MEASUREMENT OF CP VIOLATION IN MIXING AND DECAY OF STRANGE BEAUTY MESONS

Jeroen van Leerdam

Measurement of CP Violation in Mixing and Decay of Strange Beauty Mesons

Jeroen van Leerdam

Cover design by Marcel Buunk

Set in Linux Libertine using LaTeX
Printed by Gildeprint – The Netherlands
ISBN 978-94-6233-263-8

This work is part of the research programme of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). It was carried out at the National Institute for Subatomic Physics (Nikhef) in Amsterdam, the Netherlands.

VRIJE UNIVERSITEIT

Measurement of CP Violation in Mixing and Decay of Strange Beauty Mesons

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Exacte Wetenschappen
op woensdag 18 mei 2016 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Jeroen van Leerdam

geboren te Doesburg

promotor:	prof.dr. H.G. Raven
copromotor:	prof.dr. M.H.M. Merk

for Aukje

Acknowledgements

Many people have, directly or indirectly, contributed to my PhD thesis, for which I am grateful. I was given the opportunity to do both my master's and doctoral research at Nikhef, where I always received all the support I needed. My thesis could not have been written without the help from my advisors, collaboration with people in our group at Nikhef and the LHCb experiment, all the people that made it possible to study particles from LHC collisions with data from the LHCb detector, many discussions about physics and various other topics at Nikhef and Cern, people proofreading the thesis draft, and everyone who provided the necessary moral support. Thank you all.

Contents

Chapter 1

Introduction

The main objective of the work presented in this thesis is the measurement of CP violation in the decay $B_s^0 \to J/\psi\,\phi$. This measurement is a test of the current model of elementary-particle interactions. It is performed within the LHCb experiment, which studies particles produced in the proton–proton collisions of the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland. The results presented here (see also reference [1,2]) are an update of previous LHCb results [3] with a dataset that is more than three times larger.

Motivations for such a test of particle-physics models are discussed in Section 1.1. The measurement is introduced in Sections 1.3 and 1.4, after an introduction to CP violation in this context in Section 1.2.

Details of the phenomenology of the $B_s^0 \to J/\psi\,\phi$ decay are discussed in Chapter 2. A discussion of the measurement and the applied data-analysis techniques can be found in Chapter 3. Finally, results of the measurement are presented in Chapter 4.

1.1 The Standard Model of Particle Physics

1.1.1 Elementary-Particle Interactions

Interactions between quarks and leptons are described by the *Standard Model of Particle Physics* (Standard Model, SM). The Standard Model unifies the electromagnetic and weak forces in a theory of *electroweak interactions* [4], while strong interactions are described separately by *Quantum Chromodynamics* (QCD) [5]. Both theories are based on quantum field theory, in which quarks

and leptons are described as fermionic fields that interact via bosonic fields.

In the electroweak model there are four vector bosons that mediate interactions, three of which acquire a mass through the *Brout-Englert-Higgs* (BEH) *mechanism* [6]. Generating particle masses with this mechanism implies the existence of the scalar *Higgs boson*, which was recently discovered by the ATLAS and CMS experiments at the LHC [7]. The BEH mechanism is believed to be also responsible for generating the masses of quarks and leptons through so-called Yukawa interactions.

The three massive vector bosons are the W^+, the W^- and the Z^0. These particles are the carriers of the weak interaction. The remaining massless vector boson is the photon, which mediates the electromagnetic interaction.

Contrary to leptons, quarks carry the "charge" of the strong force, or *colour*. This makes them subject to strong interactions, described by QCD. The mediators of the strong interaction are gluons, which come in eight different types and are themselves colour-charged particles.

One of the features of the strong force is that it becomes weaker with increasing interaction energy. This phenomenon is called *asymptotic freedom* [8]. The increasing interaction strength with decreasing energy (or, equivalently, increasing interaction distance) makes it impossible to completely separate coloured particles. As a result, quarks and gluons are *confined* within colour-neutral objects, such as hadrons.

Low-energy QCD interactions between quarks and gluons within hadrons are so strong that they cannot be described with *perturbation theory*, where amplitudes are expanded as a series in the coupling strength of the interaction. The series would diverge and the exact solution of a calculation is not approximated by the leading terms in the series. Numerical methods may be applied to treat low-energy QCD interactions in an alternative way. Strong interactions at high energies and electroweak interactions can be described with perturbative methods.

Decays of hadrons that contain heavy quarks can be approximately described using a *factorization* approach. In the case of B_s^0 mesons, for example, the decay of the constituent b quark is treated separately from the strong interactions among the quarks in the B_s^0 meson and its decay products. The decay takes place at a relatively high energy scale set by the b-quark mass and is calculated perturbatively. The strong hadronic interactions are less energetic and non-perturbative.

For the decay of a (heavy) quark or lepton a transition to one of the lighter

quarks or leptons is required. In the Standard Model, such a *flavour change* is only possible in a weak interaction that is mediated by the $W^{\pm}$ boson [9–11]. This mixing of quark flavours in the context of meson decays is discussed in more detail in Section 1.2.1.

Another interesting property of the weak interaction is that it only couples to left-handed particle states and to right-handed antiparticle states. As a result, two transformations are needed to obtain the equivalent antiparticle interaction for a given particle interaction: A *charge-conjugation* operation (C) transforms particles into the corresponding antiparticles and a *parity* operation (P) inverts spatial coordinates, which transforms left-handed states into right-handed states.

Most interactions in the Standard Model are invariant under the combined C and P operation, which makes matter and antimatter almost symmetric. The only source of CP-symmetry violation is the flavour-changing weak interaction. Measurements of the amount of CP violation provide tests of the Standard Model description of particle interactions, which is discussed in Section 1.2.2.

In addition to the charge and parity operations there is the transformation of *time reversal* (T), which inverts the time coordinate. Quantum field theory is based on the principle of Lorentz invariance, which implies invariance under a combined C, P, and T transformation for any interaction [12]. CPT is assumed to be an exact symmetry in the studies presented in this thesis.

1.1.2 Beyond the Standard Model

Since the start of its construction in the 1960s, the Standard Model has been tested extensively by experiments. Interactions of elementary particles have been studied over a wide range of energies, both with particles that naturally occur and with particles created in laboratories. The Standard Model has proven to be a consistent description of particle physics.

Despite its success, there are strong motivations for developing descriptions of particle physics that go beyond the Standard Model. Some of these motivations are theoretical, while others are based on experimental observations.

The Standard Model contains 18 parameters, for which it predicts no values. There are nine unpredicted quark and lepton masses, the mass and vacuum expectation value of the Higgs field, four quark-mixing parameters (see Section 1.2.1), and three interaction coupling constants. These parameters

have been measured, but it would be more satisfying to have a more fundamental description without unpredicted quantities.

Another missing feature in the Standard Model is unification of all forces. While the electric and magnetic forces are unified in electromagnetism and the electromagnetic and weak forces in electroweak theory, the Standard Model does not attempt to unify the electroweak and strong forces. The gravitational force is left out completely. Although not strictly necessary to describe the physics of current experiments, the knowledge of how to describe all these phenomena consistently would significantly improve our understanding of nature.

There is a more practical theoretical issue known as the *hierarchy problem*. The mass of the Higgs boson is affected by quantum corrections from loops of other particles. These corrections are large and with only the particles in the Standard Model one would expect the Higgs mass to be much larger than the measured value [13]. The only way to get around this within the Standard Model is a careful fine-tuning of its parameters, which is considered to be unnatural.

Several experimental observations also indicate that the current Standard Model is incomplete. Experiments studying neutrinos from various sources have observed transitions between neutrinos from different generations [14]. This implies mixing between leptons of different flavours, but also non-zero neutrino masses. Neutrinos are massless in the original Standard Model, but in principle their masses can be included by a minimal extension, assuming neutrino masses have the same origin as quark and charged-lepton masses.

There is also a more exciting possibility to include neutrino masses. Since all quantum numbers of neutrinos and antineutrinos are equal, neutrinos may be their own antiparticles, which could lead to Majorana mass components [15]. This scenario would open up new possibilities to explain why neutrino masses are so much smaller than quark and charged-lepton masses. The masses of the left-handed Standard Model neutrinos could be suppressed by the very large masses of hypothetical right-handed Majorana neutrinos via a so-called *see-saw mechanism* [16].

A related observation is that the universe contains matter, but almost no antimatter. One of the required conditions to create such an imbalance is a sufficient amount of CP violation in particle interactions [17]. CP violation in the Standard Model is believed to be too small to generate the large matter–antimatter asymmetry that is observed [18].

The right-handed neutrinos introduced by a see-saw mechanism would also provide a natural way to introduce the required additional CP violation in particle interactions. Couplings of charged leptons to these neutrinos would in general be CP violating, which leads to a lepton–antilepton asymmetry in neutrino decays. This asymmetry for leptons could subsequently lead to an asymmetry for baryons [19].

There are other cosmological observations that also suggest the need for extensions of the Standard Model. If measurements of the spatial fluctuations in the cosmic microwave background are interpreted with the current cosmological models, the bulk of the matter in the universe has an unknown origin [20]. Less than a fifth is conventional matter, which consists of Standard Model particles. The remainder interacts neither strongly nor electromagnetically and is therefore termed *dark matter*. The Standard Model does not provide any particles that would be viable candidates to form dark matter.

Besides see-saw mechanisms, there are many other ideas on how to extend the Standard Model, or even to find a more fundamental theory. A popular candidate is *Supersymmetry* [21], which leads to a class of models to be tested by experiments. Supersymmetry is based on an additional symmetry between bosons and fermions, which would give all particles in the Standard Model a so-called "superpartner". Introducing these additional particles could solve the hierarchy problem, may provide a dark-matter candidate, and could be a first step towards unification of the electroweak and strong forces. See reference [22] for a recent review of the status of Supersymmetry.

To make progress in the search for a more complete theory, observations of particle interactions that cannot be described with the Standard Model are needed. There are two different approaches to search for such inconsistencies. One can assume a particular beyond-the-Standard Model theory and test its predictions in a measurement. An example of this *top-down* approach is to search for signs of the additional particles predicted by Supersymmetry. The other method is a *bottom-up* search, where the Standard Model predictions are tested instead.

In the bottom-up approach one generally searches for small deviations in predictions for well-known processes. An example is the work presented here, which is a study of the variables in the decay of a Standard Model particle, the B_s^0 meson. This particle is produced abundantly in the proton–proton collisions of the LHC and its decays have been studied previously by other collider experiments. One of the particularly interesting decays is that into a

J/ψ and a $\phi(1020)$ meson, which could exhibit CP violation [23].

CP violation in the $B_s^0 \to J/\psi\,\phi$ decay[1] as predicted by the Standard Model is very small. Studies of various extensions of the Standard Model show that it can be significantly enhanced by introducing new contributions to this process [24, 25]. The decay is also experimentally accessible, which makes it an excellent tool to search for deviations from the Standard Model prediction.

The B_s^0 meson and the $B_s^0 \to J/\psi\,\phi$ decay will be further introduced in Section 1.3. First an overview of CP violation within the context of the Standard Model is given in the next section.

1.2 Quark-Flavour Physics and CP Violation

1.2.1 Quark Mixing

The Feynman diagrams for the charged weak current that changes the flavour of a quark are shown in Figure 1.1. The subscript "L" on the quark indicates that only quark states with a left-handed chirality participate in this interaction.

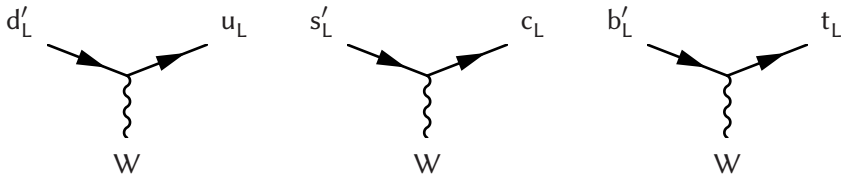


Figure 1.1: *Charged-current weak interactions of quarks.*

The primes on the d'_L, s'_L and b'_L down-type quarks in the figure indicate that the states that couple to the W boson are not the quark mass eigenstates. The down-type states in the interaction are given by a rotation of the mass eigenstates, which is described by the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [11]:

$$\begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} \equiv V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} . \tag{1.1}$$

[1] The symbol "ϕ" will be understood to mean "$\phi(1020)$" in this context.

Diagrams with W-boson interactions in which only the quark mass eigenstates occur are proportional to the corresponding matrix elements V_{ij}. A similar mechanism applies for mixing of leptons [26].

The CKM matrix is a unitary complex matrix. The number of independent parameters in the matrix is limited by the constraint of unitarity and by the fact that part of the complex phases of the matrix elements can be absorbed in the arbitrary phases of the quark fields. The representation of the CKM matrix in terms of the four remaining parameters is convention dependent. A commonly used choice is the *Wolfenstein parameterization* with real-valued V_{ud}, V_{us}, V_{cb}, and V_{tb} and a single complex phase entering in the other elements [27]. Four real parameters, λ, A, ρ, and η, are then defined by

$$\lambda \equiv \frac{|V_{\text{us}}|}{\sqrt{|V_{\text{ud}}|^2 + |V_{\text{us}}|^2}} \quad A\lambda^2 \equiv \frac{|V_{\text{cb}}|}{\sqrt{|V_{\text{ud}}|^2 + |V_{\text{us}}|^2}} \quad A\lambda^3(\rho + i\eta) \equiv V_{\text{ub}}^* \ . \tag{1.2}$$

The Wolfenstein parameterization of the CKM matrix is motivated by the orders of magnitude of matrix elements. The magnitudes of the diagonal elements V_{ud}, V_{cs} and V_{tb}, which describe the coupling between up-type and down-type quarks of the same generation, are approximately equal to one. Magnitudes of the couplings between the first and second generation are between a factor four and five smaller: $|V_{\text{us}}| \approx |V_{\text{cd}}| \approx \lambda \approx 0.23$ [28, 29].

Couplings between the second and third and between the first and third generation are suppressed by factors λ^2 and λ^3, respectively: $|V_{\text{cb}}| \approx |V_{\text{ts}}| \approx A\lambda^2 \approx 0.04$, $|V_{\text{ub}}| = A\lambda^3|\rho + i\eta| \approx 0.004$, and $|V_{\text{td}}| \approx A\lambda^3|1 - \rho - i\eta| \approx 0.009$. In this form of the Wolfenstein parameterization the complex phase is introduced with the parameters $\rho \approx 0.13$ and $\eta \approx 0.36$, where $\arg(V_{\text{ub}}^*) = \arg(\rho + i\eta) \approx 70°$.

The CKM matrix can be expanded in terms of the small parameter λ. Neglecting terms of order λ^4 and higher relative to the leading term for each element, it is approximated by [28]

$$V_{\text{CKM}} \approx \begin{pmatrix} c & \lambda & A\lambda^3\, r^* \\ -\lambda & c & A\lambda^2 \\ A\lambda^3(1 - c\, r) & -A\lambda^2(c + \lambda^2\, r) & 1 \end{pmatrix} , \tag{1.3}$$

where $c \equiv 1 - \frac{1}{2}\lambda^2$ and $r \equiv \rho + i\eta$.

Notice that only the elements V_{ub}, V_{td} and V_{ts} have non-vanishing imaginary parts in this approximation. In V_{ts} the parameter r is suppressed by

a factor λ^2, which makes the imaginary part of this element much smaller than the real part. For V_{cd} and V_{cs} the suppression factors are λ^4 and λ^6, respectively, which results in vanishing imaginary parts.

1.2.2 CP Violation

The Standard Model mixing formalism was introduced for quarks, but it also applies to antiquarks. A CP operation on the states in Figure 1.1 results in the interactions of right-handed antiquark states with the W boson.

In the transformation from quarks to antiquarks the CKM-matrix elements in the weak interaction states are replaced with their complex conjugates. If there were no complex phase in the CKM matrix, its elements would be real and the weak interactions of quarks would be invariant under the CP transformation. Introducing a complex phase breaks this invariance and gives rise to CP violation. The CKM matrix is the only source of CP violation in the Standard Model with massless neutrinos.

Although the complex phase in the CKM matrix formally breaks CP invariance, CP-violating phenomena cannot be observed directly in processes that are described by a single amplitude with W-boson interactions. The magnitude of an amplitude is squared to obtain the corresponding observable probability, which makes its phase unobservable.

CP violation can only be observed in processes where two or more amplitudes with different CKM elements interfere. The CKM elements then introduce a different *weak phase* for each contribution, which changes sign between CP-conjugate processes. Differences between the weak phases do affect the observable magnitude of the total amplitude, provided that the contributing amplitudes also have different *strong phases*, which do not change sign under a CP transformation.

This can be illustrated by considering two interfering amplitudes A_1 and A_2 with CKM factors F_1 and F_2. The asymmetry between the squared magnitudes for CP-conjugate processes is then given by

$$\begin{aligned} A_{\mathrm{CP}} &\equiv \frac{|\mathcal{A}|^2 - |\overline{\mathcal{A}}|^2}{|\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2} = \frac{|F_1\,A_1 + F_2\,A_2|^2 - |F_1^*\,A_1 + F_2^*\,A_2|^2}{|F_1\,A_1 + F_2\,A_2|^2 + |F_1^*\,A_1 + F_2^*\,A_2|^2} \\ &= \frac{-2R\,\sin(\Delta\delta)\,\sin(\Delta\phi)}{1 + R^2 + 2R\,\cos(\Delta\delta)\,\cos(\Delta\phi)}\,, \end{aligned} \tag{1.4}$$

where $R \equiv \frac{|F_2|\,|A_2|}{|F_1|\,|A_1|}$, $\Delta\delta \equiv \arg(A_2) - \arg(A_1)$ and $\Delta\phi \equiv \arg(F_2) - \arg(F_1)$. Notice that the CP asymmetry vanishes if there is no weak phase difference

$\Delta\phi$, but also if there is no strong phase difference $\Delta\delta$. The size of the asymmetry also depends on the relative magnitudes of the amplitudes and CKM factors. It vanishes in case the product $|F_2|\,|A_2|$ is much smaller than its counterpart $|F_1|\,|A_1|$.

To construct convention-independent measures of CP violation in quark interactions, the CKM-matrix unitarity constraints are used. The unitarity relation $V_{\text{CKM}}^{\dagger}V_{\text{CKM}} = V_{\text{CKM}}V_{\text{CKM}}^{\dagger} = \mathbb{1}$ gives nine constraints. Six of these are orthogonality relations, two of which are given by

$$\text{“d–b”:}\quad V_{\text{ud}}V_{\text{ub}}^{*} + V_{\text{cd}}V_{\text{cb}}^{*} + V_{\text{td}}V_{\text{tb}}^{*} = 0$$
$$\Rightarrow\ 1 + \frac{V_{\text{ud}}V_{\text{ub}}^{*}}{V_{\text{cd}}V_{\text{cb}}^{*}} + \frac{V_{\text{td}}V_{\text{tb}}^{*}}{V_{\text{cd}}V_{\text{cb}}^{*}} = 0 \tag{1.5a}$$
$$\text{“s–b”:}\quad V_{\text{us}}V_{\text{ub}}^{*} + V_{\text{cs}}V_{\text{cb}}^{*} + V_{\text{ts}}V_{\text{tb}}^{*} = 0$$
$$\Rightarrow\ 1 + \frac{V_{\text{us}}V_{\text{ub}}^{*}}{V_{\text{cs}}V_{\text{cb}}^{*}} + \frac{V_{\text{ts}}V_{\text{tb}}^{*}}{V_{\text{cs}}V_{\text{cb}}^{*}} = 0\,. \tag{1.5b}$$

Notice that the phases of the combinations of four CKM elements that appear in Equation 1.5 are of the form

$$\arg\left(\frac{V_{ij}V_{il}^{*}}{V_{kl}V_{kj}^{*}}\right) = \arg(V_{ij}V_{kl}V_{il}^{*}V_{kj}^{*})$$

Each of the quark indices on the right-hand side of this expression appears twice; once for a matrix element and once for the complex conjugate of a matrix element. Absorbing CKM phases in the quark fields gives opposite phase shifts for the two corresponding elements, which makes the phase of the four-element combination invariant and convention independent.

