overline{B}^{(-)})$, which might depend on the true initial flavour. In addition, there might be a production and/or selection asymmetry A_p that alters the overall ratio of initial B and $\bar{B}$ mesons:

$$p(B) = \frac{1 + A_p}{1 - A_p}\, p(\bar{B}), \tag{D.3}$$

where $p(B)$ and $p(\bar{B})$ are the overall probabilities that a candidate was produced as B or $\bar{B}$ mesons if no further information is given. Given these probabilities and the

tagging efficiencies, the probabilities $p(q_1, q_2|B, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2)$ and $p(q_1, q_2|\bar{B}, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2)$ can be calculated. Before giving the general expression, one specific case is chosen to show the derivation. We want to get the probability $p(q_1 = -1, q_2 = 0|B, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2)$, i.e. the probability that the first tagging algorithm identifies a $\bar{B}$ meson and the second algorithm gives no decision in the case that the true flavour corresponds to a B meson. The assumption is that the two tagging algorithms are independent, which means:

$$p(q_1 = -1, q_2 = 0|B, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2) = p(q_1 = -1|B, \overset{(-)}{\omega}_1)\, p(q_2 = 0|B, \overset{(-)}{\omega}_1). \tag{D.4}$$

Using the tagging efficiency and the definition of the mistag probabilities, this can be written as:

$$p(q_1 = -1, q_2 = 0|B, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2) = \varepsilon_1(B)\, \omega_1\, (1 - \varepsilon_2(B)). \tag{D.5}$$

The other cases can be easily calculated in the same way. It is possible to give a closed-form expression covering all of them:

$$\begin{aligned} p(q_1, q_2|\overset{(-)}{B}, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2) = &\left(\frac{1}{2}\right)^{|q_1|} \left[1 \overset{+}{(-)} q_1(1 - 2\overset{(-)}{\omega}_1)\right] \varepsilon_1(\overset{(-)}{B})^{|q_1|} \left[1 - \varepsilon_1(\overset{(-)}{B})\right]^{1-|q_1|} \\ &\cdot \left(\frac{1}{2}\right)^{|q_2|} \left[1 \overset{+}{(-)} q_2(1 - 2\overset{(-)}{\omega}_2)\right] \varepsilon_2(\overset{(-)}{B})^{|q_2|} \left[1 - \varepsilon_2(\overset{(-)}{B})\right]^{1-|q_2|}. \end{aligned} \tag{D.6}$$

Plugging this in Equation (D.1) yields:

$$\begin{aligned} \mathrm{PDF}(...|q_1, q_2, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2) = \frac{1}{N} \Bigg[&[1 + q_1(1 - 2\omega_1)]\, [1 + q_2(1 - 2\omega_2)]\, \mathrm{PDF}(...|B) \\ &+ [1 - q_1(1 - 2\bar{\omega}_1)]\, [1 - q_2(1 - 2\bar{\omega}_2)]\, \frac{p(\bar{B})}{p(B)}\, \frac{X^{\bar{B}}}{X^{B}} \\ &\cdot \left(\frac{\varepsilon_1(\bar{B})}{\varepsilon_1(B)}\right)^{|q_1|} \left(\frac{\varepsilon_2(\bar{B})}{\varepsilon_2(B)}\right)^{|q_2|} \left(\frac{1 - \varepsilon_1(\bar{B})}{1 - \varepsilon_1(B)}\right)^{1-|q_1|} \left(\frac{1 - \varepsilon_2(\bar{B})}{1 - \varepsilon_2(B)}\right)^{1-|q_2|} \mathrm{PDF}(...|\bar{B}) \Bigg], \end{aligned} \tag{D.7}$$

where N is a normalization factor depending on the tagging decisions and mistag probabilities. The production asymmetry enters via:

$$\frac{p(\bar{B})}{p(B)} = \frac{1 - A_P}{1 + A_p}, \tag{D.8}$$

and the ratio $X^{\bar{B}}/X^{B}$ can be expressed using its definition in Equation (D.2):

$$\frac{X^{\bar{B}}}{X^{B}} = \frac{p(\overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2|\bar{B})}{p(\overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2|B)}. \tag{D.9}$$

Given that the mistag probabilities are derived from estimated mistag probabilities η_1 and η_2, as it is the case in the analysis presented in this thesis, this ratio is unity if the distribution of these estimated mistag probabilities is independent of the true flavour. This is confirmed in simulation and also the production and tagging efficiency asymmetries are typically neglected, which then leads to the formulas shown in Section 6.3.

Toy studies were performed with a production asymmetry of 3% and no significant biases on top of the ones observed with no production asymmetry are visible. However, for future reference, in the following a generic extension to the PDF is presented that allows to incorporate possible production and tagging efficiency asymmetries.

It is not possible to determine all five additional parameters, A_p, $\varepsilon_1(B)$, $\varepsilon_1(\bar{B})$, $\varepsilon_2(B)$ and $\varepsilon_2(\bar{B})$. In principle, the PDF is sensitive to an asymmetry factor $A(q_1, q_2)$ for every of the nine possible combinations of the tagging decisions q_1 and q_2:

$$\begin{aligned}\mathrm{PDF}(...|q_1, q_2, \overset{(-)}{\omega}_1, \overset{(-)}{\omega}_2) = \frac{1}{N}\big[\,&[1 + q_1(1 - 2\omega_1)]\,[1 + q_2(1 - 2\omega_2)]\,\mathrm{PDF}(...|B) \\ &+ [1 - q_1(1 - 2\bar{\omega}_1)]\,[1 - q_2(1 - 2\bar{\omega}_2)]\,A(q_1, q_2)\mathrm{PDF}(...|\bar{B})\big].\end{aligned} \tag{D.10}$$

However, matching these factors with Equation (D.7) establishes the following relations:

$$A(1,1) = A(1,-1) = A(-1,1) = A(-1,-1) \tag{D.11}$$

$$A(1,0) = A(-1,0) \tag{D.12}$$

$$A(0,1) = A(0,-1) \tag{D.13}$$

$$A(0,0) = \frac{A(1,0)A(0,1)}{A(1,1)}, \tag{D.14}$$

which reduces the independent parameters to the three asymmetry factors $A(1,1)$, $A(1,0)$ and $A(0,1)$. In future, these parameters could be included in the fit and would account for any possible asymmetry in the tagging efficiencies or the production of B mesons.