The three terms in each of the orthogonality relations can be used to construct a triangle in the complex plane. The resulting *unitarity triangles* are schematically shown in Figure 1.2. Figure 1.2a shows the *d–b triangle*, which corresponds to Equation 1.5a. Its sides are defined by the CKM elements for the couplings of the up-type quarks and the down/beauty quarks. The angles α, β and γ are defined by

$$\alpha \equiv \arg\left(-\frac{V_{\text{td}}V_{\text{tb}}^{*}}{V_{\text{ud}}V_{\text{ub}}^{*}}\right) \quad \beta \equiv \arg\left(-\frac{V_{\text{cd}}V_{\text{cb}}^{*}}{V_{\text{td}}V_{\text{tb}}^{*}}\right) \quad \gamma \equiv \arg\left(-\frac{V_{\text{ud}}V_{\text{ub}}^{*}}{V_{\text{cd}}V_{\text{cb}}^{*}}\right)\,. \tag{1.6}$$

The coordinates of the triangle apex are defined as $(\overline{\rho}, \overline{\eta})$.

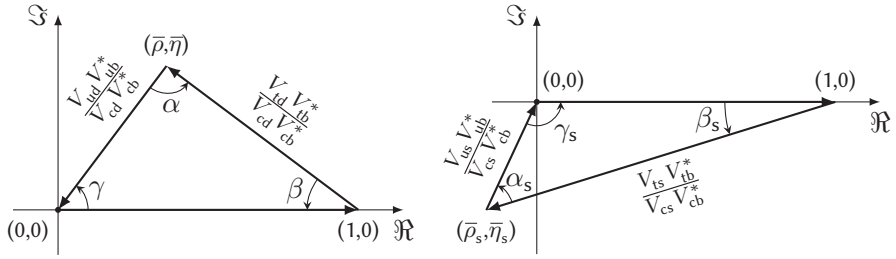


Figure 1.2: *CKM unitarity triangles (figures from [30]). (a) d–b triangle, corresponding to Equation 1.5a. (b) s–b triangle, corresponding to Equation 1.5b.*

Equation 1.5b gives the *s–b triangle* in Figure 1.2b. It is defined in a similar way as the d–b triangle, but its apex has negative real and imaginary coordinates $(\overline{\rho}_s, \overline{\eta}_s)$. The angles α_s, β_s and γ_s are defined by

$$\alpha_s \equiv \arg\left(-\frac{V_{us}V_{ub}^*}{V_{ts}V_{tb}^*}\right) \quad \beta_s \equiv \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \quad \gamma_s \equiv \arg\left(-\frac{V_{cs}V_{cb}^*}{V_{us}V_{ub}^*}\right) . \tag{1.7}$$

In Figure 1.2 the sides of the s–b and d–b triangles were scaled with factors $V_{cd}V_{cb}^*$ and $V_{cs}V_{cb}^*$, respectively, to make the first side lie along the real axis between 0 and 1. Without this scaling, the surface areas of all six possible unitarity triangles are equal and provide a convention-independent measure of the CP violation that is introduced by the CKM matrix. The areas are given by half of the *Jarlskog invariant* (J) [31], which is defined by

$$\Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*) \equiv J \sum_{m,n} \epsilon_{ikm}\,\epsilon_{jln} , \tag{1.8}$$

where ϵ is the Levi-Civita symbol (no summation over the i, j, k, and l indices implied).

Using the Wolfenstein parameterization and neglecting terms of relative order λ^4 and higher (Equation 1.3), the Jarlskog invariant can be expressed as

$$J \approx A^2\,\lambda^6\,(1 - \tfrac{1}{2}\lambda^2)\,\eta . \tag{1.9}$$

Its experimental value is approximately $3 \cdot 10^{-5}$ [28, 29]. This value is four orders of magnitude smaller than the theoretical maximum of $\frac{1}{6\sqrt{3}} \approx 0.1$ given by unitarity.

In the same relative λ^3 approximation, the only CKM-matrix elements with imaginary parts are V_{ub}, V_{td} and V_{ts}. Using the approximations from Equation 1.3 and the angle definitions from Equations 1.6 and 1.7, the angles γ, β, and β_s are approximated by

$$\gamma \approx \pi - \gamma_s \approx \quad \arg(V_{ub}^*) \quad \approx \arctan\left(\frac{\eta}{\rho}\right) \tag{1.10a}$$

$$\beta \approx -\arg(V_{td}) \quad \approx \arctan\left(\frac{(1-\frac{1}{2}\lambda^2)\,\eta}{1-(1-\frac{1}{2}\lambda^2)\,\rho}\right) \tag{1.10b}$$

$$\beta_s \approx \quad \arg(-V_{ts}) \approx \arctan\left(\frac{\lambda^2\,\eta}{1-\frac{1}{2}\lambda^2+\lambda^2\rho}\right)\,. \tag{1.10c}$$

Expanding the expressions for the angles α and α_s in the same way gives the expected relations between the three angles in a triangle:

$$\alpha \approx \pi + \arg(V_{td}) - \arg(V_{ub}^*) \approx \pi - \beta - \gamma \tag{1.11a}$$

$$\alpha_s \approx \arg(V_{ub}^*) - \arg(-V_{ts}) \approx \pi - \beta_s - \gamma_s\,. \tag{1.11b}$$

The experimental values of the CKM angles can be determined by a global fit of the Standard Model to all relevant experimental data. Two of the groups that are doing such a fit are CKMfitter and UTfit, which are using different statistical methods and slightly different sets of input data. With the experimental data that were available in the spring of 2013, the CKMfitter group finds [28]

$$\gamma = (69.7^{+1.3}_{-2.8})^\circ \qquad \beta = (21.8^{+0.8}_{-0.7})^\circ \qquad \beta_s = (1.05 \pm 0.04)^\circ\,. \tag{1.12}$$

Figures 1.3 and 1.4 show the resulting CKMfitter unitarity triangles, together with the constraints on the coordinates of the triangle apices. At lowest order in λ the apices of the d–b and s–b triangles are given by $(\overline{\rho}, \overline{\eta}) \approx (\rho, \eta)$ and $(\overline{\rho}_s, \overline{\eta}_s) \approx (-\lambda^2\,\rho,\ -\lambda^2\,\eta)$, respectively. Constraints from the fit are given by

$$\begin{aligned}(\overline{\rho}, \overline{\eta}) &= (+0.129^{+0.018}_{-0.009},\ +0.348^{+0.012}_{-0.012})\\ (\overline{\rho}_s, \overline{\eta}_s) &= (-0.0068^{+0.0005}_{-0.0010},\ -0.0185^{+0.0006}_{-0.0007})\end{aligned} \tag{1.13}$$

The UTfit collaboration finds slightly different values for the angles with the data available by summer 2013 [29]:

$$\gamma = (70.3 \pm 3.5)^\circ \qquad \beta = (22.0 \pm 0.9)^\circ \qquad \beta_s = (1.07 \pm 0.04)^\circ\,. \tag{1.14}$$

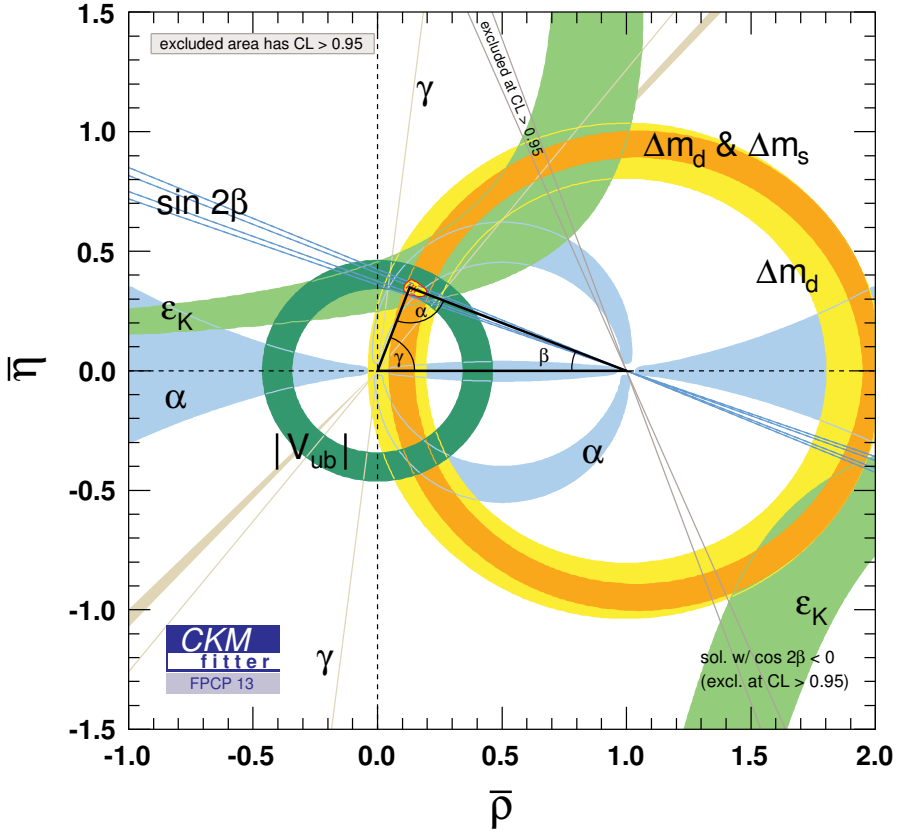


Figure 1.3: *Constraints on the d–b unitarity triangle resulting from the global Standard-Model fit by the CKMfitter group [28]. The fit estimates the coordinates of the triangle apex in the complex plane, which is given by the parameters* $\overline{\rho} \approx \rho$ *and* $\overline{\eta} \approx \eta$*. Constraints on these parameters from the measurements that are input to the fit are shown as the coloured bands.*

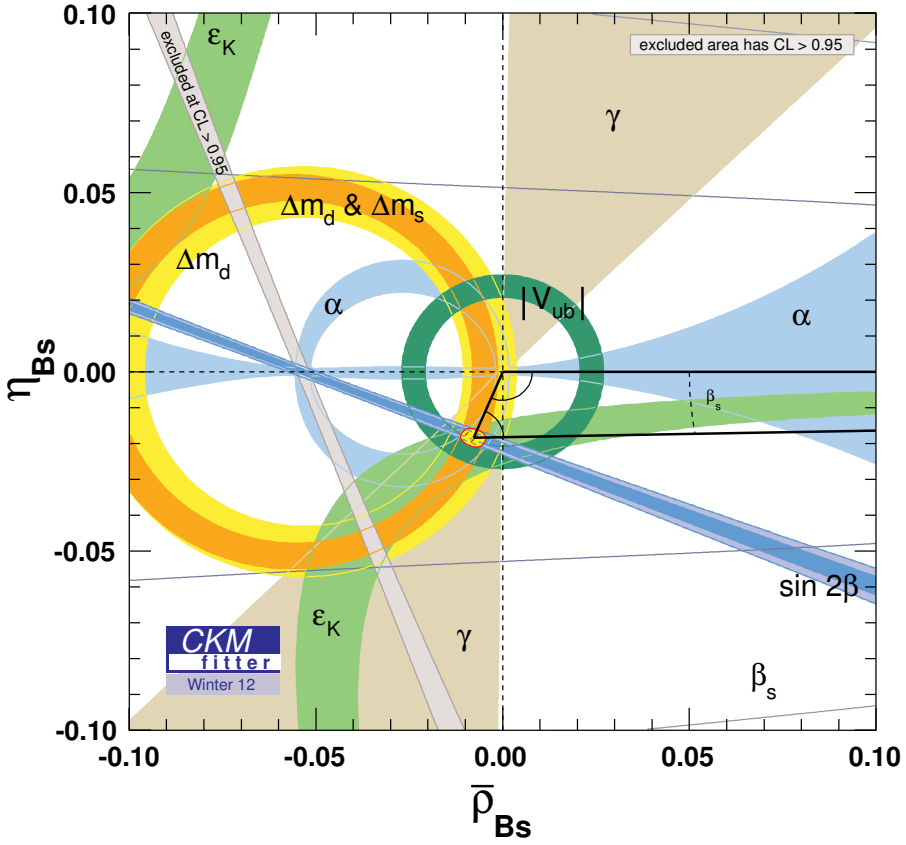


Figure 1.4: *Constraints on the s–b unitarity triangle resulting from the global Standard-Model fit by the CKMfitter group [28]. The fit estimates the coordinates of the triangle apex in the complex plane, which is given by the parameters* $\overline{\rho}_s \approx -\lambda^2\,\rho$ *and* $\overline{\eta}_s \approx -\lambda^2\,\eta$*. Constraints on these parameters from the measurements that are input to the fit are shown as the coloured bands.*

Quark-flavour changing interactions and the CKM picture of CP violation provide excellent probes for testing the Standard Model. They yield many experimental observables, which all depend on only four independent CKM parameters. If the measurement of one of these observables deviates from the value predicted by a global fit or from the value from another measurement, this would be clear evidence for physics beyond the Standard Model.

As will be discussed in the next section, the CP-violation measurement in the $B_s^0 \to J/\psi\,\phi$ decay yields the phase ϕ_s. This phase is approximately equal to $-2\beta_s$ in the Standard Model. The value of the angle β_s is very small, since the imaginary part of V_{ts} is proportional to a factor λ^2 (see Equations 1.3 and 1.10c). Even if contributions of unknown new physics to ϕ_s are small, they could still significantly add to the suppressed Standard Model contribution.

1.3 CP Violation in the $B_s^0 \to J/\psi\,\phi$ Decay

1.3.1 The B_s^0–$\overline{B}_s^0$ System

The B_s^0 meson is a QCD bound state of an anti-b quark and an s quark. Its antiparticle is the $\overline{B}_s^0$ meson (b and anti-s). The B_s^0 and $\overline{B}_s^0$ are charge-neutral particles, which makes it possible to convert one into the other. Figure 1.5 shows the lowest-order diagrams for this transition.

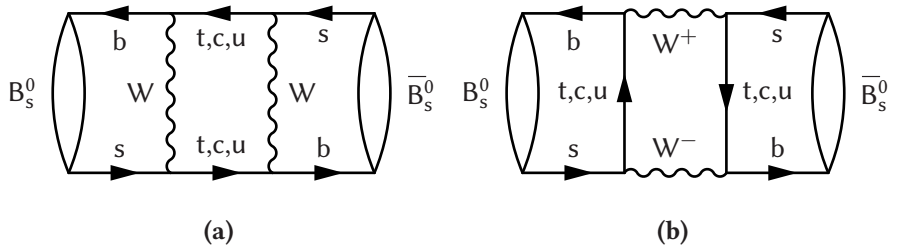


Figure 1.5: *Lowest-order diagrams for B_s^0–$\overline{B}_s^0$ mixing (figures from [3]).*

Their mixing makes the B_s^0 and $\overline{B}_s^0$ a coupled system of particles. The system comprises two eigenstates with different masses and mean lifetimes (see Section 2.1.1). A particle that is created as B_s^0 can at a later point in time be observed as either a B_s^0 or a $\overline{B}_s^0$. As a result, the distribution of the time at which the B_s^0–$\overline{B}_s^0$ system decays is not a simple exponential, as it would be in

the absence of mixing. Because the expression for the probability to observe the system as a B_s^0 or a $\overline{B}_s^0$ as a function of time contains sinusoidal terms, mixing is often referred to as "B_s^0–$\overline{B}_s^0$ oscillations".

The *decay time* is defined as the elapsed time between the production and the decay of the B_s^0–$\overline{B}_s^0$ system in its rest frame. In Section 2.1 the exact shape of the decay-time dependence will be discussed. The applied formalism is common to the D^0, B^0 and B_s^0 mesons and, to some extend, also the K^0 meson, which all mix with their respective antiparticles.

The parameters used to describe the B_s^0–$\overline{B}_s^0$ system are the mean decay width, Γ_s, the difference between the decay widths of the eigenstates, $\Delta\Gamma_s$, and the difference between the masses of the eigenstates, Δm_s. By definition, the mean lifetime of the system is given by $\tau_s \equiv \frac{1}{\Gamma_s}$. The mass difference will turn out to be the frequency of the oscillations of the B_s^0–$\overline{B}_s^0$ probability in time.

In the Standard Model, CP violation enters the mixing process through the CKM-matrix elements at W-boson vertices. The amplitudes in Figure 1.5 depend on the mass of the internal up-type quark, which makes the B_s^0–$\overline{B}_s^0$ mixing process dominated by diagrams with virtual top quarks. As a result, CP violation is small, since it requires multiple contributions with different weak phases. In terms of Equation 1.4, this corresponds to a small value of R, which suppresses the CP asymmetry.

The additional contributions that do give rise to small *CP violation in mixing* are transitions via real intermediate states into which both B_s^0 and $\overline{B}_s^0$ can decay, which will be discussed in more detail in Section 2.1. The asymmetry between the $B_s^0 \to \overline{B}_s^0$ and $\overline{B}_s^0 \to B_s^0$ rates is measured to be approximately one per cent [32]. Given the current experimental uncertainties, this is compatible with no CP violation.

Depending on the final state, there may also be different amplitudes contributing to the decay of the B_s^0 meson. Also then interference can lead to different decay rates, in this case for the $B_s^0 \to f$ and $\overline{B}_s^0 \to \bar{f}$ processes. This form is termed *CP violation in decay*.

The first significant observation of CP violation in the B_s^0 system was recently obtained by LHCb [33]. This was a measurement of CP violation in decay for $\overline{B}_s^0 \to K^+\pi^-$ and $B_s^0 \to K^-\pi^+$, where an asymmetry in the decay rates of about 30% was found.

An interesting situation occurs if both B_s^0 and $\overline{B}_s^0$ can decay into the same final state f'. In that case the processes $B_s^0 \to f'$ and $B_s^0 \to \overline{B}_s^0 \to f'$ interfere,

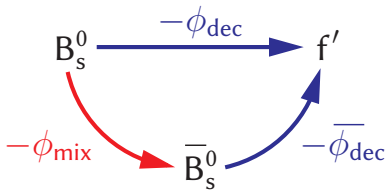


Figure 1.6: *Interference between the $B_s^0 \to f'$ and $B_s^0 \to \overline{B}_s^0 \to f'$ processes. The phases ϕ_{mix}, ϕ_{dec}, and $\overline{\phi}_{dec}$ are the relevant weak phases contributing to the two decay paths from the mixing, B_s^0 decay and $\overline{B}_s^0$ decay, respectively. This is assuming no CP violation in mixing or CP violation in decay.*

which is depicted schematically in Figure 1.6. Even without CP violation in mixing or CP violation in decay, the interference of these two decay paths may cause a difference between the rates of $B_s^0(\to \overline{B}_s^0) \to f'$ and $\overline{B}_s^0(\to B_s^0) \to \overline{f'}$. This difference is called *CP violation in the interference of decays with and decays without mixing*.

Interference between the decays with and without mixing leads to a CP asymmetry that depends on the decay time. A sinusoidal term is introduced in the differential decay rate as a function of decay time with an amplitude that depends on the amount of CP violation. This oscillation originates from the time dependence of the transition amplitudes in B_s^0–$\overline{B}_s^0$ mixing and consequently has a frequency equal to the mass difference Δm_s. In an experiment it is important to resolve the resulting oscillation in the decay-time distribution, which enables the measurement of its amplitude and hence the amount of CP violation.

A special case of this form of CP violation is the one where f' is a CP eigenstate. In this case f' and $\overline{f'}$ are identical and CP violation results in a difference between the decay rates of $B_s^0(\to \overline{B}_s^0) \to f'$ and $\overline{B}_s^0(\to B_s^0) \to f'$.

An example of such a decay is the so-called "golden mode" $B^0 \to J/\psi\, K_S^0$. The combination of CKM-matrix elements from B^0–$\overline{B}^0$ mixing and the $B^0 \to J/\psi\, K_S^0$ decay makes the amplitude of the oscillations in decay time depend on the CKM angle β. The measurement of β with this decay mode by the BaBar and Belle experiments [34] was the first observation of CP violation in B^0 decays. The result of combining all currently available measurements of the angle β [32] is consistent with the value obtained from the global Standard Model fits [28, 29].

1.3.2 $B_s^0 \to J/\psi\,\phi$ Decay

The $B_s^0 \to J/\psi\,\phi$ decay[2] of the B_s^0 system is equivalent to the $B^0 \to J/\psi\,K_S^0$ decay of the B^0 system. Its decay-time distribution depends on the angle β_s instead of β, as will be shown below.

However, in comparison to $B^0 \to J/\psi\,K_S^0$, a complication arises from the fact that both the J/ψ and the ϕ are spin-one mesons, whereas the K_S^0 is a spin-zero meson. This leads to three possible orbital angular momentum configurations of the $J/\psi\,\phi$ system, compared to one configuration for the $J/\psi\,K_S^0$ system. As a result, the $B_s^0 \to J/\psi\,\phi$ decay comprises three different CP eigenstates instead of one. The contributions of these states must be statistically separated by an analysis of the J/ψ and ϕ spin polarizations for an optimal measurement [35, 36], as will be explained in Section 2.2.

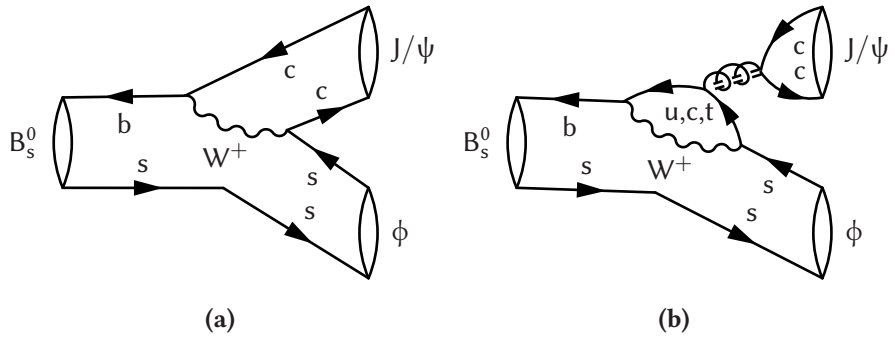


Figure 1.7: *$B_s^0 \to J/\psi\,\phi$ decay (figures from [3]): (a) tree-level diagram; (b) penguin diagram. The curled/dashed line in (b) represents a colour-neutral state, which can be a Z^0, a photon, or colour-singlet gluons.*

At quark level, the $B_s^0 \to J/\psi\,\phi$ decay is a $b \to c\bar{c}s$ *transition.* The b quark decays into an s quark and a $c\bar{c}$ pair, where the s quark forms a ϕ meson with the spectator $\bar{s}$ quark and the $c\bar{c}$ pair forms a J/ψ meson. Figure 1.7a shows the dominant Standard Model contribution. This is a tree-level diagram, which is proportional to the CKM-matrix elements V_{cs} and V_{cb}^*.

Since the CP-violation phenomenology is governed by B_s^0–$\overline{B}_s^0$ mixing and the $b \to c\bar{c}s$ transition, decays in which the $c\bar{c}$ and $s\bar{s}$ pairs form different

[2] Charge-conjugate particles, CP-conjugate decays and neutral-meson mixing are implied, unless stated otherwise.

mesons may be used in addition to $B_s^0 \to J/\psi\,\phi$. An example is the $B_s^0 \to J/\psi\,\pi^+\pi^-$ decay [37], which is also used by LHCb to measure ϕ_s.

The dominant contributions to the B_s^0–$\overline{B}_s^0$ mixing with internal top quarks (Figure 1.5) are proportional to $(V_{ts}V_{tb}^*)^2$. In combination with the tree-level decay, these give a weak phase difference between decays with and decays without mixing of

$$\begin{aligned} &\arg\left[(V_{ts}V_{tb}^*)^2\right] + \arg\left(V_{cs}^*V_{cb}\right) - \arg\left(V_{cs}V_{cb}^*\right) \\ &\qquad = 2\,\arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) = 2\beta_s\,, \end{aligned} \tag{1.15}$$

where the first contribution comes from the $B_s^0 \to \overline{B}_s^0$ transition, the second contribution from the subsequent $\overline{B}_s^0$ decay, and the third contribution from the B_s^0 decay path without mixing.

Additional contributions to the mixing and decay processes can affect this prediction of the weak phase. The quantity that is observed in this measurement is the phase ϕ_s, which is equal to $-2\beta_s$ with the above assumptions.[3]

In the Standard Model, corrections come from the mixing diagrams with real states and from higher order contributions to the decay. While the former lead to small Standard Model CP violation in mixing, these amplitudes are not expected to contribute significantly to the value of ϕ_s (see Section 2.1.3). An example of an additional contribution to the decay is the *penguin diagram* in Figure 1.7b. Although hard to estimate, small contributions from penguin diagrams to the decay are expected [38].

Apart from CP violation in the interference between decays with and decays without mixing, the $B_s^0 \to J/\psi\,\phi$ process is affected by CP violation in mixing and by CP violation in decay. As discussed above, CP violation in mixing is small and its effects are not expected to be measurable in the $B_s^0 \to J/\psi\,\phi$ measurement. CP violation in decay arises from interference between different decay amplitudes. Since the decay is dominated by the tree-level amplitude, also CP violation in decay is expected to be small in the $B_s^0 \to J/\psi\,\phi$ process.

Beyond the Standard Model, CP violation can arise from new contributions to the B_s^0–$\overline{B}_s^0$ mixing process [23,24] as well as the B_s^0 decay [25]. In gen-

[3] There are various, often contradicting notations of the phase difference in use. In this work β_s is the CKM-triangle angle (Equation 1.7) and ϕ_s the model-independent observable (see Sections 2.2 and 2.3.1). If only dominant Standard Model contributions to both mixing and decay are considered, ϕ_s is equal to $-2\beta_s$.

eral it is expected that potential corrections to the Standard Model are largest in processes with virtual particles in loops at lowest order. Loop processes are smaller than tree-level processes, which makes Standard Model contributions compete with potential new contributions at the same level. New particles must be heavy and cannot be created in decays of Standard Model particles. In loop processes, however, they can enter as virtual particles and replace, for example, the internal top quarks in the B_s^0–$\overline{B}_s^0$ mixing diagrams. For these reasons, the focus of the measurement will be on mixing-induced CP violation rather than CP violation that arises from the $B_s^0 \to J/\psi\,\phi$ decay, which is dominated by the tree-level amplitude in Figure 1.7a.

If the effect of new contributions on the value of ϕ_s is large compared to the Standard Model prediction, the corresponding deviation can be revealed by the measurement of a large non-zero value of ϕ_s. In this case an exact estimate of the Standard Model value is not required and subleading penguin contributions can be neglected. Such a large deviation could be detected by a low-statistics measurement, which would not have the precision to distinguish between the cases of no CP violation, the Standard Model, and a small deviation from the Standard Model.

Measurements of ϕ_s with the $B_s^0 \to J/\psi\,\phi$ decay prior to this thesis have been performed by the D0 [39], CDF [40], ATLAS [41], and LHCb [3] experiments. LHCb also measured ϕ_s in the $B_s^0 \to J/\psi\,\pi^+\pi^-$ decay channel [42]. All measurements and their combination [32] are compatible with no CP violation and also with the value of β_s from the Standard Model fit. The LHCb measurement in the $B_s^0 \to J/\psi\,\phi$ channel yielded

$$\phi_s = +0.07 \pm 0.09\,(\text{stat.}) \pm 0.01\,(\text{syst.})\ \text{rad}\,, \tag{1.16}$$

where the first uncertainty is statistical and the second systematic. This value is to be compared with the Standard Model estimate $-2\beta_s = -0.0368^{+0.0013}_{-0.0014}$ rad (see Equation 1.12 and reference [28]).

A measurement of the decay-time distribution in $B_s^0 \to J/\psi\,\phi$ is not only sensitive to the phase ϕ_s, but also to the lifetime parameters of the B_s^0–$\overline{B}_s^0$ system. The previously mentioned CP violation measurements yielded estimates of Γ_s and $\Delta\Gamma_s$ as well. These parameters were also estimated by CMS in a measurement that assumed no CP violation [43].

In the Standard Model the parameter Γ_s is expected to be equal to the B^0 decay width, Γ_d, up to relative corrections of the order of 10^{-3} [44]. Γ_d is measured to be $0.6583 \pm 0.0030\,\text{ps}^{-1}$ [32]. A prediction of the decay-width

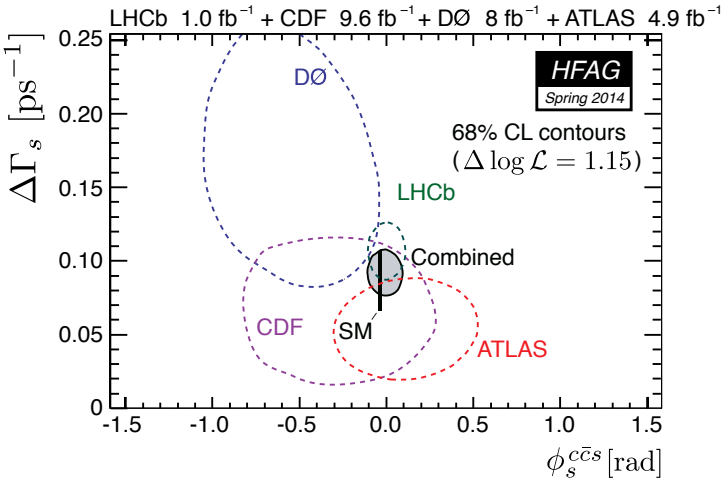


Figure 1.8: *Combination of ϕ_s (here represented as $\phi_s^{c\bar{c}s}$) and $\Delta\Gamma_s$ measurements by HFAG [32]. The estimates at 68% confidence level (CL) by the different experiments are shown by the dashed contours. Note that the LHCb contour is a combination of measurements in the $B_s^0 \to J/\psi\phi$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays. The combined 68% confidence region is shown by the shaded area and the Standard Model prediction by the vertical bar.*

difference $\Delta\Gamma_s$ yields $0.087 \pm 0.021\,\text{ps}^{-1}$ [44]. The measurements of Γ_s and $\Delta\Gamma_s$ and their combination [32] are compatible with these predictions. The estimates from the LHCb measurement are given by

$$\Gamma_s = 0.663 \pm 0.005\,(\text{stat.}) \pm 0.006\,(\text{syst.})\ \text{ps}^{-1} \tag{1.17a}$$

$$\Delta\Gamma_s = 0.100 \pm 0.016\,(\text{stat.}) \pm 0.003\,(\text{syst.})\ \text{ps}^{-1}\,. \tag{1.17b}$$

A graphical representation of the status of ϕ_s and $\Delta\Gamma_s$ measurements in the spring of 2014 by the Heavy Flavour Averaging Group (HFAG) is shown in Figure 1.8. The LHCb contribution is a combination of the $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ results with data from 2011, which form roughly one third of the currently available dataset. The 68% confidence-level (CL) contour of the combined result is consistent with the region that represents the Standard Model prediction.

As it is clear from Figure 1.8 that possible effects from non-Standard Model physics on ϕ_s and $\Delta\Gamma_s$ cannot be large, more precise measurements are required to probe them. With the currently available LHCb data, combination

of the ϕ_s estimates from the two decay channels is expected to reach a precision of approximately 0.04 rad, which is roughly equal to the value of the Standard Model prediction. An improvement of this precision by an order of magnitude is expected with future LHCb data.

For a measurement with this precision, effects have to be taken into account that have not been considered for previous measurements. A precise estimate of Standard Model CP violation is required, which also includes higher order penguin contributions to the $b \to c\bar{c}s$ transition. A complication arises from the fact that these additional contributions affect the value of ϕ_s differently for the $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ decays and for the different intermediate states that contribute to the decays [38]. Moreover, physics beyond the Standard Model may also affect all intermediate states differently [25]. These effects need to be considered not only in theoretical predictions, but also in measurements.

In the $B_s^0 \to J/\psi\,\phi$ measurement that is presented in this thesis the dependence of CP violation on the intermediate state is taken into account for the first time. The value of ϕ_s is estimated separately for each of the CP eigenstates in the decay, which is possible if the contributions of the states are statistically separated. The decay model for this measurement is presented in Chapter 2.

1.3.3 The $\mu^+\mu^-\,K^+K^-$ Final State

The $J/\psi\,\phi$ system is unstable and only its decay products can be detected. There are many possibilities for both the J/ψ and ϕ mesons to decay [45]. The only final state that will be considered here is $\mu^+\mu^-\,K^+K^-$, where the muon pair originates from the J/ψ decay and the kaon pair from the ϕ decay. All four particles can be detected efficiently and with low background by the LHCb detector, as will be described in Section 1.4. This makes $B_s^0 \to J/\psi(\to \mu^+\mu^-)\,\phi(\to K^+K^-)$ the optimal channel to measure CP violation in B_s^0 decays with a $b \to c\bar{c}s$ transition.

Since the $\mu^+\mu^-\,K^+K^-$ final state can also be reached through resonances other than the J/ψ and the ϕ, interference with other processes needs to be considered. To optimize for the detection of the J/ψ and ϕ resonances, a region in kinematic phase space is selected where this intermediate state dominates.

The measured invariant mass of the muon pair is required to be between 3.03 and 3.15 GeV/c^2 (see also Section 3.2). This restriction is assumed to select

only those $B_s^0 \to \mu^+\mu^- K^+K^-$ decays with muons coming from a J/ψ, which has a mass of approximately 3.10 GeV/c^2 [45].

Also the invariant-mass range of the K^+K^- pair is restricted, but there the situation is more complicated. An analysis of the resonant components in the K^+K^- system [46] has shown that there is a mass region where the $\phi(1020)$ dominates, but other contributions cannot be neglected.

The K^+K^--mass spectrum in $B_s^0 \to J/\psi K^+K^-$ decays is shown in Figure 1.9. The contribution of the $\phi(1020)$ is represented by the dashed, magenta curve. Notice that part this peak is not visible, because of the truncated vertical scale.

For the $B_s^0 \to J/\psi\,\phi$ CP-violation measurement, only K^+K^- pairs with a mass between 0.99 and 1.05 GeV/c^2 are selected. In this region the $\phi(1020)$ clearly dominates, but there are also contributions from the $f_0(980)$ and from K^+K^- pairs that do not originate from the decay of a resonance.

For both the $f_0(980)$ and non-resonant contributions the K^+K^- system is in a state of zero orbital angular momentum and hence these contributions are commonly referred to as the *K^+K^- S-wave*. Because the $\phi(1020)$ is a spin-one particle, the orbital angular momentum of the K^+K^- system has an orbital angular momentum equal to one for the $J/\psi\,\phi$ intermediate state. Therefore, the $B_s^0 \to J/\psi\,\phi$ contribution is sometimes termed *K^+K^- P-wave*. Since the $J/\psi\,\phi$ and K^+K^- S-wave processes are observed simultaneously in the measurement, the analysed decay is referred to as $B_s^0 \to J/\psi K^+K^-$.

In addition to separating the three components of the $B_s^0 \to J/\psi\,\phi$ process, these components are also separated from the K^+K^- S-wave contribution. The latter could be accomplished on a statistical basis with an analysis of the K^+K^--mass distribution. However, this distribution does not discriminate between the CP eigenstates of the $B_s^0 \to J/\psi\,\phi$ decay, so a different approach is chosen.

The spatial distributions of the final-state particles are different for the $B_s^0 \to J/\psi\,\phi$ and K^+K^- S-wave contributions, which enables a statistical separation of the two components. An analysis of the spatial distributions also separates $J/\psi\,\phi$ spin-polarization states and hence the three CP eigenstates of this system.

The directions of the final-state particles are specified with respect to the momentum directions of the $\mu^+\mu^-$ and K^+K^- systems in the centre-of-mass system of the B_s^0 meson. This is done with three *decay angles*, which can be computed given the four-momenta of the final-state particles. These angles are included in the model of the decay, in addition to the decay time. The

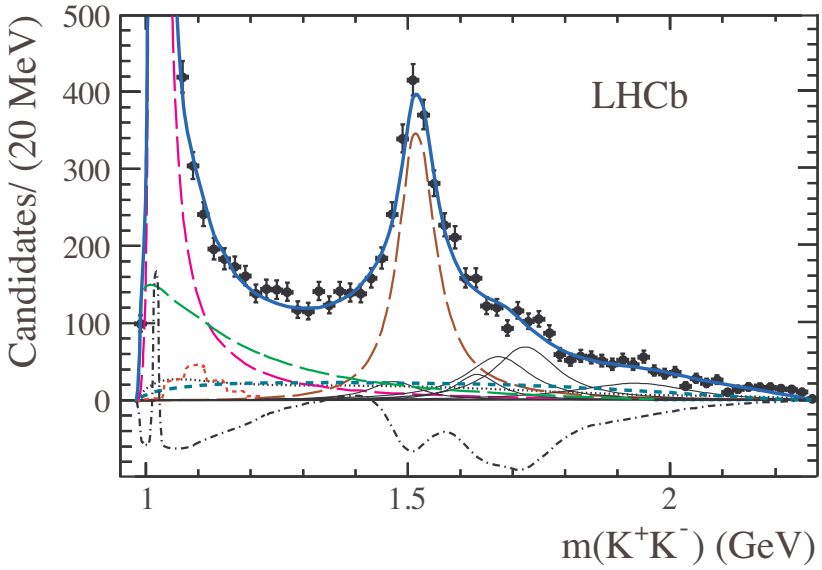


Figure 1.9: *K^+K^--mass spectrum in $B_s^0 \to J/\psi K^+K^-$ decays [46]. The black points represent a histogram of decay candidates. A model of the mass distribution is shown as the blue curve. The largest contributions to the distribution come from the $\phi(1020)$, $f_2'(1525)$, and $f_0(980)$ resonances, which are shown as the magenta, brown, and green, long-dashed lines, respectively. Notice that a large part of the $\phi(1020)$ peak is not visible, because of the truncated vertical scale. Other resonances (heavier ϕ, f_0, and f_2 states) are shown as the thin, black curves and a non-resonant contribution as the dashed, cyan curve. The contribution from interferences between the resonances is represented by the dotted-dashed, black line. The small dotted, black and dashed, red contributions are backgrounds of four particles that do not originate from a $B_s^0 \to J/\psi K^+K^-$ decay.*

formalism for the angular dependence of the decay is described in detail in Section 2.4 and Appendix A.

1.4 The $B_s^0 \to J/\psi K^+K^-$ Decay in LHCb

1.4.1 B_s^0-Meson Production at the Large Hadron Collider

The B_s^0 mesons that are used for the $B_s^0 \to J/\psi K^+K^-$ measurement are produced in the proton–proton collisions of the Large Hadron Collider (LHC) [47]. The LHC accelerates protons to high energies and makes them collide head-on at designated points. For the measurement presented here, data from the LHC data-taking runs in 2011 and 2012 were used. In 2011 colliding protons had an energy of 3.5 TeV, while their energy was increased to 4 TeV for 2012 [48]. This gave a centre-of-mass energy in proton–proton collisions of 7 and 8 TeV, respectively.

Protons are brought to an energy of 0.45 TeV by a series of pre-accelerators before they are injected into the LHC. There they are stored in two opposite, circular beams and accelerated to the collision energy. Each beam contained 1380 bunches of 10^{11} protons for most of the 2011 and 2012 runs [48].