E Decay-time resolution

Wrong primary vertex component

As described in Section 7.1.2, the decay-time distribution of the sample of prompt fake $B_s^0 \to J/\psi\,\phi$ candidates contains a component that arises mainly from prompt J/ψ mesons that are associated to a wrong primary vertex (PV). In order to correctly extract the decay-time resolution shape this component has to be included in the respective fit. An event mixing technique was employed to determine the decay-time distribution of such wrongly associated J/ψ mesons. For each reconstructed event in the prompt data sample, a PV from another event was added. Then, the reconstruction is repeated with the original PV replaced with the new one. Figure E.1 shows the decay-time distribution of $B_s^0 \to J/\psi\,\phi$ candidates constructed from these J/ψ candidates. A double-sided double-exponential given by:

$$E(t) = f\frac{1}{N_1}e^{-|t|/\tau_1} + (1-f)\frac{1}{N_2}e^{-|t|\tau_2} \tag{E.1}$$

is used to model the observed distribution. Besides the normalization factors N_1 and N_2, it is described by the two decay parameters τ_1 and τ_2 and the relative fraction f between the two exponential functions. A maximum likelihood fit yields the following values for these parameters: $\tau_1 = (0.377 \pm 0.012)$ ps, $\tau_2 = (1.83 \pm 0.11)$ ps and $f = 0.776 \pm 0.018$.

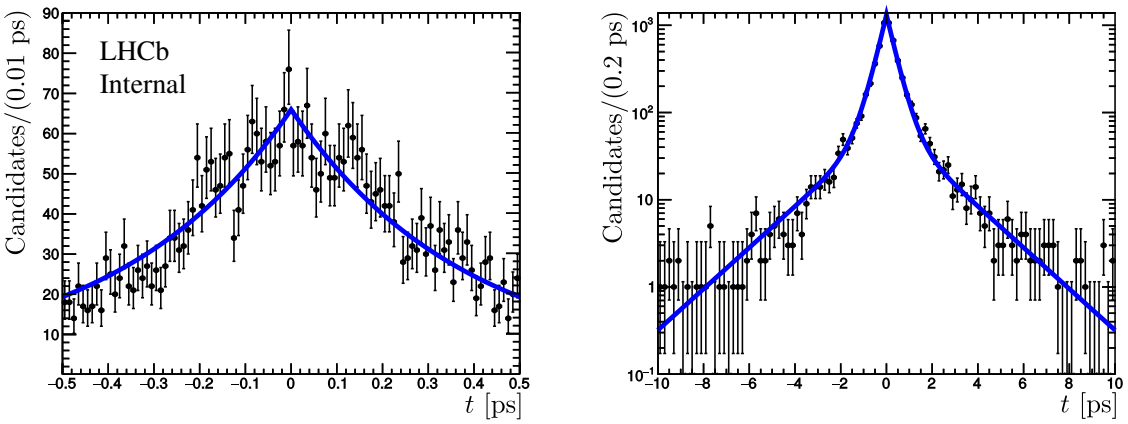


Figure E.1: Shape of the decay time distribution due to wrongly associated primary vertices. The right plot shows a wider decay time range and provides a logarithmic y axis. Figure taken from Ref. [65].

F Fit validation

Effective single Gaussian resolution

As described in Section 8.3, toy studies are performed to validate the usage of an effective single Gaussian instead of a triple Gaussian resolution function. Figure F.1 shows the relevant pull distributions for the S-wave related parameters.