When the collision energy is reached, the beams are tuned to collide bunches at four distinct *interaction points*, where the detectors of the LHC experiments are located. The moment at which two bunches meet is called a *bunch crossing*. If proton–proton collisions in a bunch crossing produce particles that are of interest for an experiment, this is called an *event*. After typically ten hours of collisions the number of remaining protons in the beams becomes too small to maintain a sufficiently high collision rate and the beams are dumped to restart the cycle of acceleration and collisions.

The mean number of collisions per bunch crossing that produce detectable particles at the LHCb interaction point varied between one and two in 2011 and 2012. This number is determined by the probability that two protons interact when they pass each other and by the probability that the protons get close enough to enable such an interaction. These probabilities are represented by two effective quantities: the proton–proton *cross section* and the beam *luminosity*.

The luminosity is an effective density of protons that meet each other per unit of time across a surface perpendicular to the beams at the interaction point. This quantity is fully determined by the beam parameters, in particular by the number of protons per bunch and the number of bunch crossings per

unit time. To get a quantity that represents the total number of protons that passed each other per unit surface, the luminosity is integrated over time. The integrated luminosity is $1\,\text{fb}^{-1}$ for 2011 and $2\,\text{fb}^{-1}$ for 2012.

To obtain the number of interactions per unit time, the luminosity is multiplied by the surface area of the effective proton cross section. This surface area represents the strength of the interaction between the colliding particles, which typically depends on the centre-of-mass energy of the collision.

The total proton–proton cross section is a sum of the cross sections for all possible interactions, which can be either elastic or inelastic. While elastic interactions do not break up the colliding protons, inelastic interactions do and create new particles. The latter type is relevant for the production of B_s^0 and $\overline{B}_s^0$ mesons.

The sum of the cross sections for B_s^0 and $\overline{B}_s^0$ production in LHCb is measured to be approximately $1 \cdot 10^{10}$ fb at an energy of 7 TeV [49]. Assuming the cross section at 8 TeV is approximately the same, this gives a total of $3 \cdot 10^{10}$ produced B_s^0 and $\overline{B}_s^0$ mesons in 2011 and 2012. A fraction of $3 \cdot 10^{-5}$ decays into $J/\psi(\to \mu^+\mu^-)\,\phi(\to K^+K^-)$ [45], which yields an expected number of $9 \cdot 10^5$ decays. Including particle-detection inefficiencies and selection of usable decays, about $9 \cdot 10^4$ decays are left for analysis (see Section 3.2).

To produce B_s^0 and $\overline{B}_s^0$ mesons, (anti)beauty and (anti)strange quarks must be created. Beauty quarks are predominantly produced as $b\bar{b}$ pairs. Beauty hadrons are formed by combining the b and $\bar{b}$ quarks with lighter quarks that are produced at a later stage. If this lighter quark is an s ($\bar{s}$) quark, the result is a B_s^0 ($\overline{B}_s^0$) meson.

In the inelastic collisions that are relevant for $b\bar{b}$ production, the constituents of the colliding protons interact. These constituents can be gluons or (anti)quarks, which are collectively called *partons*. A proton consists of three valence quarks: two up quarks and one down quark. In addition it contains virtual partons, created by low-energy, non-perturbative QCD interactions. The centre-of-mass energy of the collision is large enough to make the lifetime of these virtual states larger than the time needed to also interact with partons from the other proton.

The colliding protons break up in the parton interaction, leaving coloured remnants. Quarks and gluons created in the interaction and these proton remnants recombine into colourless hadrons. This process is called *hadronization*. Tens of hadrons can be created in a single proton–proton collision at the LHC.

The energy carried by a parton is a fraction of the proton energy. For

$b\bar{b}$ production the energy fractions of the interacting partons must be minimally the ratio of the energy corresponding to beauty-hadron masses and the collision energy, which is of the order of 10^{-3}. However, in most collisions the fractions are different for the two interacting partons. This gives the produced particles additional energy and leads to large boosts along the beam direction.

Conceptually the parton interaction can be separated into two parts. First there are high-energy interactions, which are described by perturbative QCD. The proton energy is converted into particle masses, which may be much larger than the proton mass. Beauty quarks are created at this stage.

Interactions that follow are less energetic and create only lighter quarks in quark–antiquark pairs. This non-perturbative process is called *fragmentation*. A B_s^0 meson is formed with the strange quark from a created $s\bar{s}$ pair. The remaining antistrange quark forms a hadron with other quarks in the fragmentation process.

The B_s^0 production mechanism can be exploited to determine whether the produced meson was a B_s^0 or a $\overline{B}{}_s^0$. This procedure is called *flavour tagging*. The flavour information cannot be inferred directly from the decay into $\mu^+\mu^- K^+K^-$, since both B_s^0 and $\overline{B}{}_s^0$ decay into this final state. For the CP-violation measurement in $B_s^0 \to J/\psi\, K^+K^-$, however, it is crucial to disentangle the two contributions, because the important features in the decay-time distribution cancel in their sum (see Sections 2.1.2 and 2.3).

Flavour tagging relies on the fact that both the beauty and strange quarks that form an (anti-)B_s^0 meson are predominantly produced in quark–antiquark pairs. Even if the flavour of the produced meson cannot be determined from its decay products, it can still be inferred from the charges of the (anti)b and (anti)s quarks that remain of the required $b\bar{b}$ and $s\bar{s}$ pairs. This is schematically shown in Figure 1.10.

The charge of the beauty quark on the "opposite side of the event" can be determined from the charges of the electron, muon, or kaon in the decay of the hadron it formed. This procedure is termed *opposite-side tagging* and is described in reference [50].

The efficiency of opposite-side tagging is limited. The required charged particles are produced in many, but not all beauty decays. In addition, the decay products may not be detectable by the LHCb detector. There is also a probability that the determined flavour is wrong. A neutral beauty meson may convert into its antiparticle before it decays, which gives the opposite

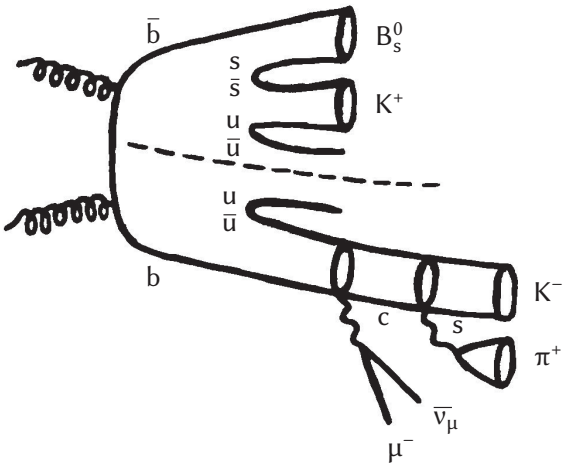


Figure 1.10: *Production and decay mechanisms exploited by flavour-tagging algorithms. Same-side tagging determines the flavour of the B_s^0 or $\bar{B}_s^0$ meson from the charge of the kaon that is produced close to it (top half of the figure). Opposite-side tagging determines the flavour from the charges of the decay products of the "opposite" b or $\bar{b}$ quark in the event (bottom half of the figure).*

tag. A background of charged particles from elsewhere in the event gives additional wrong tags. The resulting effective fraction of B_s^0 mesons with a correct opposite-side tag in the $B_s^0 \to J/\psi K^+K^-$ measurement is 2.6% (see Section 3.6).

When combined with an (anti)up quark, the remaining quark from the $s\bar{s}$ pair forms a charged kaon. In *same-side kaon tagging* the flavour of the B_s^0 meson is estimated by determination of the charges of kaons that are close to it in momentum space [51]. This procedure is limited by other possibilities for fragmentation and hadronization of the strange quark and by the efficiency of detecting the correct kaon. The effective fraction of B_s^0 mesons with a correct same-side tag is 1.3% in the $B_s^0 \to J/\psi K^+K^-$ measurement.

1.4.2 $B_s^0 \to J/\psi K^+K^-$ Decays in LHCb

LHCb is one of the four large experiments that analyse the particles produced by LHC collisions. The LHCb detector measures particle trajectories in the "forward" direction, that is, along the direction of one of the proton beams.

Particles produced with momenta at angles between typically 1 and 15 degrees from the beam direction can be detected. Although this *acceptance region* covers only a small part of possible particle directions, it contains roughly a quarter of all produced beauty quarks. This makes the LHCb design efficient for studies of hadrons containing beauty quarks.

The particles that are produced in an LHC collision get a momentum from the boost introduced by asymmetric parton momenta, but also from the centre-of-mass energy that is available in the parton interaction. B_s^0 mesons that are selected for the analysis of $B_s^0 \to J/\psi K^+K^-$ decays have an average momentum of the order of a hundred GeV/c (see Figure 1.11a). With the 5.4 GeV/c^2 B_s^0-meson mass [45], this gives an average Lorentz factor of about 20.

The B_s^0 meson has a relatively large mean lifetime of about 1.5 ps [32]. As a result of the significant lifetime and boost, B_s^0 mesons cover a typical distance of several millimetres before they decay. The measurement of this *flight distance* is used to infer the decay time. It is determined by reconstructing the positions of the points of production and decay. The measured distribution for B_s^0 mesons that are used in the $B_s^0 \to J/\psi K^+K^-$ measurement is shown in Figure 1.11b.

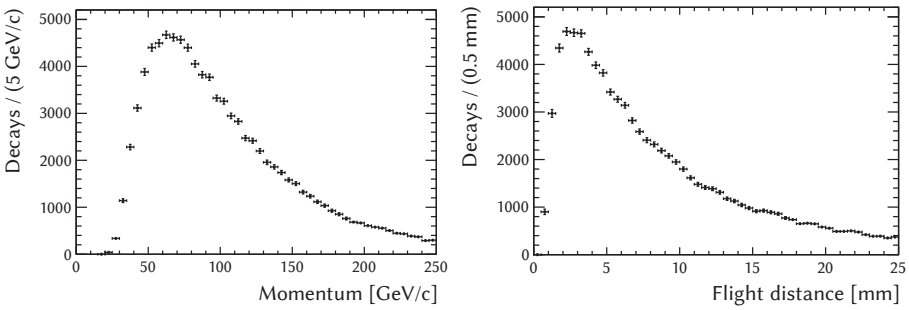


Figure 1.11: *Histograms of (a) the reconstructed momenta and (b) the flight distances of B_s^0 mesons in the $B_s^0 \to J/\psi K^+K^-$ measurement.*

Since LHC bunches have a finite size and the positions of the colliding protons in their respective bunches are not known a priori, the interaction point has to be measured for each collision. This is done by measuring the trajectories of produced particles and extrapolating them to the point of their

common origin. This point is termed the *primary vertex*. The *secondary vertex* is the point where the B_s^0 decayed, which is measured with only the trajectories of the four decay B_s^0 products.

Figure 1.12 schematically shows the vertices in the $B_s^0 \to J/\psi K^+K^-$ decay. The vertices are depicted by ellipsoids with a size that represents the uncertainty in the vertex position. The resolution on the vertex position is better for the primary vertex than for the secondary vertex, because a larger number of particles is available for the reconstruction of the former.

The secondary vertex is constructed from the $\mu^+\mu^-$ vertex and the K^+K^- vertex, where the former dominates. This is caused by the fact that only K^+K^- invariant masses in a range between 0.99 and 1.05 GeV/c^2 are selected, which is just above the threshold of twice the kaon mass. As a result, the opening angle between the K^+ and K^- trajectories is small, which gives the measurement of the point of common origin a large uncertainty.

The decay time can be calculated from the flight distance, provided that also the B_s^0 momentum is known:

$$t = \frac{1}{\gamma(v)}\frac{d}{v} = \frac{md}{\gamma(v)mv} = \frac{md}{|\boldsymbol{p}|} \ , \tag{1.18}$$

where t is the decay time in the rest system of the B_s^0, d the flight distance, v the B_s^0 velocity, γ the corresponding Lorentz factor, $\boldsymbol{p}$ the momentum, and m the B_s^0 mass. The momentum is reconstructed as a vector sum of the decay-product momenta, which are also required to calculate the decay angles. Momenta of charged particles are inferred from the curvature of their tracks in the magnetic field of the detector.

The B_s^0 momentum and flight distance are estimated by combining information from measured particle momenta and extrapolated positions [53]. Uncertainties in the measurements are propagated and ultimately lead to an uncertainty in the estimated decay time. This decay-time resolution is estimated for each B_s^0 decay and taken into account in the analysis (see Section 3.3.1). The effective resolution is about 0.05 ps. This suffices to resolve the oscillations in the $B_s^0 \to J/\psi K^+K^-$ decay-time distribution, which have a period of $2\pi/\Delta m_s \approx 0.35$ ps.

Before the vertex positions and momenta can be determined, the four particles from a $B_s^0 \to J/\psi K^+K^-$ decay have to be selected from all particles that are produced in an LHC event. Since the particle multiplicity in inelastic proton–proton collisions is high, there is a significant probability to select four particles that do not originate from a B_s^0 decay.

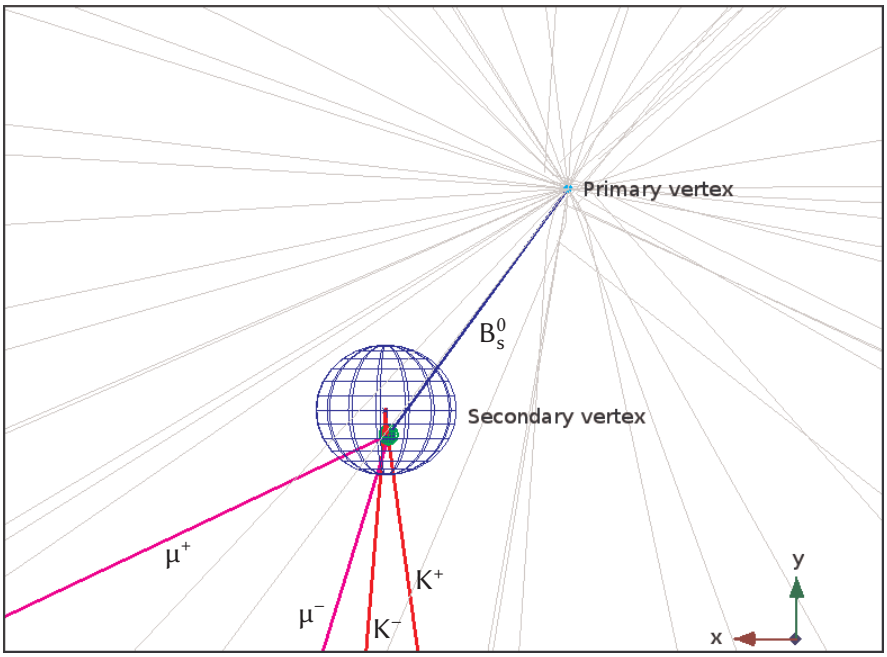


Figure 1.12: *Vertices in a $B_s^0 \to J/\psi K^+K^-$ decay (figure from [52]). The vertices are indicated by ellipsoids with a size that represents the uncertainty in the vertex position. The primary vertex is indicated by the cyan ellipsoid, the $\mu^+\mu^-$ vertex with the small green ellipsoid and the K^+K^- vertex with the large blue ellipsoid. The secondary vertex is the combination of the $\mu^+\mu^-$ and K^+K^- vertices.*

Combinations of four particles that are considered for analysis are called *decay candidates*. Candidates can either be real $B_s^0 \to J/\psi K^+K^-$ decays (*signal*) or combinations that fake the decay signature (*background*). Distributions of decay time and decay angles for background events are subtracted from the measured distributions to obtain the net signal contribution (see Section 3.2).

There are several categories of background decay candidates. Most background is *combinatorial*, which means that candidates are formed from random combinations of four particles. Often all four particles originate directly from a primary vertex. This combinatorial background is called *prompt*. For non-prompt candidates some of the particles originate from a secondary vertex in the decay of a long-lived particle, for example a beauty hadron. Since

the $\mu^+\mu^-$ vertex has a relatively small uncertainty in comparison with the K^+K^- vertex, this often involves a $J/\psi \to \mu^+\mu^-$ decay combined with kaons from the primary vertex.

To reduce the amount of prompt background that is considered for further analysis, candidates for which the secondary vertex is too close to the primary vertex are rejected. This leads to a detection inefficiency for candidates with a small flight distance and, consequently, small decay time. This effect is visible in Figure 1.11b, where the number of decays goes to zero at small decay time.

Another background category is *misidentified background*, which comprises decays where one or more particles have been incorrectly identified. The two misidentified backgrounds that are taken into account in the $B_s^0 \to J/\psi K^+K^-$ measurement are $B^0 \to J/\psi K^{*0}(\to K^+\pi^-)$ and $\Lambda_b^0 \to J/\psi p K^-$. In the B^0 decay the pion is misidentified as a kaon and in the Λ_b^0 decay the proton as a kaon.

There are different stages of event and decay-candidate selection. Each stage is optimized to find signal decays with high efficiency and at the same time keep the background at a manageable level. The first three selection stages are *triggers* [54], which are designed to select events of interest *online*, before detector data are stored for further analysis. Two *offline* stages select decay candidates in the stored events.

The first trigger is called *Level 0* (L0). It is implemented in the detector hardware and designed to bring the rate of incoming events down to less than 1 MHz, which is the maximum frequency at which the detector can be read out. In proton collisions that create beauty hadrons, particles are created in all directions in the centre-of-mass frame of the parton interaction. As a result, particles in these collisions typically have significant momentum components in the direction transverse to the proton beams. This is used by the L0 trigger, which selects only events that contain particles above a transverse-momentum threshold.

Events that are selected by L0 are processed by the *High Level Trigger* (HLT). This trigger is implemented in software and consists of two stages: HLT1 and HLT2. In the first stage a minimal reconstruction of particle trajectories and vertices is performed. For the $B_s^0 \to J/\psi(\to \mu^+\mu^-) K^+K^-$ measurement, events are selected if they contain the signatures of the muons in this decay.

HLT1 selects roughly 40 kHz of events, which are subsequently processed by HLT2. At this stage a more complete particle reconstruction is performed.

Events are selected for the $B_s^0 \to J/\psi K^+K^-$ measurement if they contain $\mu^+\mu^-$ pairs that are compatible with a J/ψ decay at a significant distance from a primary vertex. Events for all LHCb measurements are stored at a rate of 3–5 kHz.

The first stage of the offline selection fully reconstructs the particle trajectories and vertices in the remaining events and selects combinations of muon and kaon pairs that form $B_s^0 \to J/\psi K^+K^-$ candidates. A set of loose selection criteria is applied to filter out part of the background, which reduces the number of remaining candidates to a manageable level. This procedure is termed *stripping*.

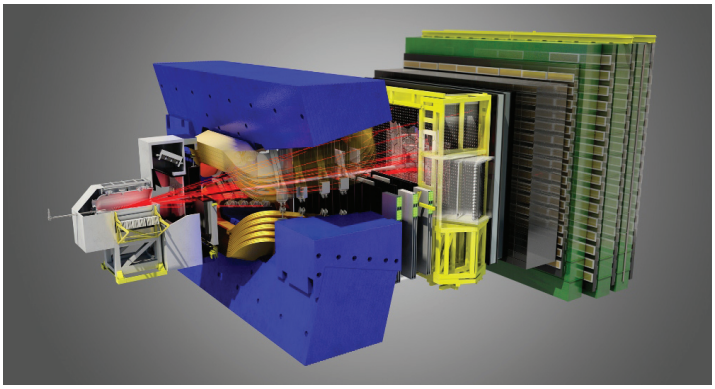
Stripping yields about twelve million $B_s^0 \to J/\psi K^+K^-$ decay candidates for 2011 and 2012. A second set of stricter offline selection criteria is applied, optimized to find the best compromise between selecting as many signal and as few background candidates as possible. At this stage about 230 thousand candidates are retained for further analysis, of which roughly 40% is signal (see Section 3.2).