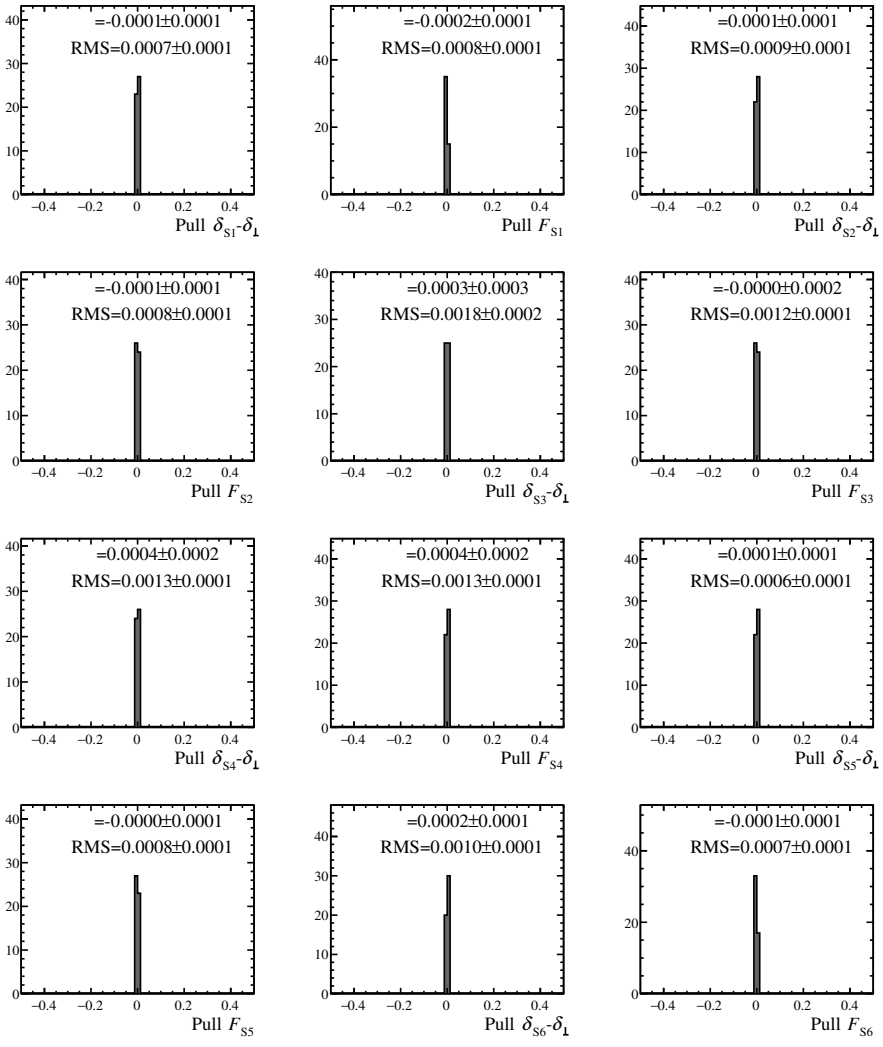


Figure F.1: Pull distributions of the S-wave related physics parameters observed in the toy studies for the single Gaussian resolution. The difference between the fit result obtained with single Gaussian and triple Gaussian resolution model are shown in units of the statistical uncertainty. The mean and the root mean square are given.

G Fit results

G.1 Likelihood scans

Figures G.1 and G.2 show the likelihood profiles for all freely floating parameters in the final decay-time and angular fit. While the main parameters show nearly parabolic, i.e. Gaussian, behavior, some of the strong phases and the S-wave fractions clearly deviate from this shape.

G.2 Correlation matrix

Table G.1 shows the correlations of the main physics parameters how they are obtained from the maximum likelihood fit to the decay-time and angular distributions of the $B_s^0 \to J/\psi\,\phi$ sample.

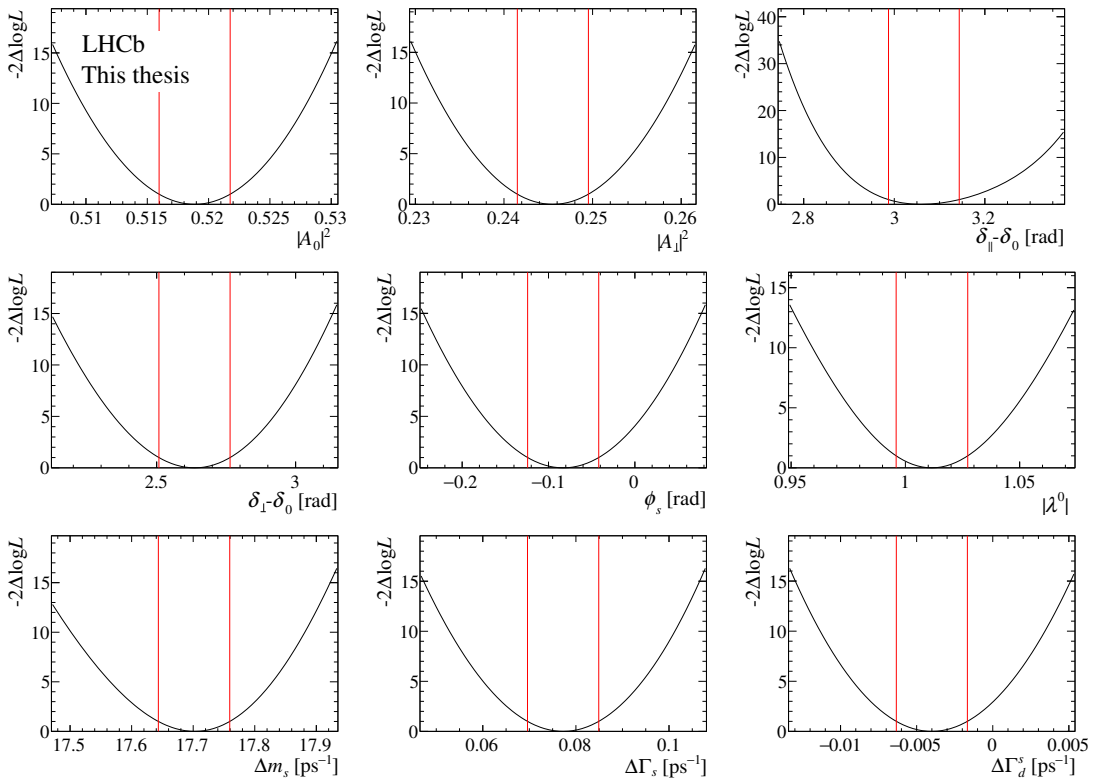


Figure G.1: Likelihood profiles of the main parameters.