The total set of selection requirements for particles and decay candidates removes background, but also a significant part of the signal. Since the efficiency of the selection depends on the decay time and the decay angles, the observed distributions of these variables do not directly reflect the underlying true distributions. This effect has to be taken into account in the model of the decay (see Sections 3.3.2 and 3.4).

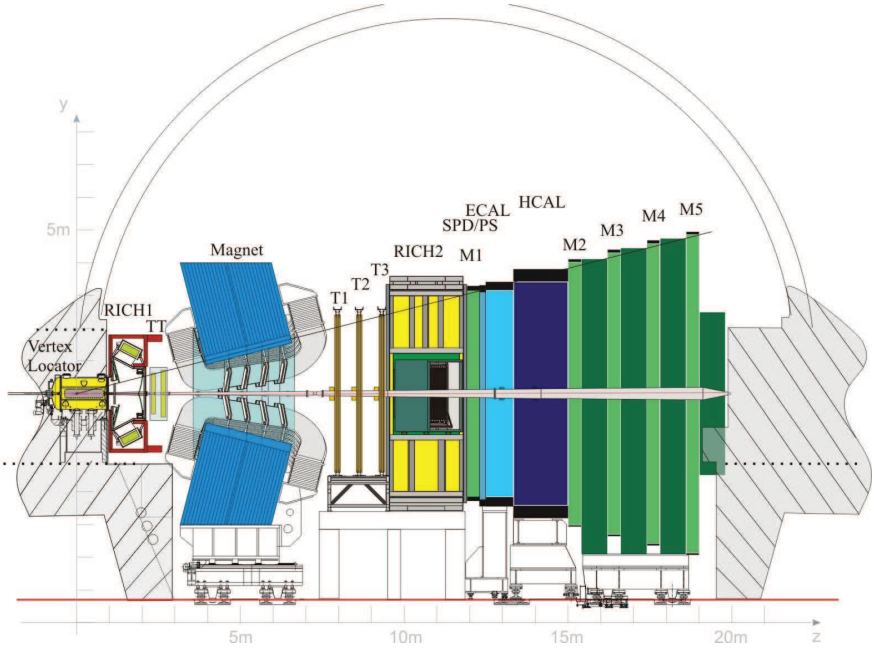
1.4.3 The LHCb Detector

The LHCb detector was mainly built to measure beauty- and strange-hadron decays into particles with an electric charge. It is capable of measuring trajectories of charged particles and the locations of primary and secondary vertices with a precision that is good enough to resolve oscillations in decay-time distributions with a frequency $\Delta m_s \approx 18\,\text{ps}$. Another important feature of the detector is its ability to identify different types of particles.

The detector and its performance are described in detail in references [55] and [56]. A schematic view is shown in Figure 1.13. The subdetectors are placed around the LHC beams, which come in from the left and right. The interaction point is contained within the subdetector labelled "Vertex Locator" on the left of Figure 1.13b. Particles that are produced within the LHCb acceptance region traverse the various subdetectors from the left to the right. This is indicated by the red lines in Figure 1.13a.



(a)



(b)

Figure 1.13: *Schematic views of the LHCb detector (courtesy of CERN). (a) Three-dimensional view, where particles from the proton–proton collision are depicted by red lines. (b) Cross sectional view of the various subdetectors, which are discussed in the text.*

A magnetic field with an integrated strength of 4 Tm is produced by a dipole magnet between approximately 3 and 8 m from the interaction point. The saddle-shaped coils of the magnet are surrounded by the magnet yoke, which is indicated by the blue block in Figure 1.13. The magnetic field is pointing either upwards or downwards, bending charged-particle trajectories in the horizontal plane. During the two years of LHC running, the direction of the magnetic field was reversed several times to control systematic uncertainties related to the measurement of particle trajectories.

Trajectories of charged particles are reconstructed by measurement of the positions at which the particles traverse the various *tracking detectors* of LHCb. These measurements are all based on electromagnetic interactions of the particle with detector material. The signal from a particle interaction with sensitive detector material is called a *hit*. The collection of hits that represent the trajectory of a particle is a *track*.

The LHCb tracking system consists of four components. The *Vertex Locator* (Velo) and the *Tracker Turicensis* (TT) before the magnet and the *Inner Tracker* (IT) and *Outer Tracker* (OT) behind the magnet. Together, the IT and OT form the three tracking stations labelled by "T1", "T2", and "T3" in Figure 1.13b. The inner region of roughly $0.5\,\mathrm{m}^2$ around the proton beams is covered by the IT and the region up to approximately 3 m from the beams by the OT.

Hit information from all tracking detectors is combined to build tracks. The momentum of a particle can be inferred from the radius of curvature of the track in the magnetic field, given the charge of the particle. Since most charged particles that live long enough to reach the detector are electrons, muons, pions, kaons, and protons, a unit charge can be assumed. The sign of the charge is determined from the direction of the track curvature.

The performance of the tracking system is shown in Figures 1.14a and 1.14b. The former figure shows the relative resolution of the momentum measurement for muon tracks in $J/\psi \to \mu^+\mu^-$ decays, which varies from roughly 0.5% to 1.1% as a function of momentum over a range of 2 to 300 GeV/c. The resulting resolution of the dimuon invariant-mass measurement is shown in the latter figure. The relative mass resolution is about 0.5% up to $10\,\mathrm{GeV}/c^2$, but increases to 1.9% at the Z^0 mass.

Vertices are reconstructed by extrapolating tracks to the point where they are closest together. The most precise information on vertex locations comes from the Velo, which detects particles very close to the interaction point down

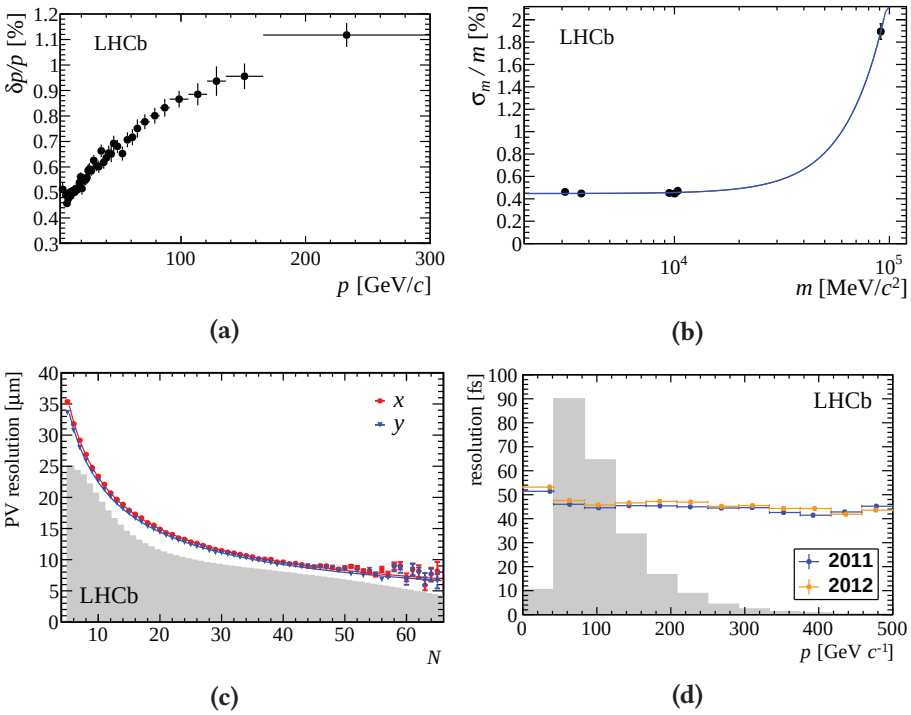


Figure 1.14: *Performance of the LHCb detector [56]. (a) Relative momentum resolution as a function of momentum for muon tracks in $J/\psi \to \mu^+\mu^-$ events. (b) Relative $\mu^+\mu^-$ invariant-mass resolution as a function of mass, measured for the J/ψ, $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$, and Z^0 resonances. (c) Primary vertex resolution as a function of the number of tracks in the vertex, separately for the two directions in the plane transverse to the beam direction. (d) Decay-time resolution of prompt $B_s^0 \to \mu^+\mu^- K^+K^-$ background as a function of the combined $\mu^+\mu^- K^+K^-$ momentum, separately for 2011 and 2012. In (c) and (d) the distribution of respectively vertices and background decay candidates in the relevant variable is indicated by the shaded histogram.*

to a radius of 8 mm around the beams.

Figure 1.14c shows the resolution of primary vertices in LHCb as a function of the number of tracks originating from the vertex. A minimum of five tracks is required, where the resolution is about 35 μm. The resolution improves with the number of tracks to less than 10 μm for more than 40 tracks.

Since only four tracks originate from the secondary vertex, its resolution dominates the uncertainty of the decay-time measurement. For $B_s^0 \to \mu^+\mu^- K^+K^-$ events the decay-time resolution is measured with prompt background, for which the true decay time is equal to zero (see Section 3.3.1). The resulting resolution as a function of $\mu^+\mu^- K^+K^-$ momentum is shown in Figure 1.14d. Its value varies between roughly 0.04 ps and 0.05 ps.

Particles are identified with other subsystems. Electrons produce signals in the *Scintillator Pad Detector* (SPD), *Preshower detector* (PS) and the *Electromagnetic Calorimeter* (ECal). The electron energy is absorbed by these detectors and no signal is produced in the *Hadronic Calorimeter* (HCal). Pions, kaons, and protons predominantly lose energy in the HCal and less in the ECal. Energy measurements from the calorimeters are also used in the L0 trigger to select events with high-energy particles in the transverse direction.

Muons do not lose significant energy in the HCal. Therefore, they make it through the calorimeters, and produce hits in the four *muon stations* behind the HCal (M2–M5 in Figure 1.13b). These features distinguish muons from hadrons. The *Muon System* is completed by the station M1 before the calorimeters to provide an optimal momentum estimate for muons.

Two *Ring Imaging Cherenkov* (RICH) detectors are used to distinguish the different charged hadrons. The first RICH detector (RICH1) is located between the Velo and the magnet and the other (RICH2) behind the IT/OT. These detectors determine the velocities of charged particles by measuring the angle at which Cherenkov light is emitted when the particle traverses the RICH radiator materials. Combined with the momentum measurement this gives an estimate of the particle mass to distinguish between pions, kaons, and protons.

Figure 1.15 shows the information from the different subdetectors for an event that contains a $B_s^0 \to J/\psi K^+K^-$ decay candidate. The coloured lines show reconstructed tracks. Three tracks are combined with hits in the muons stations and represent particles that have been identified as muons.

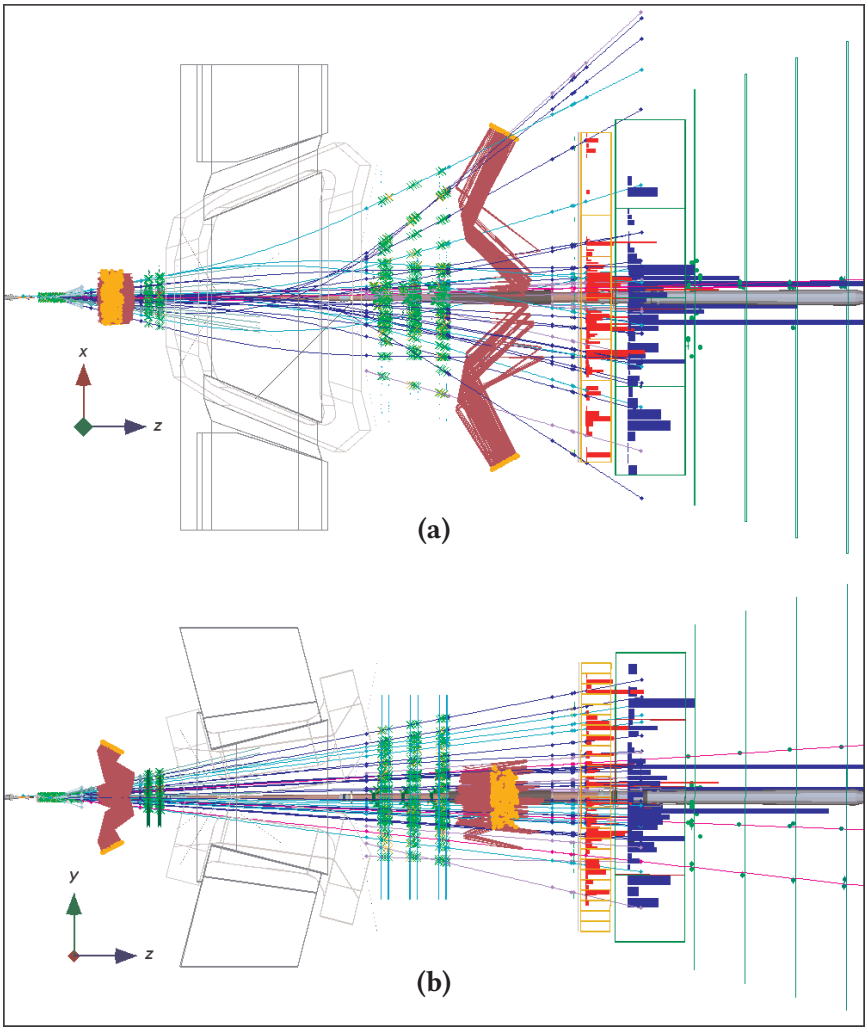


Figure 1.15: *Reconstructed particle trajectories in the LHCb detector for an event with a $B_s^0 \to J/\psi K^+K^-$ decay candidate (figure from [52]): (a) a view from the top and (b) a view from the side, equivalent to Figure 1.13. Particle tracks are shown by the coloured lines. The green crosses indicate hits in the tracking detectors and the red and blue bars energy deposits in the calorimeters. Hits in the muon stations are represented by green dots. Cherenkov photons in the RICH detectors are shown by the purple lines and orange hits.*

Chapter 2

Phenomenology

2.1 Mixing and Decay of the B_s^0–$\overline{B}_s^0$ System

2.1.1 Mixing

A meson produced in a pure B_s^0 or $\overline{B}_s^0$ state will evolve in time and become a mixture of these two flavour states. If the time coordinate in the B_s^0–$\overline{B}_s^0$ centre-of-mass system is given by t, the state of the system (Ψ) can be written as

$$|\Psi(t)\rangle = a(t)\,|B_s^0\rangle + b(t)\,|\overline{B}_s^0\rangle\ . \tag{2.1}$$

The states $|B_s^0\rangle \equiv |s\bar{b}\rangle$ and $|\overline{B}_s^0\rangle \equiv |\bar{s}b\rangle$ are the *flavour eigenstates* of the system and a and b are coefficients that describe its time dependence.

Assuming that the time scale of interest is much larger than the time scale of strong interactions, the Weisskopf-Wigner approximation can be used and the decay of the system has an exponential time dependence [57]. The time evolution then follows from a Schrödinger equation with a constant Hamiltonian, which is given by[1]

$$i\,\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \boldsymbol{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}\ . \tag{2.2}$$

Without loss of generality, the Hamiltonian matrix $\boldsymbol{H}$ can be written as the sum of a Hermitian matrix and an anti-Hermitian matrix: $\boldsymbol{H} \equiv \boldsymbol{M} -$

[1] All equations in this chapter are expressed in *natural units*, in which the reduced Planck constant and the speed of light are equal to one ($\hbar \equiv c \equiv 1$).

$\frac{i}{2}\,\boldsymbol{\Gamma}$, where the *mass matrix* $\boldsymbol{M}$ and the *decay matrix* $\boldsymbol{\Gamma}$ are both Hermitian. Assuming CPT invariance, the mass and lifetime of a particle are equal to the respective mass and lifetime of the corresponding anti-particle. This results in equal diagonal elements of both the mass and decay matrices:

$$\boldsymbol{H} \equiv \begin{pmatrix} H_0 & H_{12} \\ H_{21} & H_0 \end{pmatrix} = \boldsymbol{M} - \tfrac{i}{2}\,\boldsymbol{\Gamma} \equiv \begin{pmatrix} M_{\mathrm{s}} & M_{12} \\ M_{12}^* & M_{\mathrm{s}} \end{pmatrix} - \tfrac{i}{2} \begin{pmatrix} \Gamma_{\mathrm{s}} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{\mathrm{s}} \end{pmatrix} , \quad (2.3)$$

where M_{s} is the $\mathrm{B_s^0}$ mass and Γ_{s} the $\mathrm{B_s^0}$ decay width.

Mixing of the $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_{\mathrm{s}}^{0}$ states is governed by the off-diagonal elements of the Hamiltonian. The parameter M_{12} is the *dispersive* part of H_{12}. This part originates from contributions of virtual intermediate states to the mixing process and is dominated by diagrams with virtual top quarks (Figure 1.5 on page 14). Γ_{12} is the *absorptive* part, which originates from contributions of real states into which both $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_{\mathrm{s}}^{0}$ can decay.

The absorptive part of the mixing process is dominated by tree-level $\mathrm{b} \to \mathrm{c\bar{c}s}$ transitions [44]. Following the arguments from Section 1.3.2, this means that effects of physics beyond the Standard Model in Γ_{12} are expected to be small. The virtual loop process that dominates M_{12}, on the other hand, can be affected significantly and may lead to a deviation in the phase ϕ_{s} with respect to the Standard Model prediction.

To solve Equation 2.2 and obtain expressions for the time evolution of the $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_{\mathrm{s}}^{0}$ states, the system is decoupled with a transformation of the flavour eigenstates that diagonalizes the Hamiltonian matrix. The decoupled states are *mass eigenstates*, which have definite mass and lifetime. With transformation matrix $\boldsymbol{P}$, diagonalized Hamiltonian $\boldsymbol{H'}$ and mass-eigenstate coefficients a' and b', the transformation is specified by

$$\boldsymbol{H'} = \boldsymbol{P}^{-1}\,\boldsymbol{H}\,\boldsymbol{P} \qquad \text{and} \qquad \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} . \quad (2.4)$$

The two eigenvalues of the matrix $\boldsymbol{H}$ are $H_0 \mp \sqrt{H_{12}H_{21}}$, with corresponding eigenvectors $\left(\sqrt{H_{12}} \quad \mp\sqrt{H_{21}}\right)$. The eigenvalues become the diagonal elements of the decoupled Hamiltonian $\boldsymbol{H'}$ and the transformation matrix is constructed from the eigenvectors:

$$\boldsymbol{H'} \equiv \begin{pmatrix} M_{\mathrm{L}} & 0 \\ 0 & M_{\mathrm{H}} \end{pmatrix} - \tfrac{i}{2} \begin{pmatrix} \Gamma_{\mathrm{L}} & 0 \\ 0 & \Gamma_{\mathrm{H}} \end{pmatrix} \quad (2.5a)$$

$$= \begin{pmatrix} H_0 - \sqrt{H_{12}H_{21}} & 0 \\ 0 & H_0 + \sqrt{H_{12}H_{21}} \end{pmatrix} \tag{2.5b}$$

$$\boldsymbol{P} = \begin{pmatrix} \sqrt{H_{12}} & \sqrt{H_{12}} \\ -\sqrt{H_{21}} & +\sqrt{H_{21}} \end{pmatrix} , \tag{2.5c}$$

where the subscript L (light) is used for the state with the smaller mass and the subscript H (heavy) for the state with the larger mass.

Mass and decay parameters of B_s^0 and $\overline{B}_s^0$ are related to the masses and decay widths of the diagonalized B_L and B_H states by Equations 2.5a and 2.5b. Taking the sum and the difference of the two values on the diagonal of $\boldsymbol{H'}$ yields

$$H_0 = \tfrac{1}{2}\,(M_H + M_L) - \tfrac{i}{4}\,(\Gamma_L + \Gamma_H) \tag{2.6a}$$

$$\sqrt{H_{12}H_{21}} = \tfrac{1}{2}\,(M_H - M_L) + \tfrac{i}{4}\,(\Gamma_L - \Gamma_H)\,. \tag{2.6b}$$

Using also Equation 2.3, it is found that

$$M_s \equiv \quad \Re(H_0) \;=\; \tfrac{1}{2}\,(M_H + M_L) \tag{2.7a}$$

$$\Gamma_s \equiv -2\,\Im(H_0) \;=\; \tfrac{1}{2}\,(\Gamma_L + \Gamma_H) \tag{2.7b}$$

and the following quantities can be defined:

$$\begin{aligned} \Delta m_s &\equiv M_H - M_L = 2\,\Re\Big(\sqrt{H_{12}H_{21}}\Big) \\ &= 2\,\Re\Bigg(\sqrt{(M_{12} - \tfrac{i}{2}\,\Gamma_{12})\,(M_{12}^* - \tfrac{i}{2}\,\Gamma_{12}^*)}\Bigg) \end{aligned} \tag{2.8a}$$

$$\begin{aligned} \Delta\Gamma_s &\equiv \Gamma_L - \Gamma_H = 4\,\Im\Big(\sqrt{H_{12}H_{21}}\Big) \\ &= 4\,\Im\Bigg(\sqrt{(M_{12} - \tfrac{i}{2}\,\Gamma_{12})\,(M_{12}^* - \tfrac{i}{2}\,\Gamma_{12}^*)}\Bigg)\;, \end{aligned} \tag{2.8b}$$

where a convention is chosen in which $\Delta\Gamma_s$ is positive in the Standard Model. With these definitions of Δm_s and $\Delta\Gamma_s$, the expression for $\sqrt{H_{12}H_{21}}$ reads

$$\sqrt{H_{12}H_{21}} = \sqrt{(M_{12} - \tfrac{i}{2}\,\Gamma_{12})\,(M_{12}^* - \tfrac{i}{2}\,\Gamma_{12}^*)} = \tfrac{1}{2}\,\Delta m_s + \tfrac{i}{4}\,\Delta\Gamma_s\,. \tag{2.9}$$

Squaring this equation gives

$$4\,|M_{12}|^2 - |\Gamma_{12}|^2 - 4i\,\Re(M_{12}\,\Gamma_{12}^*) = \Delta m_s^2 - \tfrac{1}{4}\,\Delta\Gamma_s^2 + i\,\Delta m_s\,\Delta\Gamma_s\,, \tag{2.10}$$

which results in the relations

$$\Delta m_s^2 - \tfrac{1}{4}\,\Delta\Gamma_s^2 = 4\,|M_{12}|^2 - |\Gamma_{12}|^2 \tag{2.11a}$$

$$\Delta m_s\,\Delta\Gamma_s = -4\,\Re(M_{12}\,\Gamma_{12}^*) = 4\,|M_{12}|\,|\Gamma_{12}|\,\cos\phi_{12}\ . \tag{2.11b}$$

The parameter ϕ_{12} is defined as the phase difference between M_{12} and Γ_{12}: $\phi_{12} \equiv \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right)$.

To find the transformation between flavour eigenstates and mass eigenstates, the state of Equation 2.1 is expressed in matrix form and the transformation of Equation 2.4 is applied:

$$\begin{aligned} |\Psi(t)\rangle &= \begin{pmatrix} a'(t) & b'(t) \end{pmatrix} \begin{pmatrix} |\mathrm{B_L}\rangle \\ |\mathrm{B_H}\rangle \end{pmatrix} \\ &= \begin{pmatrix} a(t) & b(t) \end{pmatrix} \begin{pmatrix} |\mathrm{B_s^0}\rangle \\ |\overline{\mathrm{B}}_\mathrm{s}^0\rangle \end{pmatrix} = \begin{pmatrix} a'(t) & b'(t) \end{pmatrix} \boldsymbol{P}^\mathrm{T} \begin{pmatrix} |\mathrm{B_s^0}\rangle \\ |\overline{\mathrm{B}}_\mathrm{s}^0\rangle \end{pmatrix}\ . \end{aligned} \tag{2.12}$$

By comparing the first and second line of Equation 2.12 it can be seen that the transformation matrix between flavour eigenstates and mass eigenstates is the transpose of the matrix that diagonalizes the Hamiltonian (Equation 2.5c).

Introducing the complex parameters p and q and normalizing the mass eigenstates, the transformation to flavour eigenstates can be expressed as

$$\begin{pmatrix} |\mathrm{B_L}\rangle \\ |\mathrm{B_H}\rangle \end{pmatrix} = \boldsymbol{P}^\mathrm{T} \begin{pmatrix} |\mathrm{B_s^0}\rangle \\ |\overline{\mathrm{B}}_\mathrm{s}^0\rangle \end{pmatrix} = \begin{pmatrix} p & +q \\ p & -q \end{pmatrix} \begin{pmatrix} |\mathrm{B_s^0}\rangle \\ |\overline{\mathrm{B}}_\mathrm{s}^0\rangle \end{pmatrix} \qquad \text{with } |p|^2 + |q|^2 \equiv 1\ . \tag{2.13}$$

Combining Equations 2.3, 2.5c, and 2.13 results in

$$\frac{q}{p} = -\sqrt{\frac{H_{21}}{H_{12}}} = -\sqrt{\frac{M_{12}^* - \frac{i}{2}\,\Gamma_{12}^*}{M_{12} - \frac{i}{2}\,\Gamma_{12}}}\quad . \tag{2.14}$$

The time evolution of the mass eigenstates is obtained by solving the Schrödinger equation with the diagonal Hamiltonian of Equation 2.5, which gives decoupled exponential decays for $\mathrm{B_L}$ and $\mathrm{B_H}$. Transforming back to the flavour basis then gives the coupled time evolution of $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}_\mathrm{s}^0$:

$$\begin{aligned} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} &= e^{-i\,\boldsymbol{H}\,t} \begin{pmatrix} a(0) \\ b(0) \end{pmatrix} \\ &= \boldsymbol{P} \begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} = \boldsymbol{P}\, e^{-i\,\boldsymbol{H'}\,t} \begin{pmatrix} a'(0) \\ b'(0) \end{pmatrix} = \boldsymbol{P}\, e^{-i\,\boldsymbol{H'}\,t}\, \boldsymbol{P}^{-1} \begin{pmatrix} a(0) \\ b(0) \end{pmatrix} \end{aligned} \tag{2.15a}$$

$$
\begin{aligned}
e^{-i\boldsymbol{H}t} &= \boldsymbol{P}\, e^{-i\boldsymbol{H'}t}\, \boldsymbol{P}^{-1} \\
&= \begin{pmatrix} p & p \\ q & -q \end{pmatrix} \begin{pmatrix} e^{-i\,(M_L - i\,\Gamma_L/2)\,t} & 0 \\ 0 & e^{-i\,(M_H - i\,\Gamma_H/2)\,t} \end{pmatrix} \begin{pmatrix} q & p \\ q & -p \end{pmatrix} \frac{1}{2\,p\,q} \\
&= \begin{pmatrix} g_+(t) & \frac{p}{q}\, g_-(t) \\ \frac{q}{p}\, g_-(t) & g_+(t) \end{pmatrix} , \qquad (2.15b)
\end{aligned}
$$

where the functions $g_\pm$ are given by

$$
\begin{aligned}
g_\pm(t) &\equiv \tfrac{1}{2} \left(e^{-i\,(M_L - i\,\Gamma_L/2)\,t} \pm e^{-i\,(M_H - i\,\Gamma_H/2)\,t} \right) \qquad (2.16) \\
&= \tfrac{1}{2}\, e^{-i\,M_s\,t}\, e^{-\Gamma_s\,t/2} \left(e^{+i\,\Delta m_s\,t/2}\, e^{-\Delta\Gamma_s\,t/4} \pm e^{-i\,\Delta m_s\,t/2}\, e^{+\Delta\Gamma_s\,t/4} \right) .
\end{aligned}
$$

The time evolution for mesons produced as B_s^0 ($a(0) = 1$ and $b(0) = 0$) and mesons produced as $\overline{B}_s^0$ ($a(0) = 0$ and $b(0) = 1$) can now be inferred from Equations 2.1 and 2.15:

$$
|\Psi_{B_s^0}(t)\rangle = g_+(t)\, |B_s^0\rangle + \frac{q}{p}\, g_-(t)\, |\overline{B}_s^0\rangle \qquad (2.17a)
$$

$$
|\Psi_{\overline{B}_s^0}(t)\rangle = g_+(t)\, |\overline{B}_s^0\rangle + \frac{p}{q}\, g_-(t)\, |B_s^0\rangle \qquad (2.17b)
$$

2.1.2 Mixing and Decay

Time-dependent amplitudes for mixing and decay of the B_s^0–$\overline{B}_s^0$ system are obtained by combining the state of Equation 2.1 with the amplitudes for the decays of $|B_s^0\rangle$ and $|\overline{B}_s^0\rangle$. Assuming the system is produced as either $|B_s^0\rangle$ or $|\overline{B}_s^0\rangle$, the required time-dependent states are given by Equation 2.17. The decay amplitude of a decay of $|B_s^0\rangle$ into a final state $|f\rangle$ is labelled by $\mathcal{A}_f$. The amplitude for a system produced as a B_s^0, being in a B_s^0 state at the time of decay and decaying into $|f\rangle$ is given by $\langle B_s^0|\Psi_{B_s^0}(t)\rangle\, \mathcal{A}_f \propto g_+(t)\, \mathcal{A}_f$.

If also $|\overline{B}_s^0\rangle$ can decay into the final state there is a second contribution to the $B_s^0 \to f$ process, which is proportional to the decay amplitude labelled by $\overline{\mathcal{A}}_f$. Also considering decays into the CP conjugate of the final state, $|\bar{f}\rangle$, the amplitudes of the four possible combinations of initial and final states are

given by

$$\begin{aligned}
\mathcal{A}(\mathrm{B_s^0}\to \mathrm{f}) &\propto \quad g_+\,\mathcal{A}_\mathrm{f} + \frac{q}{p}\,g_-\,\overline{\mathcal{A}}_\mathrm{f} \quad ; \qquad \mathcal{A}(\mathrm{B_s^0}\to \bar{\mathrm{f}}) \propto \frac{q}{p}\Big(g_-\,\overline{\mathcal{A}}_{\bar{\mathrm{f}}} + \frac{p}{q}\,g_+\,\mathcal{A}_{\bar{\mathrm{f}}}\Big) \\
\mathcal{A}(\overline{\mathrm{B}}_\mathrm{s}^0\to \mathrm{f}) &\propto \frac{p}{q}\Big(g_-\,\mathcal{A}_\mathrm{f} + \frac{q}{p}\,g_+\,\overline{\mathcal{A}}_\mathrm{f}\Big)\,; \qquad \mathcal{A}(\overline{\mathrm{B}}_\mathrm{s}^0\to \bar{\mathrm{f}}) \propto \quad g_+\,\overline{\mathcal{A}}_{\bar{\mathrm{f}}} + \frac{p}{q}\,g_-\,\mathcal{A}_{\bar{\mathrm{f}}}\,.
\end{aligned} \tag{2.18}$$

Notice that the amplitudes for the processes with final state $|\bar{\mathrm{f}}\rangle$ have the same structure as the amplitudes for the processes with final state $|\mathrm{f}\rangle$. $\mathcal{A}(\overline{\mathrm{B}}_\mathrm{s}^0\to \bar{\mathrm{f}})$ and $\mathcal{A}(\mathrm{B_s^0}\to \bar{\mathrm{f}})$ can be obtained from $\mathcal{A}(\mathrm{B_s^0}\to \mathrm{f})$ and $\mathcal{A}(\overline{\mathrm{B}}_\mathrm{s}^0\to \mathrm{f})$, respectively, by interchanging p and q, replacing $\mathcal{A}_\mathrm{f}$ by $\overline{\mathcal{A}}_{\bar{\mathrm{f}}}$ and replacing $\overline{\mathcal{A}}_\mathrm{f}$ by $\mathcal{A}_{\bar{\mathrm{f}}}$. The amplitude for $\overline{\mathrm{B}}_\mathrm{s}^0\to \mathrm{f}$ $(\mathrm{B_s^0}\to \bar{\mathrm{f}})$ is obtained from the amplitude for $\mathrm{B_s^0}\to \mathrm{f}$ $(\overline{\mathrm{B}}_\mathrm{s}^0\to \bar{\mathrm{f}})$ by interchanging g_+ and g_- and multiplying by a factor $\frac{p}{q}$ $\left(\frac{q}{p}\right)$.

The magnitudes of the amplitudes in Equation 2.18 are squared to obtain an expression for the differential decay rates in time. For the $\mathrm{B_s^0}\to \mathrm{f}$ amplitude this yields

$$\begin{aligned}
\frac{\mathrm{d}\Gamma(\mathrm{B_s^0}\to \mathrm{f})}{\mathrm{d}t} &\propto \Big|\mathcal{A}(\mathrm{B_s^0}\to \mathrm{f})\Big|^2 \\
&\propto |g_+|^2\,|\mathcal{A}_\mathrm{f}|^2 + \left|\frac{q}{p}\right|^2 |g_-|^2\,|\overline{\mathcal{A}}_\mathrm{f}|^2 + \frac{q}{p}\,g_+^*\,g_-\,\mathcal{A}_\mathrm{f}^*\,\overline{\mathcal{A}}_\mathrm{f} + \left(\frac{q}{p}\,g_+^*\,g_-\,\mathcal{A}_\mathrm{f}^*\,\overline{\mathcal{A}}_\mathrm{f}\right)^* \\
&\propto |g_+|^2\,|\mathcal{A}_\mathrm{f}|^2 + \left|\frac{q}{p}\right|^2 |g_-|^2\,|\overline{\mathcal{A}}_\mathrm{f}|^2 \\
&\qquad + 2\,\Re(g_+^*\,g_-)\,\Re\left(\frac{q}{p}\,\mathcal{A}_\mathrm{f}^*\,\overline{\mathcal{A}}_\mathrm{f}\right) - 2\,\Im(g_+^*\,g_-)\,\Im\left(\frac{q}{p}\,\mathcal{A}_\mathrm{f}^*\,\overline{\mathcal{A}}_\mathrm{f}\right)\,.
\end{aligned} \tag{2.19}$$

Using the definition of $g_\pm$ from Equation 2.16, the required products of g_+ and g_- are given by

$$|g_\pm|^2 = \tfrac{1}{2}\,e^{-\Gamma_\mathrm{s}\,t}\Big[\ \cosh\big(\tfrac{1}{2}\Delta\Gamma_\mathrm{s}\,t\big)\ \pm \cos(\Delta m_\mathrm{s}\,t)\Big] \tag{2.20a}$$

$$g_+^*\,g_- = \tfrac{1}{2}\,e^{-\Gamma_\mathrm{s}\,t}\Big[-\sinh\big(\tfrac{1}{2}\Delta\Gamma_\mathrm{s}\,t\big) + i\ \sin(\Delta m_\mathrm{s}\,t)\Big]\,, \tag{2.20b}$$

which yields

$$\begin{aligned}\frac{\mathrm{d}\Gamma(\mathrm{B}_\mathrm{s}^0\to\mathrm{f})}{\mathrm{d}t} \propto \tfrac{1}{2}\, e^{-\Gamma_\mathrm{s} t}\Bigg[&\left(|\mathcal{A}_\mathrm{f}|^2 + \left|\frac{q}{p}\right|^2 |\overline{\mathcal{A}}_\mathrm{f}|^2\right) \cosh\!\left(\tfrac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right) \\ &+ \left(|\mathcal{A}_\mathrm{f}|^2 - \left|\frac{q}{p}\right|^2 |\overline{\mathcal{A}}_\mathrm{f}|^2\right) \cos(\Delta m_\mathrm{s}\, t) \\ &- 2\,\Re\!\left(\frac{q}{p}\,\mathcal{A}_\mathrm{f}^*\,\overline{\mathcal{A}}_\mathrm{f}\right) \sinh\!\left(\tfrac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right) \\ &- 2\,\Im\!\left(\frac{q}{p}\,\mathcal{A}_\mathrm{f}^*\,\overline{\mathcal{A}}_\mathrm{f}\right) \sin(\Delta m_\mathrm{s}\, t)\Bigg] .\end{aligned} \tag{2.21}$$

With the definitions

$$\lambda_\mathrm{f} \equiv \frac{q}{p}\frac{\overline{\mathcal{A}}_\mathrm{f}}{\mathcal{A}_\mathrm{f}} \tag{2.22a}$$

$$C_\mathrm{f} \equiv \frac{1-|\lambda_\mathrm{f}|^2}{1+|\lambda_\mathrm{f}|^2} \qquad D_\mathrm{f} \equiv -\frac{2\,\Re(\lambda_\mathrm{f})}{1+|\lambda_\mathrm{f}|^2} \qquad S_\mathrm{f} \equiv \frac{2\,\Im(\lambda_\mathrm{f})}{1+|\lambda_\mathrm{f}|^2}\,, \tag{2.22b}$$

the differential decay rate can be expressed as

$$\begin{aligned}\frac{\mathrm{d}\Gamma(\mathrm{B}_\mathrm{s}^0\to\mathrm{f})}{\mathrm{d}t} \propto \tfrac{1}{2}\,|\mathcal{A}_\mathrm{f}|^2\,(1+|\lambda_\mathrm{f}|^2)\,e^{-\Gamma_\mathrm{s} t} \\ \times\Big[\cosh\!\left(\tfrac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right) + C_\mathrm{f}\cos(\Delta m_\mathrm{s}\, t) \\ + D_\mathrm{f}\sinh\!\left(\tfrac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right) - S_\mathrm{f}\sin(\Delta m_\mathrm{s}\, t)\Big]\,,\end{aligned} \tag{2.23}$$

Expressions for the differential rates of the remaining three processes can be obtained by applying the aforementioned relations between the four amplitudes. From Equation 2.20 it can be seen that interchanging g_+ and g_- results in a sign change of the $\cos(\Delta m_\mathrm{s}\, t)$ and $\sin(\Delta m_\mathrm{s}\, t)$ terms. For the $|\bar{\mathrm{f}}\rangle$ final state the parameter λ_f goes to $\overline{\lambda}_{\bar{\mathrm{f}}}$:

$$\overline{\lambda}_{\bar{\mathrm{f}}} \equiv \frac{p}{q}\frac{\mathcal{A}_{\bar{\mathrm{f}}}}{\overline{\mathcal{A}}_{\bar{\mathrm{f}}}} \tag{2.24a}$$

$$C_{\bar{\mathrm{f}}} \equiv \frac{1-|\overline{\lambda}_{\bar{\mathrm{f}}}|^2}{1+|\overline{\lambda}_{\bar{\mathrm{f}}}|^2} \qquad D_{\bar{\mathrm{f}}} \equiv -\frac{2\,\Re(\overline{\lambda}_{\bar{\mathrm{f}}})}{1+|\overline{\lambda}_{\bar{\mathrm{f}}}|^2} \qquad S_{\bar{\mathrm{f}}} \equiv \frac{2\,\Im(\overline{\lambda}_{\bar{\mathrm{f}}})}{1+|\overline{\lambda}_{\bar{\mathrm{f}}}|^2}\,. \tag{2.24b}$$

In addition, an asymmetry for CP-violation in mixing is defined as

$$C_{\text{mix}} \equiv \frac{1 - \left|\frac{q}{p}\right|^2}{1 + \left|\frac{q}{p}\right|^2} \,. \tag{2.25}$$

This parameter is not to be confused with the complex-valued asymmetry ϵ, which is generally used to parameterize CP violation in the neutral-kaon system and is given by

$$\epsilon \equiv \frac{1 - \frac{q}{p}}{1 + \frac{q}{p}} \qquad C_{\text{mix}} \approx 2\,\Re(\epsilon) \,, \tag{2.26}$$

where the approximation is first order in $|\epsilon|$ for values close to zero.