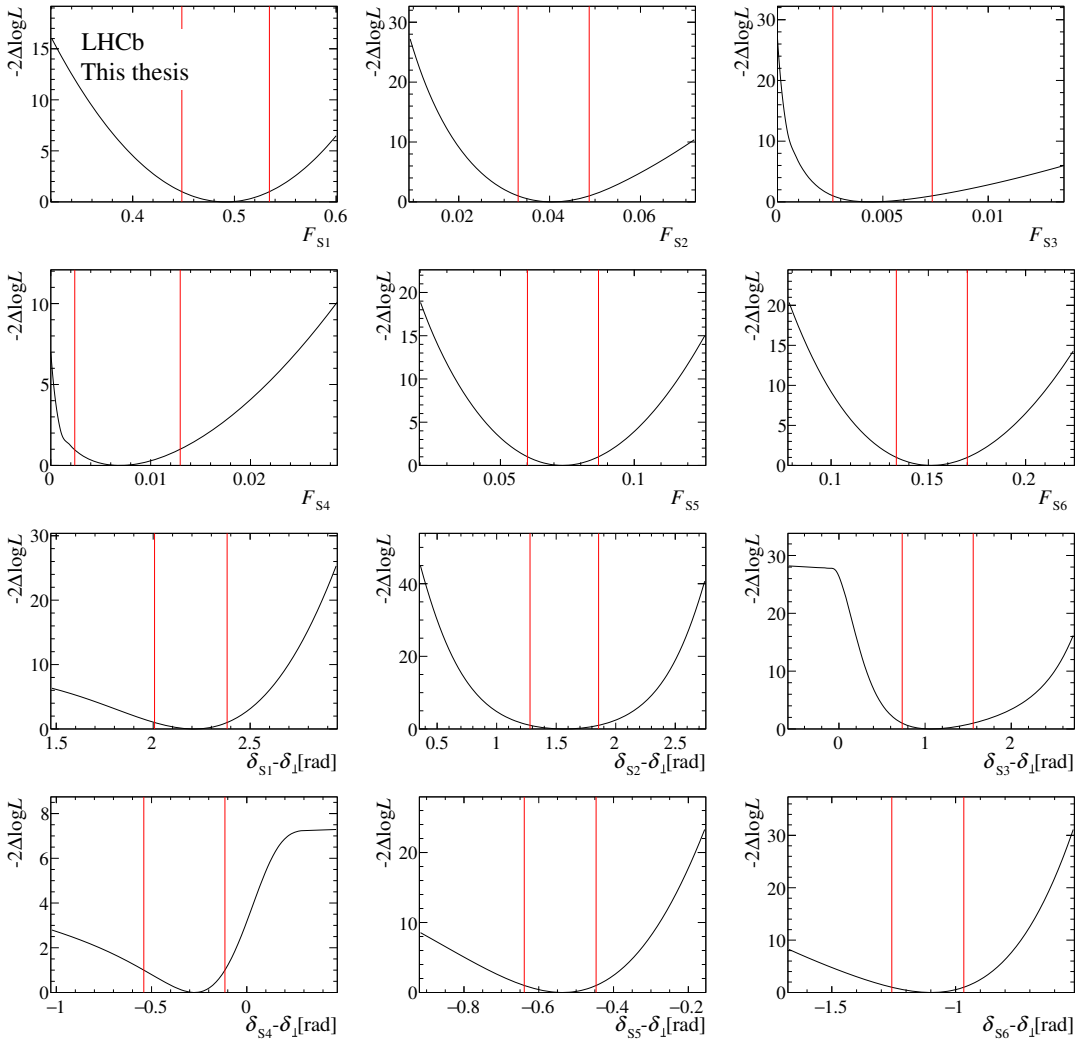


Figure G.2: Likelihood profiles of the parameters related to the S-wave component.

Table G.1: Correlations of the main physics parameters observed in the final time and angular fit.

	ϕ_s	$\lvert\lambda\rvert$	$\Delta\Gamma_s$	Δm_s	$\Delta\Gamma_d^s$	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel - \delta_0$	$\delta_\perp - \delta_0$
ϕ_s	1.00	0.19	0.00	0.00	-0.03	-0.03	0.03	0.01	-0.01
$\lvert\lambda\rvert$		1.00	-0.04	-0.04	0.01	0.02	-0.02	0.01	0.02
$\Delta\Gamma_s$			1.00	0.02	-0.47	-0.69	0.63	-0.03	0.01
Δm_s				1.00	-0.01	-0.03	0.01	-0.05	0.74
$\Delta\Gamma_d^s$					1.00	0.40	-0.32	0.03	0.00
$\lvert A_\perp\rvert^2$						1.00	-0.60	0.16	0.01
$\lvert A_0\rvert^2$							1.00	0.00	0.01
$\delta_\parallel - \delta_0$								1.00	0.20
$\delta_\perp - \delta_0$									1.00

Table G.2: Comparison of the fit results presented in this thesis and the results published by the LHCb collaboration in Ref. [63]. Only statistical uncertainties are shown.

Parameter	This thesis	Official result
$\phi_s\,[\mathrm{rad}]$	-0.083 ± 0.041	-0.083 ± 0.041
$\lvert\lambda\rvert$	1.012 ± 0.016	1.012 ± 0.016
$\Delta\Gamma_s\,[\mathrm{ps}^{-1}]$	$0.0773\pm^{0.0076}_{0.0077}$	0.0773 ± 0.0077
$\Delta m_s\,[\mathrm{ps}^{-1}]$	$17.702\pm^{0.057}_{0.059}$	17.703 ± 0.059
$\Delta\Gamma_d^s\,[\mathrm{ps}^{-1}]$	-0.0040 ± 0.0023	-0.0041 ± 0.0024
$\lvert A_\perp\rvert^2$	0.2455 ± 0.0040	0.2456 ± 0.0040
$\lvert A_0\rvert^2$	0.5189 ± 0.0029	0.5186 ± 0.0029
$\delta_\parallel-\delta_0\,[\mathrm{rad}]$	$3.060\pm^{0.084}_{0.073}$	$3.062\pm^{0.082}_{0.074}$
$\delta_\perp-\delta_0\,[\mathrm{rad}]$	2.64 ± 0.13	2.64 ± 0.13
F_{S1}	0.491 ± 0.043	–
F_{S2}	$0.0406\pm^{0.0081}_{0.0075}$	–
F_{S3}	$0.0044\pm^{0.0029}_{0.0018}$	–
F_{S4}	$0.0069\pm^{0.0061}_{0.0046}$	–
F_{S5}	0.073 ± 0.013	–
F_{S6}	$0.151\pm^{0.019}_{0.018}$	–
$\delta_{S1}-\delta_\perp\,[\mathrm{rad}]$	$2.21\pm^{0.17}_{0.20}$	–
$\delta_{S2}-\delta_\perp\,[\mathrm{rad}]$	1.55 ± 0.29	–
$\delta_{S3}-\delta_\perp\,[\mathrm{rad}]$	$1.07\pm^{0.49}_{0.34}$	–
$\delta_{S4}-\delta_\perp\,[\mathrm{rad}]$	$-0.28\pm^{0.16}_{0.27}$	–
$\delta_{S5}-\delta_\perp\,[\mathrm{rad}]$	$-0.536\pm^{0.090}_{0.103}$	–
$\delta_{S6}-\delta_\perp\,[\mathrm{rad}]$	$-1.10\pm^{0.13}_{0.16}$	–

G.3 Comparison to the official LHCb result

The analysis presented in this thesis is published by the LHCb collaboration, see [63]. As described in Section 3.2, the results shown in Chapter 9 serve as a partial independent cross-check of the official LHCb results. Table G.2 compares the main parameters of the final fit. No significant deviation is observed.