With the parameter C_{mix}, the factors $\left|\frac{p}{q}\right|^2$ and $\left|\frac{q}{p}\right|^2$ for the $\overline{\text{B}}{}^0_{\text{s}} \to \text{f}$ and $\text{B}^0_{\text{s}} \to \bar{\text{f}}$ decay rates, respectively, are given by

$$\left|\frac{p}{q}\right|^2 = \frac{1 + C_{\text{mix}}}{1 - C_{\text{mix}}} \qquad \left|\frac{q}{p}\right|^2 = \frac{1 - C_{\text{mix}}}{1 + C_{\text{mix}}} \,. \tag{2.27}$$

The differential decay rates can now be expressed as

$$\begin{aligned} \frac{\mathrm{d}\Gamma(\text{f})}{\mathrm{d}t} \propto{}& \tfrac{1}{2}\,|\mathcal{A}_{\text{f}}|^2\,(1 + |\lambda_{\text{f}}|^2)\,\frac{1 - q_{\text{f}}\,C_{\text{mix}}}{1 - C_{\text{mix}}}\,e^{-\Gamma_{\text{s}}\,t} \\ &\times \Big[\cosh\!\big(\tfrac{1}{2}\Delta\Gamma_{\text{s}}\,t\big) + q_{\text{f}}\,C_{\text{f}}\,\cos(\Delta m_{\text{s}}\,t) \\ &\quad + D_{\text{f}}\,\sinh\!\big(\tfrac{1}{2}\Delta\Gamma_{\text{s}}\,t\big) - q_{\text{f}}\,S_{\text{f}}\,\sin(\Delta m_{\text{s}}\,t)\Big] \end{aligned} \tag{2.28a}$$

$$\begin{aligned} \frac{\mathrm{d}\Gamma\big(\overline{\text{f}}\big)}{\mathrm{d}t} \propto{}& \tfrac{1}{2}\,|\overline{\mathcal{A}}_{\bar{\text{f}}}|^2\,(1 + |\overline{\lambda}_{\bar{\text{f}}}|^2)\,\frac{1 - q_{\text{f}}\,C_{\text{mix}}}{1 + C_{\text{mix}}}\,e^{-\Gamma_{\text{s}}\,t} \\ &\times \Big[\cosh\!\big(\tfrac{1}{2}\Delta\Gamma_{\text{s}}\,t\big) - q_{\text{f}}\,C_{\bar{\text{f}}}\,\cos(\Delta m_{\text{s}}\,t) \\ &\quad + D_{\bar{\text{f}}}\,\sinh\!\big(\tfrac{1}{2}\Delta\Gamma_{\text{s}}\,t\big) + q_{\text{f}}\,S_{\bar{\text{f}}}\,\sin(\Delta m_{\text{s}}\,t)\Big] \,, \end{aligned} \tag{2.28b}$$

where the variable q_{f} takes the value +1 for a B^0_{s} initial state and −1 for a $\overline{\text{B}}{}^0_{\text{s}}$ initial state.

For flavour-specific final states, where $|\text{B}^0_{\text{s}}\rangle$ can only decay into $|\text{f}\rangle$ and $|\overline{\text{B}}{}^0_{\text{s}}\rangle$ only into $|\bar{\text{f}}\rangle$, the parameters λ_{f} and $\overline{\lambda}_{\bar{\text{f}}}$ vanish. In that case D_{f}, S_{f}, $D_{\bar{\text{f}}}$ and $S_{\bar{\text{f}}}$

are equal to zero and C_f and $C_{\bar{f}}$ are equal to one. Only the $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ and $\cos(\Delta m_s\, t)$ terms then remain in the expressions for the differential decay rates. An example of this case are the CP-conjugate processes $B_s^0 \to D_s^- \pi^+$ and $\overline{B}_s^0 \to D_s^+ \pi^-$.

Another special case is a CP eigenstate, for which $|f\rangle$ and $|\bar{f}\rangle$ are the same. As a result, $|\mathcal{A}_{\bar{f}}| = |\mathcal{A}_f|$, $|\overline{\mathcal{A}}_{\bar{f}}| = |\overline{\mathcal{A}}_f|$, $\overline{\lambda}_{\bar{f}} = \frac{1}{\lambda_f}$, $C_{\bar{f}} = -C_f$, $D_{\bar{f}} = +D_f$, and $S_{\bar{f}} = -S_f$, which makes Equations 2.28a and 2.28b identical. The $B_s^0 \to J/\psi\, K^+K^-$ decay proceeds via several intermediate CP eigenstates and its decay-time dependence will be discussed in Sections 2.2 and 2.3.

2.1.3 CP-Violation Observables

All three types of CP violation discussed in Section 1.3.1 affect the differential decay rates in Equation 2.28. Notice that the expressions are CP symmetric if $C_{mix} = 0$, $|\overline{\mathcal{A}}_{\bar{f}}| = |\mathcal{A}_f|$, and $\overline{\lambda}_{\bar{f}} = \lambda_f$. In that case the decay rates of $B_s^0 \to f$ and $\overline{B}_s^0 \to \bar{f}$ are equal and the decay rates of $\overline{B}_s^0 \to f$ and $B_s^0 \to \bar{f}$ are equal.

CP violation in mixing would give rise to a $\left|\frac{q}{p}\right|$ that is not equal to one. As a result, C_{mix} would have a nonzero value and the magnitudes of λ_f and $\overline{\lambda}_{\bar{f}}$ would be different. CP violation in decay leads to $|\mathcal{A}_f| \neq |\overline{\mathcal{A}}_{\bar{f}}|$ and/or $|\overline{\mathcal{A}}_f| \neq |\mathcal{A}_{\bar{f}}|$, which also gives a difference in $|\lambda_f|$ and $|\overline{\lambda}_{\bar{f}}|$.

Although neither the phase of the ratio $\frac{q}{p}$ nor the phases of the amplitudes are directly observable, the phases of both λ_f and $\overline{\lambda}_{\bar{f}}$ are. CP violation in the interference of decays with and without mixing gives λ_f and $\overline{\lambda}_{\bar{f}}$ different phases, resulting in $D_f \neq D_{\bar{f}}$ and $S_f \neq S_{\bar{f}}$.

If the final state is a CP eigenstate, λ_f and $\overline{\lambda}_{\bar{f}}$ can only be equal if they are real and equal to (minus) one, as a consequence of the relation $\overline{\lambda}_{\bar{f}} = \frac{1}{\lambda_f}$. In other words, CP symmetry is violated if $|\lambda_f| \neq 1$ or $\Im(\lambda_f) \neq 0$ in this case. For the $B_s^0 \to J/\psi\, K^+K^-$ decay the complex phases of the λ_f parameters, and therefore their imaginary parts, are parameterized with ϕ_s (see Sections 2.2 and 2.3.1). The contribution from B_s^0–$\overline{B}_s^0$ mixing to ϕ_s is introduced by the phase of the ratio $\frac{q}{p}$.

The CP-violation observables ϕ_s and C_{mix} are related to the parameters Δm_s and $\Delta\Gamma_s$ through the mixing process. The parameters $\frac{q}{p}$, Δm_s, and $\Delta\Gamma_s$ are all defined by the off-diagonal elements of the mixing Hamiltonian matrix, as shown in Equations 2.7–2.11 and 2.14. The observed parameter values allow for an approximation of these expressions, which simplifies the relations between the parameters.

The parameters Δm_s and $\Delta\Gamma_s$ are measured to be $17.768 \pm 0.024\,\text{ps}^{-1}$ [58] and $0.091 \pm 0.011\,\text{ps}^{-1}$ [32], respectively, and hence $|\Delta\Gamma_s| \ll \Delta m_s$. In the Standard Model the magnitude of Γ_{12} is predicted to be approximately equal to $0.04\,\text{ps}^{-1}$ [44]. Since effects from potential contributions beyond the Standard Model on Γ_{12} are expected to be relatively small, also the relation $|\Gamma_{12}| \ll \Delta m_s$ holds.

Using Equation 2.11a with the above parameter values, the magnitude of M_{12} is approximated by $|M_{12}| \approx \frac{1}{2}\Delta m_s$, and consequently $|\Gamma_{12}| \ll |M_{12}|$. This last inequality enables an expansion in the ratio $\left|\frac{\Gamma_{12}}{M_{12}}\right| \approx 5 \cdot 10^{-3}$ for the parameters $\frac{q}{p}$, Δm_s, and $\Delta\Gamma_s$. Starting from Equations 2.14, 2.8a, and 2.8b the following expansions can be derived:

$$\frac{q}{p} = -e^{-i\phi_M}\left[1 - \tfrac{1}{2}\left|\frac{\Gamma_{12}}{M_{12}}\right|\sin\phi_{12} + \tfrac{1}{8}\left|\frac{\Gamma_{12}}{M_{12}}\right|^2 \sin^2\phi_{12} + \tfrac{i}{4}\left|\frac{\Gamma_{12}}{M_{12}}\right|^2 \sin 2\phi_{12} + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^3\right)\right] \tag{2.29a}$$

$$\Delta m_s = 2\,|M_{12}|\left[1 - \tfrac{1}{8}\left|\frac{\Gamma_{12}}{M_{12}}\right|^2 \sin^2\phi_{12} + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^4\right)\right] \tag{2.29b}$$

$$\Delta\Gamma_s = 2\,|\Gamma_{12}|\cos\phi_{12}\left[1 + \tfrac{1}{8}\left|\frac{\Gamma_{12}}{M_{12}}\right|^2 \sin^2\phi_{12} + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^4\right)\right], \tag{2.29c}$$

where $\phi_M \equiv \arg(M_{12})$ and $\phi_{12} \equiv \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right)$. Notice that the product of the third order Δm_s and $\Delta\Gamma_s$ approximations is equal to $4\,|M_{12}||\Gamma_{12}|\cos\phi_{12}$ at that order, as required by the relation in Equation 2.11b.

The phase ϕ_M is convention dependent and cannot be observed directly, but contributes to the phases of λ_f and $\overline{\lambda}_{\bar{f}}$ (see Equations 2.22a, 2.24a, and 2.29a). In a first order approximation in $\left|\frac{\Gamma_{12}}{M_{12}}\right|$, this is the only mixing-induced contribution to these phases, since other complex phases only enter at order $\left|\frac{\Gamma_{12}}{M_{12}}\right|^2$. Consequently, only (the phase of) M_{12} is expected to contribute significantly to CP violation in the interference of decays with and without mixing and not Γ_{12}.

In the Standard Model, ϕ_M arises from the phases of the V_{ts} and V_{tb} CKM-matrix elements in the dominant contribution to the mixing process. In combination with the phases in the $B_s^0 \to J/\psi K^+K^-$ decay process, this leads to $\phi_s \approx -2\beta_s = -0.0368^{+0.0013}_{-0.0014}$ rad, as shown in Section 1.3.2, Equation 1.15.

The phase difference ϕ_{12} is directly observable and governs CP violation

in mixing. The Standard Model prediction is $\phi_{12} = 0.0038 \pm 0.0010\,\text{rad}$ [44]. With Equations 2.25 and 2.29a, the CP asymmetry C_{mix} is approximated at first order by

$$C_{\text{mix}} \equiv \frac{1 - \left|\frac{q}{p}\right|^2}{1 + \left|\frac{q}{p}\right|^2} \approx \frac{1}{2}\left|\frac{\Gamma_{12}}{M_{12}}\right| \sin\phi_{12} \,. \tag{2.30}$$

This relates to the asymmetry between the $\overline{\text{B}}{}^0_{\text{s}} \to \text{f}$ and $\text{B}^0_{\text{s}} \to \bar{\text{f}}$ rates in flavour-specific decays, a_{fs}, as

$$a_{\text{fs}} \equiv \frac{\left|\frac{p}{q}\right|^2 - \left|\frac{q}{p}\right|^2}{\left|\frac{p}{q}\right|^2 + \left|\frac{q}{p}\right|^2} = \frac{(1 + C_{\text{mix}})^2 - (1 - C_{\text{mix}})^2}{(1 + C_{\text{mix}})^2 + (1 - C_{\text{mix}})^2} \approx 2\,C_{\text{mix}} \,, \tag{2.31}$$

which is predicted to be $(1.9 \pm 0.3) \cdot 10^{-5}$ in the Standard Model. With Equations 2.29, 2.30, and 2.31, an approximate relation between a_{fs}, Δm_{s}, and $\Delta\Gamma_{\text{s}}$ can be derived:

$$a_{\text{fs}} \approx \left|\frac{\Gamma_{12}}{M_{12}}\right| \sin\phi_{12} \approx \frac{\Delta\Gamma_{\text{s}}}{\Delta m_{\text{s}}} \tan\phi_{12} \,. \tag{2.32}$$

Under the assumption that physics beyond the Standard Model only affects M_{12} and not Γ_{12}, deviations in ϕ_M and ϕ_{12} must be equal. Also assuming that CP violation in $\text{B}^0_{\text{s}} \to \text{J}/\psi\,\text{K}^+\text{K}^-$ is fully induced by mixing (M_{12}), the same deviation should also be found in ϕ_{s}.

With these assumptions, Equation 2.32 can be used as a constraint in the ϕ_{s}–$\Delta\Gamma_{\text{s}}$ plane, given the measured values of a_{fs} and Δm_{s} and with the difference ϕ_{s}–ϕ_{12} fixed to the Standard Model prediction. The resulting 68% and 95% confidence-level contours are shown in Figure 2.1, together with the 68% confidence-level contour from the combination of direct measurements (Figure 1.8 on page 20). With the current precision, the measurements are compatible at 95% confidence level.

2.2 Differential Decay Rate

With four particles in the final state, the kinematics of the $\text{B}^0_{\text{s}} \to \text{J}/\psi(\to \mu^+\mu^-)\,\text{K}^+\text{K}^-$ decay are described by sixteen variables. All final-state particles are assumed to be on their mass shell, which gives four relations between the particle energies and three-momenta and only twelve degrees of freedom remain. The final-state momenta can be expressed in terms of the four-momentum of the B^0_{s}, three Euler angles that describe the orientation of the

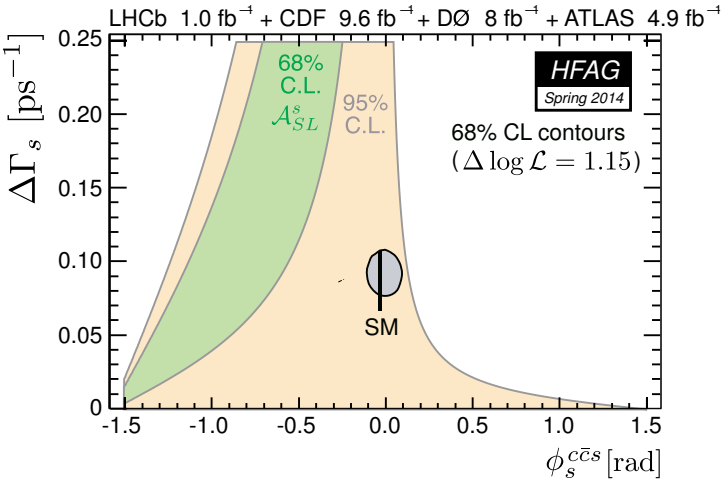


Figure 2.1: *Constraints in the ϕ_s–$\Delta\Gamma_s$ plane from measurements of CP violation in mixing (a_{fs}, here represented as the semi-leptonic observable $\mathcal{A}^s_{SL}$) by HFAG [32], under the assumptions mentioned in the text. Both the 68% (green) and 95% (orange) confidence-level contours are shown. The confidence-level (CL) contour from the combination of direct ϕ_s (here represented as $\phi_s^{c\bar{c}s}$) and $\Delta\Gamma_s$ measurements is shown as the grey area (see also Figure 1.8 on page 20). The Standard Model prediction is represented by the black bar.*

J/ψ and K^+K^- momenta in the rest frame of the B_s^0, the invariant masses of the J/ψ and the K^+K^- system, and three decay angles that describe the orientations of the final state momenta relative to the J/ψ and K^+K^- momenta.

As a result of the fact that the B_s^0 is a spinless particle, the $B_s^0 \to J/\psi K^+K^-$ decay amplitude does not depend on the B_s^0 momentum and the orientation of the J/ψ and K^+K^- momenta (see also Appendix A). The corresponding variables can be integrated over and only five degrees of freedom remain. One of these is eliminated by treating the J/ψ as an on-shell particle, which makes its mass a constant. This also has the advantage that the $B_s^0 \to J/\psi(\to \mu^+\mu^-) K^+K^-$ decay can be described as a $B_s^0 \to J/\psi K^+K^-$ decay followed by a $J/\psi \to \mu^+\mu^-$ decay.

The combined differential decay rate in terms of the four remaining variables can be expressed as [45]

$$\mathrm{d}\Gamma \propto |\mathcal{A}(B_s^0 \to f)|^2 \, \mathrm{d}m^2_{J/\psi K^+} \, \mathrm{d}m^2_{KK} \, \mathrm{d}\cos\theta_\mu \, \mathrm{d}\varphi_h \,, \tag{2.33}$$

where two squared invariant masses and two angles are used instead of one invariant mass and three decay angles. The variables $m_{J/\psi K^+}$ and m_{KK} are the invariant masses of the $J/\psi K^+$ and K^+K^- systems, respectively. The angles θ_μ and φ_h specify the direction of the muons in the rest frame of the J/ψ with a spherical coordinate system.

A change of variables to the K^+K^- mass and three decay angles requires the introduction of an additional variable, θ_K. The three angles θ_K, θ_μ, and φ_h form the set of so-called *helicity angles*, named after the formalism that is used to describe the angular dependence of the decay (see Section 2.4 and Appendix A). These decay angles specify the directions of the four final state particles, given the B_s^0, K^+K^-, and J/ψ invariant masses.

The definition of the helicity angles is shown in Figure 2.2. Starting from the rest frame of the B_s^0, the *helicity axis* is defined by the directions of the J/ψ and K^+K^- momenta. The two polar angles θ_K and θ_μ are defined in the K^+K^- and J/ψ rest frames, respectively, boosting along this axis. θ_K is the angle between the direction of the K^+ and the helicity axis and θ_μ the angle between the direction of the μ^+ and the helicity axis. The azimuthal angle φ_h specifies the relative orientation of the K^+K^- and $\mu^+\mu^-$ decay planes. It is defined as the angle between the "K^+ side" of the K^+K^- plane and the "μ^- side" of the $\mu^+\mu^-$ plane, using a right-handed rotation around the helicity axis that is positive in the direction of the J/ψ momentum. Notice that θ_μ and φ_h form a spherical coordinate system in the J/ψ rest frame, as was required for Equation 2.33.