G.4 Asymmetry plot

In the following, the procedure to create Figure 9.2 is presented. In order to obtain the largest possible visible oscillating asymmetry, the tagging, angular and

decay-time resolution information of every candidate has to be taken into account. For every candidate, an angular weight $w_{\text{ang}}(\Omega)$ is calculated according to:

$$w_{\text{ang}}(\Omega) = \frac{(|A_0|^2 f_1(\Omega) + |A_\perp|^2 f_2(\Omega)) - (|A_\parallel|^2 f_3(\Omega) + |A_S|^2 f_7(\Omega))}{(|A_0|^2 f_1(\Omega) + |A_\perp|^2 f_2(\Omega)) + (|A_\parallel|^2 f_3(\Omega) + |A_S|^2 f_7(\Omega))}, \tag{G.1}$$

where Ω represents the three helicity angles and f_i are the angular functions defined in Table 1.3. This weight reflects the amount of even or odd CP component for a given set of angles. Furthermore, a weight is defined to take into account the tagging information. Based on Equations (6.8) and (6.9) the weight is defined as:

$$w_{\text{tag}} = p(B_s^0) - p(\overline{B}_s^0), \tag{G.2}$$

and reflects on one hand the tag decision, but also gives more weight to candidates that have a smaller probability for a wrongly assigned tag. Finally, a weight is defined following the definition of the dilution factor for the decay-time resolution in Equation (7.4):

$$w_{\text{res}} = e^{-\sigma_t^2 \frac{\Delta m_s}{2}}, \tag{G.3}$$

where σ_t is the calibrated decay-time resolution of the respective candidate.

These three weights are multiplied, and the asymmetry between weighted candidates with positive and negative such weights is plotted. Each of these categories is normalized before calculating the time-dependent asymmetry. The projection of the fit is obtained by calculating for every candidate the PDF value as a function of the decay time given the tagging outputs, the decay-time resolution and the helicity angles. These functions are weighted and split in two categories using the same weights as the respective candidate. After adding the functions within one category, an asymmetry is calculated.

H Systematic uncertainties

Tables H.1 to H.3 show the systematic uncertainties on the S-wave related parameters that were discussed in Chapter 10.

Table H.1: Systematic uncertainties for some of the final parameters. For comparison, the respective statistical uncertainty is given in the first row.

Source	F_{S1}	F_{S2}	F_{S3}	F_{S4}	F_{S5}	F_{S6}
Statistical uncertainty	0.043	$^{+0.0081}_{-0.0075}$	$^{+0.0029}_{-0.0018}$	$^{+0.0061}_{-0.0046}$	0.013	$^{+0.019}_{-0.018}$
Mass parametrization	0.005	0.0006	0.0001	0.0008	0.003	0.005
Mass factorization	0.006	0.0006	-	0.0004	0.002	0.003
Multiple candidates	0.002	0.0012	0.0001	0.0007	0.001	0.001
Time res.: prompt	-	-	-	-	-	-
Time res.: mean offset	-	-	-	0.0001	-	-
Time res.: wrong PV	-	0.0001	-	0.0001	-	-
Dec. time acc.: statistical	-	-	-	-	-	-
Dec. time acc.: knot position	-	-	-	-	-	-
Dec. time acc.: PDF weighting	-	-	-	-	-	-
Dec. time acc.: kinematic weighting	-	-	-	-	-	-
Ang. acc.: statistical	0.001	0.0003	0.0001	0.0003	0.001	0.001
Ang. acc.: sim. correction	0.002	0.0003	0.0001	0.0006	0.001	0.001
Ang. acc.: Neglected dependencies	0.006	0.0007	0.0002	0.0001	0.001	0.002
C_{SP} factors	0.001	0.0031	0.0004	0.0004	0.001	0.007
Fit bias	0.001	0.0032	0.0012	0.0009	-	-
Length scale	-	-	-	-	-	-
Simulation truth matching	-	0.0002	0.0001	0.0002	0.001	0.001
Quadratic sum of syst. unc.	0.010	0.0048	0.0013	0.0017	0.004	0.009

Table H.2: Systematic uncertainties for some of the final parameters. For comparison, the respective statistical uncertainty is given in the first row.

Source	$\delta_{S1} - \delta_{\perp}$[rad]	$\delta_{S2} - \delta_{\perp}$[rad]	$\delta_{S3} - \delta_{\perp}$[rad]
Statistical uncertainty	$^{+0.17}_{-0.20}$	0.29	$^{+0.49}_{-0.34}$
Mass parametrization	0.01	-	0.01
Mass factorization	0.08	0.03	0.02
Multiple candidates	0.01	-	0.02
Time res.: prompt	-	-	-
Time res.: mean offset	-	-	-
Time res.: wrong PV	-	-	-
Dec. time acc.: statistical	-	-	-
Dec. time acc.: knot position	-	-	-
Dec. time acc.: PDF weighting	-	-	-
Dec. time acc.: kinematic weighting	-	-	-
Ang. acc.: statistical	-	0.01	0.01
Ang. acc.: sim. correction	0.01	0.02	0.03
Ang. acc.: Neglected dependencies	0.01	0.02	0.03
C_{SP} factors	0.18	0.02	0.03
Fit bias	0.04	0.02	0.08
Length scale	-	-	-
Simulation truth matching	-	-	-
Quadratic sum of syst. unc.	0.20	0.05	0.10

Table H.3: Systematic uncertainties for some of the final parameters. For comparison, the respective statistical uncertainty is given in the first row.

Source	$\delta_{S4} - \delta_{\perp}$[rad]	$\delta_{S5} - \delta_{\perp}$[rad]	$\delta_{S6} - \delta_{\perp}$[rad]
Statistical uncertainty	$^{+0.16}_{-0.27}$	$^{+0.090}_{-0.103}$	$^{+0.13}_{-0.16}$
Mass parametrization	0.01	0.010	0.01
Mass factorization	0.01	0.007	0.03
Multiple candidates	0.01	0.004	0.01
Time res.: prompt	-	0.001	-
Time res.: mean offset	-	0.006	-
Time res.: wrong PV	-	0.001	-
Dec. time acc.: statistical	-	-	-
Dec. time acc.: knot position	-	-	-
Dec. time acc.: PDF weighting	-	-	-
Dec. time acc.: kinematic weighting	-	-	-
Ang. acc.: statistical	0.01	0.004	-
Ang. acc.: sim. correction	0.01	0.005	-
Ang. acc.: Neglected dependencies	0.03	0.010	0.01
C_{SP} factors	0.01	0.010	0.10
Fit bias	0.05	0.010	0.04
Length scale	-	-	-
Simulation truth matching	-	0.002	-
Quadratic sum of syst. unc.	0.06	0.023	0.11

I Final results

Table I.1 shows the measured values with associated statistical and systematic uncertainties for the parameters determined in the maximum likelihood fit to the decay-time and angular distributions.

Table I.1: Fit results for the freely floating parameters of the maximum likelihood fit to the decay-time and angular distributions. In addition, systematic uncertainties for every parameter are given.

Parameter	Fit result	
$\phi_s\,[\mathrm{rad}]$	$-0.083\ \pm 0.041_{\mathrm{stat}}$	$\pm 0.006_{\mathrm{syst}}$
$\lvert\lambda\rvert$	$1.012\ \pm 0.016_{\mathrm{stat}}$	$\pm 0.006_{\mathrm{syst}}$
$\Delta\Gamma_s\,[\mathrm{ps}^{-1}]$	$0.0773\ \pm^{0.0076}_{0.0077\mathrm{stat}}$	$\pm 0.0026_{\mathrm{syst}}$
$\Delta m_s\,[\mathrm{ps}^{-1}]$	$17.702\ \pm^{0.057}_{0.059\mathrm{stat}}$	$\pm 0.018_{\mathrm{syst}}$
$\Delta\Gamma_d^s\,[\mathrm{ps}^{-1}]$	$-0.0040\ \pm 0.0023_{\mathrm{stat}}$	$\pm 0.0015_{\mathrm{syst}}$
$\lvert A_\perp\rvert^2$	$0.2455\ \pm 0.0040_{\mathrm{stat}}$	$\pm 0.0019_{\mathrm{syst}}$
$\lvert A_0\rvert^2$	$0.5189\ \pm 0.0029_{\mathrm{stat}}$	$\pm 0.0024_{\mathrm{syst}}$
$\delta_\parallel - \delta_0\,[\mathrm{rad}]$	$3.060\ \pm^{0.084}_{0.073\mathrm{stat}}$	$\pm 0.040_{\mathrm{syst}}$
$\delta_\perp - \delta_0\,[\mathrm{rad}]$	$2.64\ \pm 0.13_{\mathrm{stat}}$	$\pm 0.10_{\mathrm{syst}}$
F_{S1}	$0.491\ \pm 0.043_{\mathrm{stat}}$	$\pm 0.010_{\mathrm{syst}}$
F_{S2}	$0.0406\ \pm^{0.0081}_{0.0075\mathrm{stat}}$	$\pm 0.0048_{\mathrm{syst}}$
F_{S3}	$0.0044\ \pm^{0.0029}_{0.0018\mathrm{stat}}$	$\pm 0.0013_{\mathrm{syst}}$
F_{S4}	$0.0069\ \pm^{0.0061}_{0.0046\mathrm{stat}}$	$\pm 0.0017_{\mathrm{syst}}$
F_{S5}	$0.073\ \pm 0.013_{\mathrm{stat}}$	$\pm 0.004_{\mathrm{syst}}$
F_{S6}	$0.151\ \pm^{0.019}_{0.018\mathrm{stat}}$	$\pm 0.009_{\mathrm{syst}}$
$\delta_{S1} - \delta_\perp\,[\mathrm{rad}]$	$2.21\ \pm^{0.17}_{0.20\mathrm{stat}}$	$\pm 0.20_{\mathrm{syst}}$
$\delta_{S2} - \delta_\perp\,[\mathrm{rad}]$	$1.55\ \pm 0.29_{\mathrm{stat}}$	$\pm 0.05_{\mathrm{syst}}$
$\delta_{S3} - \delta_\perp\,[\mathrm{rad}]$	$1.07\ \pm^{0.49}_{0.34\mathrm{stat}}$	$\pm 0.10_{\mathrm{syst}}$
$\delta_{S4} - \delta_\perp\,[\mathrm{rad}]$	$-0.28\ \pm^{0.16}_{0.27\mathrm{stat}}$	$\pm 0.06_{\mathrm{syst}}$
$\delta_{S5} - \delta_\perp\,[\mathrm{rad}]$	$-0.536\ \pm^{0.090}_{0.103\mathrm{stat}}$	$\pm 0.023_{\mathrm{syst}}$
$\delta_{S6} - \delta_\perp\,[\mathrm{rad}]$	$-1.10\ \pm^{0.13}_{0.16\mathrm{stat}}$	$\pm 0.11_{\mathrm{syst}}$

Bibliography

[1] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[2] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[3] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[4] Planck, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209.

[5] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801, arXiv:1411.3104.

[6] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^+K^-$ decays*, Phys. Lett. **B767** (2017) 177, arXiv:1610.09476.

[7] S. S. for the LHCb collaboration, *Fast, Parallel and Parametrized Kalman Filters for the LHCb upgrade*, Connecting The Dots / Intelligent Trackers 2017, 3, 2017. URL: https://indico.cern.ch/event/577003/contributions/2476577/, last visited on 03/05/2019.

[8] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[9] E. Noether, *Invariante Variationsprobleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235.

[10] T. D. Lee and C. N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104** (1956) 254.

[11] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[12] L. D. Landau, *On the conservation laws for weak interactions*, Nucl. Phys. **3** (1957) 127.

[13] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[14] M. Kobayashi and T. Maskawa, *CP-Violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[15] U. Nierste, *Three Lectures on Meson Mixing and CKM phenomenology*, in *Heavy quark physics. Proceedings, Helmholtz International School, HQP08, Dubna, Russia, August 11-21, 2008*, pp. 1–38, 2009, arXiv:0904.1869.

[16] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895, arXiv:1612.07233, updated results and plots available at https://hflav.web.cern.ch, last visited on 06/05/2019.

[17] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the B_s^0–$\overline{B}_s^0$ oscillation frequency in the decay $B_s^0 \to D_s^- \pi^+$*, New J. Phys. **15** (2013) 053021, arXiv:1304.4741.

[18] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in B_s^0–$\overline{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, arXiv:1605.09768.

[19] CKMfitter group, J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ new physics*, Phys. Rev. **D91** (2015) 073007, arXiv:1501.05013, updated results and plots available at http://ckmfitter.in2p3.fr/, last visited on 10/04/2019.

[20] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi \pi^+\pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600.

[21] Y. Xie, P. Clarke, G. Cowan, and F. Muheim, *Determination of $2\beta_s$ in $B_s^0 \to J/\psi K^+K^-$ Decays in the Presence of a K^+K^- S-Wave Contribution*, JHEP **09** (2009) 074, arXiv:0908.3627.

[22] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in $\overline{B}_s^0 \to J/\psi \pi^-\pi^-$ decays*, Phys. Lett. **B736** (2014) 186, arXiv:1405.4140.

[23] ATLAS, G. Aad *et al.*, *Measurement of the CP-violating phase ϕ_s and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi\phi$ decays in ATLAS*, JHEP **08** (2016) 147, arXiv:1601.03297.

[24] CMS, V. Khachatryan *et al.*, *Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B_s^0 \to J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B757** (2016) 97, arXiv:1507.07527.

[25] M. Artuso, G. Borissov, and A. Lenz, *CP violation in the B_s^0 system*, Rev. Mod. Phys. **88** (2016) 045002, arXiv:1511.09466.

[26] K. De Bruyn and R. Fleischer, *A Roadmap to Control Penguin Effects in $B_d^0 \to J/\psi K_S^0$ and $B_s^0 \to J/\psi\phi$*, JHEP **03** (2015) 145, arXiv:1412.6834.

[27] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters and polarisation fractions in $B_s^0 \to J/\psi\,\overline{K}^{*0}$ decays*, JHEP **11** (2015) 082, arXiv:1509.00400.

[28] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015) 073007, arXiv:1501.05013.

[29] C.-W. Chiang *et al.*, *New Physics in $B_s^0 \to J/\psi\,\phi$: A General Analysis*, JHEP **04** (2010) 031, arXiv:0910.2929.

[30] X. Liu, W. Wang, and Y. Xie, *Penguin pollution in $B \to J/\psi V$ decays and impact on the extraction of the $B_s^0 - \overline{B}_s^0$ mixing phase*, Phys. Rev. **D89** (2014) 094010, arXiv:1309.0313.

[31] A. Lenz, *Lifetimes and heavy quark expansion*, Int. J. Mod. Phys. **A30** (2015) 1543005, arXiv:1405.3601.

[32] M. Kirk, A. Lenz, and T. Rauh, *Dimension-six matrix elements for meson mixing and lifetimes from sum rules*, JHEP **12** (2017) 068, arXiv:1711.02100.

[33] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[34] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[35] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[36] CMS, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[37] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[38] R. Lindner, *LHCb layout 2/LHCb schema 2*, LHCb Collection, URL: `http://cdsweb.cern.ch/record/1087860`, last visited on 10/04/2019.

[39] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, URL: `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`, last visited on 11/04/2019.

[40] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011.

[41] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024.

[42] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029.

[43] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007.

[44] LHCb, *LHCb webpage*, URL: `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html`, last visited on 10/04/2019.

[45] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, Tech. Rep. LHCb-PUB-2011-001, CERN, Geneva, Jan, 2011. LHCb.

[46] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[47] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030.

[48] LHCb, *LHCb public webpage*, URL: `https://lhcb-public.web.cern.ch/lhcb-public/en/Detector/Trackers2-en.html`, last visited on 10/04/2019.

[49] G. A. Cowan, *Performance of the lhcb silicon tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. Proceedings of the 8th International Hiroshima Symposium on the Development and Application of Semiconductor Tracking Detectors **699** (2013) 156 .

[50] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[51] LHCb RICH, J. He, *Real-time calibration and alignment of the LHCb RICH detectors*, Nucl. Instrum. Meth. **A876** (2017) 13, arXiv:1611.00296.

[52] LHCb RICH, A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, Nucl. Instrum. Meth. **A876** (2017) 221, arXiv:1703.08152.

[53] J. R. Harrison, *Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays*, PhD thesis, Manchester U., 2014, CERN-THESIS-2014-068.

[54] 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.

[55] B. Sciascia, *LHCb Run 2 Trigger Performance*, PoS **BEAUTY2016** (2016) 029. 7.

[56] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The LHCb Turbo Stream*, Journal of Physics: Conference Series **664** (2015) 082004.

[57] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007.

[58] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016.

[59] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[60] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[61] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[62] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[63] LHCb collaboration, R. Aaij *et al.*, *Updated measurement of time-dependent CP-violating observables in $B_s^0 \to J/\psi K^+K^-$ decays*, LHCb-PAPER-2019-013, in preparation, to be submitted to JHEP.

[64] E. G. for the LHCb collaboration, *Mixing and time-dependant CP-Violation in beauty at LHCb*, Rencontres de Moriond EW, 3, 2019. URL: `http://moriond.in2p3.fr/2019/EW/Program.html`, last visited on 03/05/2019.

[65] S. Benson *et al.*, *Measurements of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays in the low K^+K^- mass range with 13 TeV data*, , LHCb-ANA-2017-028.

[66] R. E. Schapire, *The strength of weak learnability*, Machine Learning **5** (1990) 197.

[67] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, The Annals of Statistics **29** (2001) 1189.

[68] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040.

[69] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016) , `arXiv:1608.05806`, URL: `https://github.com/arogozhnikov/hep_ml`, last visited on 03/05/2019.

[70] G. Cowan, *Statistical data analysis*, Oxford University Press, USA, 1998.

[71] L. Kish, *Survey sampling*, Wiley classics library, J. Wiley, 1965.

[72] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.

[73] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[74] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[75] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[76] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.

[77] 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.

[78] D. Fazzini, *Flavour Tagging in the LHCb experiment*, in *Proceedings, 6th Large Hadron Collider Physics Conference (LHCP 2018): Bologna, Italy, June 4-9, 2018*, vol. LHCP2018, p. 230, 2018. doi: 10.22323/1.321.0230.

[79] LHCb collaboration, R. Aaij *et al.*, *B flavour tagging using charm decays at the LHCb experiment*, JINST **10** (2015) P10005, arXiv:1507.07892.

[80] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.

[81] LHCb collaboration, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, arXiv:1602.07252.

[82] LHCb collaboration, R. Aaij *et al.*, *Measurements of the B^+, B^0, B_s^0 meson and Λ_b^0 baryon lifetimes*, JHEP **04** (2014) 114, arXiv:1402.2554.

[83] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[84] E. Cohen, R. F. Riesenfeld, and G. Elbe, *Geometric Modeling with Splines: An Introduction*, A. K. Peters, Ltd., Natick, MA, USA, 2001.

[85] C. De Boor, *A practical guide to splines; rev. ed.*, Applied mathematical sciences, Springer, Berlin, 2001.

[86] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, arXiv:1407.0748.

[87] G. A. Cowan, D. C. Craik, and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **214** (2017) 239, arXiv:1612.07489.

[88] V. Baru *et al.*, *Flatte-like distributions and the $a_0(980)/f_0(980)$ mesons*, Eur. Phys. J. **A23** (2005) 523, arXiv:nucl-th/0410099.

[89] LHCb collaboration, R. Aaij *et al.*, *Resonances and CP-violation in $\overline{B}{}^0_s$ and $B^0_s \to J/\psi K^+K^-$ decays in the mass region above the $\phi(1020)$*, JHEP **08** (2017) 037, arXiv:1704.08217.

[90] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Statist. **9** (1938) 60.

[91] LHCb collaboration, R. Aaij *et al.*, *Measurement of resonant and CP components in $\overline{B}{}^0_s \to J/\psi \pi^+\pi^-$ decays*, Phys. Rev. **D89** (2014) 092006, arXiv:1402.6248.

[92] LHCb, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s from $B^0_s \to J/\psi\pi^+\pi^-$ decays in 13 TeV pp collisions*, Submitted to: Phys. Lett. B (2019) arXiv:1903.05530.

[93] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP violating phase and decay-width difference in $B^0_s \to \psi(2S)\phi$ decays*, Phys. Lett. **B762** (2016) 253, arXiv:1608.04855.

[94] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in $\overline{B}{}^0_s \to D^+_s D^-_s$ decays*, Phys. Rev. Lett. **113** (2014) 211801, arXiv:1409.4619.

[95] ATLAS Collaboration, *Measurement of the CP violation phase ϕ_s in $B_s \to J/\psi\phi$ decays in ATLAS at 13 TeV*, Tech. Rep. ATLAS-CONF-2019-009, CERN, Geneva, Mar, 2019.

Acknowledgments

This thesis would have not been possible without the help of many people. First of all I have to thank my supervisor Prof. Stephanie Hansmann-Menzemer. She gave me the chance to start with my PhD in the LHCb group in Heidelberg and introduced me to the topic presented in this thesis. I am grateful for her trust in my work that gave me a lot of freedom in what I did. On the other side I profited a lot from her valuable feedback, no matter if this was on analyses or on the way how to present them. She also promoted me in my application for the International Max Planck Research School for Precision Tests of Fundamental Symmetries, which supported me for most of the time of my PhD studies. I also thank Prof. Ulrich Uwer as my second supervisor for his valuable comments during the regular LHCb meetings we had in Heidelberg.

Next, I want to thank the members of my thesis committee, Prof. Ulrich Glasmacher, Prof. Tilman Plehn and especially Prof. Klaus Reygers who agreed to be the second referee for my thesis.

I also thank Dr. Michel De Cian and Dr. Sevda Esen who supported me a lot during my time in the LHCb Heidelberg group. Besides the analysis presented here, I also developed a fast Kalman filter for the LHCb upgrade and Michel always provided extremely valuable guidance in these studies. Sevda I want to thank for her patience when I started to work on what is presented in this thesis. During that time I had a lot of questions, and it was of great help to have her in Heidelberg.

An analysis as presented here is hardly feasible without a collaboration of many people contributing to it. I had the fortune to be part of a team that always gave me the feeling to be a valuable member of it. Although also other people have contributed, I want to express special thanks to Diego, Francesca, Greig, Jennifer, Katya, Konstantin, Liming, Miriam, Sevda, Veronika and Wenhua for their always fruitful collaboration.

At this point, I would also like to thank all members of the LHCb group in Heidelberg. I am very grateful to have met all these nice people and enjoyed the time a lot.

Last but for sure not least I thank my family and my friends who always supported me and also showed interest for what I was doing. In this context, I am particular thankful to my wife Marissa for her trustfulness in my decision to start my PhD.