The change of variables from $m^2_{J/\psi K^+}$ and m^2_{KK} to m_{KK} and $\cos\theta_K$ introduces a Jacobian determinant in Equation 2.33. Given that m^2_{KK} does not depend on $\cos\theta_K$ when expressed in terms of the new variables, this determinant reads $\frac{\partial m^2_{KK}}{\partial m_{KK}} \frac{\partial m^2_{J/\psi K^+}}{\partial \cos\theta_K}$. Exploiting kinematic relations between the final-state momenta an expression for $m^2_{J/\psi K^+}$ can be derived:

$$m^2_{J/\psi K^+} = \tfrac{1}{2}\,(m^2_{B_s^0} + m^2_{J/\psi} + 2\,m^2_{K^+} - m^2_{KK}) + 2\,\frac{m_{B_s^0}}{m_{KK}}\,|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|\,\cos\theta_K\,, \quad (2.34)$$

where $m_{B_s^0}$ is the B_s^0 mass, $\boldsymbol{p}_{J/\psi}$ the J/ψ three-momentum in the rest frame of the B_s^0, and $\boldsymbol{p}_{K^+}$ the kaon three-momentum in the K^+K^- rest frame. With the derivative of this expression with respect to $\cos\theta_K$ the determinant is given by $4\,m_{B_s^0}\,|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|$. Including the determinant, ignoring constant factors, the differential decay rate can now be expressed as

$$\mathrm{d}\Gamma \propto |\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|\,|\mathcal{A}(B_s^0 \to f)|^2\,\mathrm{d}m_{KK}\,\mathrm{d}\cos\theta_K\,\mathrm{d}\cos\theta_\mu\,\mathrm{d}\varphi_h\,. \quad (2.35)$$

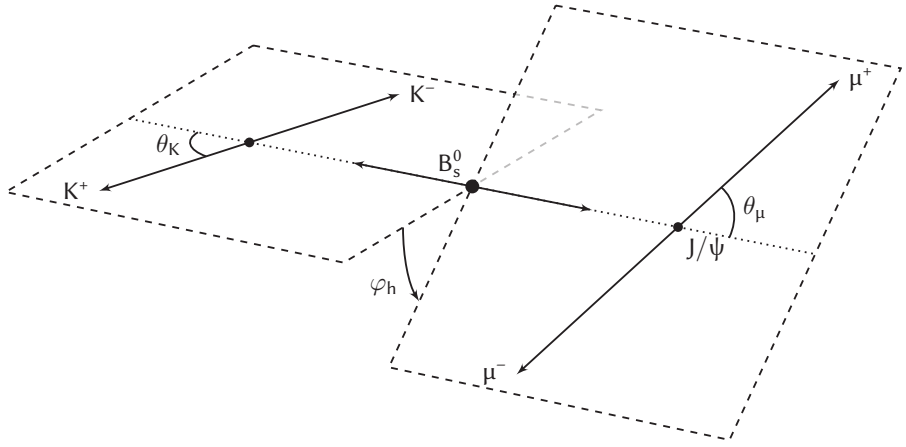


Figure 2.2: *Decay angles in the helicity formalism (figure from [30]).*

The factors $|\boldsymbol{p}_{J/\psi}|$ and $|\boldsymbol{p}_{K^+}|$ depend on the K^+K^- invariant mass, but not on the decay angles and can be absorbed in the m_{KK}-dependent part of $|\mathcal{A}(B_s^0 \to f)|^2$ (see Section 2.5).

The squared amplitude for the $B_s^0 \to J/\psi\, K^+K^-$ process can be obtained by working out expressions for $|\mathcal{A}_f|^2$, $|\overline{\mathcal{A}}_f|^2$ and $\mathcal{A}_f^*\, \overline{\mathcal{A}}_f$ and combining these with the decay time dependence of Equation 2.21. The resulting expression is ordered by the four time-dependent functions, $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$, $\sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$, $\cos(\Delta m_s\, t)$, and $\sin(\Delta m_s\, t)$. Each of the four function coefficients depends on the kinematic variables. See Reference [59] for an example of this approach for $B_q^0 \to J/\psi\, h^+h^-$ processes in general.

However, it is often more convenient to order terms in the expression for the differential decay rate by intermediate resonant or angular-momentum state rather than by decay-time dependence. Each term then has a distinct dependence on kinematic variables. For the CP-violation analysis, a set of intermediate CP eigenstates is used to describe the decay.

The $B_s^0 \to J/\psi\, \phi$ process is the decay of a spin-zero particle into two spin-one particles, which gives three intermediate angular-momentum states. Describing the decay in terms of linear spin-polarization, or *transversity* states, there is one state where the polarization vectors of the J/ψ and the ϕ point along the helicity axis and there are two states where the vectors are transverse with respect to the helicity axis. The former polarization is called *longitudinal* and is indicated with "0". The two transverse polarizations are *parallel*

("$\|$") and *perpendicular* ("$\perp$"), indicating the relative orientations of the two vectors in the transverse plane.

The behaviour of the $J/\psi\,\phi$ system under a CP transformation depends on its orbital angular-momentum configuration. Since both the J/ψ and the $\phi(1020)$ particles are odd under a CP transformation, the CP of the system is equal to the parity of the orbital angular-momentum state. For the longitudinal and parallel polarizations the state of the $J/\psi\,\phi$ orbital angular momentum is either an S-wave ($L = 0$) or a D-wave ($L = 2$), which are even under a parity transformation. The perpendicular polarization state corresponds to a P-wave ($L = 1$), which is odd under parity.

A fourth intermediate state is formed by the K^+K^- S-wave, indicated with "S". For this state the K^+K^- system has no angular momentum. Because the decaying B_s^0 meson is spinless, the orbital angular momentum of the $J/\psi\,K^+K^-$ system needs to be in a P-wave configuration, which compensates for the spin of the J/ψ. As a result, the K^+K^- S-wave state is odd under CP.

To derive an expression for the squared amplitude in terms of intermediate states it is necessary to go back to the expression for the $B_s^0 \to f$ amplitude in Equation 2.18 and expand the decay amplitudes:

$$\mathcal{A}(B_s^0 \to f) \propto g_+\,\mathcal{A}_f + \frac{q}{p}\,g_-\,\overline{\mathcal{A}}_f = g_+ \sum_i \mathcal{A}_i + \frac{q}{p}\,g_- \sum_j \overline{\mathcal{A}}_j \;, \tag{2.36}$$

where $\mathcal{A}_i$ and $\overline{\mathcal{A}}_j$ are the amplitudes for the $|B_s^0\rangle$ and $|\overline{B}_s^0\rangle$ decays through intermediate states i and j, respectively. The dependence of the $|B_s^0\rangle$ decay amplitude on final state kinematics is given by

$$\mathcal{A}_i = A_i\;\mathcal{H}_i(\Omega)\;\mathcal{R}_i(m_{KK}) \;, \tag{2.37}$$

where A_i is the complex-valued coefficient for amplitude $\mathcal{A}_i$ (also termed *transversity amplitude*), $\mathcal{H}_i$ is the amplitude's dependence on decay angles (denoted by Ω), and $\mathcal{R}_i$ is a model of the dependence on the K^+K^- invariant mass.

The $|\overline{B}_s^0\rangle$ decay proceeds through the same intermediate states as the $|B_s^0\rangle$ decay and the respective amplitudes have the same kinematic dependence. Therefore, the $|\overline{B}_s^0\rangle$ amplitude is obtained by replacing the coefficient A_i with $\overline{A}_i$. The total $B_s^0 \to f$ amplitude can now be expressed as

$$\mathcal{A}(B_s^0 \to f) \propto \sum_i \left(g_+ + \lambda_s^i\,g_-\right)\mathcal{A}_i = \sum_i \mathcal{T}_i\,\mathcal{A}_i \;, \tag{2.38}$$

where $\mathcal{T}_i \equiv g_+ + \lambda_{\mathrm{s}}^i\, g_-$ represents the decay-time dependence and the parameter λ_{s}^i is defined in accordance with λ_{f} (Equation 2.22a):

$$\lambda_{\mathrm{s}}^i \equiv \frac{q}{p}\frac{\overline{A}_i}{A_i} \tag{2.39}$$

The amplitude for the $\overline{\mathrm{B}}_{\mathrm{s}}^0 \to \mathrm{f}$ decay is obtained by interchanging g_+ and g_- and multiplying by a factor $\frac{p}{q}$ (see Equation 2.18). Notice that this only affects the decay-time part of Equation 2.38, i.e. the factor $\mathcal{T}_i$.

For intermediate CP eigenstates, the parameter λ_{s}^i can be expressed as

$$\lambda_{\mathrm{s}}^i = \eta_i\, |\lambda_{\mathrm{s}}^i|\, e^{-i\,\phi_{\mathrm{s}}^i}\,, \tag{2.40}$$

where $\eta_i = \pm 1$ is the CP eigenvalue of the state, $|\lambda_{\mathrm{s}}^i| = \left|\frac{q}{p}\right| \left|\frac{\overline{A}_i}{A_i}\right|$, and the CP-violating phase ϕ_{s}^i is given by $\phi_{\mathrm{s}}^i \equiv -\arg\left(\frac{1}{\eta_i}\lambda_{\mathrm{s}}^i\right)$. If CP symmetry is violated by the same amount for all intermediate states, the combination $\frac{1}{\eta_i}\lambda_{\mathrm{s}}^i$ does not depend on the state and $\phi_{\mathrm{s}}^i \to \phi_{\mathrm{s}}$.

Squaring the magnitude of $\mathcal{A}(\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{f})$ gives an expression that contains products of the terms in Equation 2.38 of the form $A_i^* A_j\, \mathcal{T}_i^* \mathcal{T}_j\, \mathcal{H}_i^* \mathcal{H}_j\, \mathcal{R}_i^* \mathcal{R}_j$. Combining terms with indices ij and ji, which are each others complex conjugates, gives

$$|\mathcal{A}(\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{f})|^2 \propto \sum_i |A_i\, \mathcal{T}_i\, \mathcal{H}_i\, \mathcal{R}_i|^2 + \sum_{i \neq j} \Re(A_i^* A_j\, \mathcal{T}_i^* \mathcal{T}_j\, \mathcal{H}_i^* \mathcal{H}_j\, \mathcal{R}_i^* \mathcal{R}_j)\,, \tag{2.41}$$

where the indices i and j run over the intermediate states 0, $\parallel$, $\perp$, and S. The first sum in this expression contains the *diagonal terms* with the squared magnitude of the contribution of each state and the second sum contains the *interference terms* for the different combinations of two states. The dependence on decay time, decay angles, and invariant $\mathrm{K}^+\mathrm{K}^-$ mass of this expression are discussed in the following sections.

2.3 Decay-Time Distribution

The distribution in decay time follows from the products $\mathcal{T}_i^* \mathcal{T}_j$ (see Equation 2.38). With $|g_\pm|^2$ and $g_+^*\, g_-$ from Equation 2.20, these products are given by

$$\mathcal{T}_i^* \mathcal{T}_j = \Big[|g_+|^2 + \lambda_{\mathrm{s}}^{i\,*}\, \lambda_{\mathrm{s}}^j\, |g_-|^2 + \lambda_{\mathrm{s}}^{i\,*}\, (g_+^*\, g_-)^* + \lambda_{\mathrm{s}}^j\, g_+^*\, g_- \Big]$$

$$= \tfrac{1}{2}\, e^{-\Gamma_s t} \Big[(1 + \lambda_s^{i*}\lambda_s^j)\, \cosh\!\big(\tfrac{1}{2}\Delta\Gamma_s\, t\big) + (1 - \lambda_s^{i*}\lambda_s^j)\, \cos(\Delta m_s\, t) - (\lambda_s^{i*} + \lambda_s^j)\, \sinh\!\big(\tfrac{1}{2}\Delta\Gamma_s\, t\big) - i(\lambda_s^{i*} - \lambda_s^j)\, \sin(\Delta m_s\, t) \Big] \,. \quad (2.42)$$

Defining

$$\begin{aligned} C_{ij}^{\pm} &\equiv \frac{1 \pm \lambda_s^{i*}\lambda_s^j}{\sqrt{(1 + |\lambda_s^i|^2)(1 + |\lambda_s^j|^2)}} \\ D_{ij} &\equiv \frac{-(\lambda_s^{i*} + \lambda_s^j)}{\sqrt{(1 + |\lambda_s^i|^2)(1 + |\lambda_s^j|^2)}} \\ S_{ij} &\equiv \frac{+i\,(\lambda_s^{i*} - \lambda_s^j)}{\sqrt{(1 + |\lambda_s^i|^2)(1 + |\lambda_s^j|^2)}} \end{aligned} \quad (2.43)$$

in accordance with Equation 2.22b and a "CP-average" decay amplitude

$$\begin{aligned} A_i^{\rm CP} &\equiv \tfrac{1}{\sqrt{2}}\, A_i \sqrt{1 + |\lambda_s^i|^2} \\ \big|A_i^{\rm CP}\big|^2 &= \tfrac{1}{2}\, |A_i|^2 \left(1 + |\lambda_s^i|^2\right) = \tfrac{1}{2} \left(|A_i|^2 + \big|\tfrac{q}{p}\big|^2\, |\bar{A}_i|^2 \right) \,, \end{aligned} \quad (2.44)$$

the combination $A_i^* A_j\, \mathcal{T}_i^* \mathcal{T}_j$ can be expressed as

$$\begin{aligned} A_i^* A_j\, \mathcal{T}_i^* \mathcal{T}_j = A_i^{\rm CP*} A_j^{\rm CP}\, e^{-\Gamma_s t} \\ \times \Big[C_{ij}^{+}\, \cosh\!\big(\tfrac{1}{2}\Delta\Gamma_s\, t\big) + C_{ij}^{-}\, \cos(\Delta m_s\, t) \\ + D_{ij}\, \sinh\!\big(\tfrac{1}{2}\Delta\Gamma_s\, t\big) - S_{ij}\, \sin(\Delta m_s\, t) \Big] \,. \end{aligned} \quad (2.45)$$

In general, the coefficients $C_{ij}^{\pm}$, D_{ij}, and S_{ij} are complex numbers and hence the terms $A_i^* A_j\, \mathcal{T}_i^* \mathcal{T}_j$ have both real and imaginary parts. Notice that if $i = j$, the coefficients are real:

$$\begin{aligned} C_{ii}^{+} &= 1 & C_{ii}^{-} &= \frac{1 - |\lambda_s^i|^2}{1 + |\lambda_s^i|^2} \\ D_{ii} &= -\frac{2\,\Re(\lambda_s^i)}{1 + |\lambda_s^i|^2} & S_{ii} &= \frac{2\,\Im(\lambda_s^i)}{1 + |\lambda_s^i|^2} \,, \end{aligned} \quad (2.46)$$

as required by the relation $A_i^* A_i\, \mathcal{T}_i^* \mathcal{T}_i \;= |A_i\, \mathcal{T}_i|^2$.

The equivalent expression for the $\overline{\mathrm{B}}{}^0_{\mathrm{s}} \to \mathrm{f}$ decay is obtained by changing the signs of the $\cos(\Delta m_{\mathrm{s}}\, t)$ and $\sin(\Delta m_{\mathrm{s}}\, t)$ terms and multiplying by a factor $\left|\frac{p}{q}\right|^2$ (see Equations 2.18 and 2.20). Using the definition of C_{mix} from Equation 2.25, the time dependence of the $\mathrm{B}^0_{\mathrm{s}} \to \mathrm{f}$ and $\overline{\mathrm{B}}{}^0_{\mathrm{s}} \to \mathrm{f}$ decays is given by

$$\begin{aligned} A_i^* A_j\, \mathcal{T}^*_{i,q_{\mathrm{f}}} \mathcal{T}_{j,q_{\mathrm{f}}} = A_i^{\mathrm{CP}*} A_j^{\mathrm{CP}}\, \frac{1 - q_{\mathrm{f}}\, C_{\mathrm{mix}}}{1 - C_{\mathrm{mix}}}\, e^{-\Gamma_{\mathrm{s}}\, t} \\ \times \Big[C_{ij}^{+}\, \cosh\!\big(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\big) + q_{\mathrm{f}}\, C_{ij}^{-}\, \cos(\Delta m_{\mathrm{s}}\, t) \\ + D_{ij}\, \sinh\!\big(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\big) - q_{\mathrm{f}}\, S_{ij}\, \sin(\Delta m_{\mathrm{s}}\, t) \Big]\,, \end{aligned} \tag{2.47}$$

where the expressions have become dependent on the flavour of the decaying beauty meson, q_{f} (see Equation 2.28). In Sections 2.4, 2.5, and 2.6 only the expression for the differential decay rate of the $\mathrm{B}^0_{\mathrm{s}}$ decay will be discussed, but the dependence on q_{f} can be reintroduced at any point by multiplying the $\cos(\Delta m_{\mathrm{s}}\, t)$ and $\sin(\Delta m_{\mathrm{s}}\, t)$ terms by q_{f} and the total rate by $\frac{1-q_{\mathrm{f}}\, C_{\mathrm{mix}}}{1-C_{\mathrm{mix}}}$.

2.3.1 Common CP Violation

In case CP symmetry is violated equally for all intermediate states, the common parameter $\lambda_{\mathrm{s}} \equiv \frac{1}{\eta_i}\lambda^i_{\mathrm{s}}$ can be defined, with $\phi_{\mathrm{s}} = -\arg(\lambda_{\mathrm{s}})$. The corresponding real-valued parameters in the decay-time distribution are then given by

$$C_{\mathrm{s}} \equiv \frac{1 - |\lambda_{\mathrm{s}}|^2}{1 + |\lambda_{\mathrm{s}}|^2} \qquad D_{\mathrm{s}} \equiv -\frac{2\,\Re(\lambda_{\mathrm{s}})}{1 + |\lambda_{\mathrm{s}}|^2} \qquad S_{\mathrm{s}} \equiv \frac{2\,\Im(\lambda_{\mathrm{s}})}{1 + |\lambda_{\mathrm{s}}|^2}\,. \tag{2.48}$$

For states i and j that are both CP even or both CP odd ($\eta_i = \eta_j$), the coefficients from Equations 2.43, 2.45, and 2.47 reduce to

$$C_{ij}^{+} \to 1 \qquad C_{ij}^{-} \to C_{\mathrm{s}} \qquad D_{ij} \to \eta_i\, D_{\mathrm{s}} \qquad S_{ij} \to \eta_i\, S_{\mathrm{s}}\,, \tag{2.49a}$$

Similarly, for $\eta_i = -\eta_j$, the coefficients reduce to

$$C_{ij}^{+} \to C_{\mathrm{s}} \qquad C_{ij}^{-} \to 1 \qquad D_{ij} \to i\,\eta_i\, S_{\mathrm{s}} \qquad S_{ij} \to -i\,\eta_i\, D_{\mathrm{s}}\,. \tag{2.49b}$$

Notice that in this special case the coefficients for $\eta_i = \eta_j$ are all real, the coefficients $C_{ij}^{\pm}$ for $\eta_i = -\eta_j$ are real, and the coefficients D_{ij} and S_{ij} for $\eta_i = -\eta_j$ are imaginary. As a result, simultaneously flipping the signs of D_{s} and $\sinh\!\big(\frac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\big)$, which is accomplished by the operations $\phi_{\mathrm{s}} \longrightarrow \pi - \phi_{\mathrm{s}}$ and $\Delta\Gamma_{\mathrm{s}} \longrightarrow -\Delta\Gamma_{\mathrm{s}}$, gives the complex conjugate of the product $\mathcal{T}_i^* \mathcal{T}_j$. This turns out to give a discrete ambiguity in the parameter values, as will be explained in Section 2.6.2.

2.3.2 Alternative Parameterization

The coefficients of the time-dependent functions in Equation 2.47 are functions of λ_{s}^{i} and hence contain both the effects of CP violation in mixing and CP violation in decay. These two effects can be separated with an alternative parameterization, which has been implemented in the software that was used for data analysis in the measurement presented in this thesis.

Starting again from Equation 2.42, an alternative expression for the product $\mathcal{T}_{i,q_{\mathrm{f}}}^{*}\,\mathcal{T}_{j,q_{\mathrm{f}}}$ is derived by splitting the λ_{s}^{i} parameter into a $\frac{q}{p}$ factor and an amplitude factor. Defining

$$\begin{aligned} P_{ij}^{\pm} &\equiv \frac{1}{2}\left[1 \pm \left(\frac{\bar{A}_i}{A_i}\right)^{*}\,\frac{\bar{A}_j}{A_j}\right] \\ P_{ij}^{\Re} \equiv \frac{1}{2}\left[\left(\frac{\bar{A}_i}{A_i}\right)^{*} + \frac{\bar{A}_j}{A_j}\right] &\qquad P_{ij}^{\Im} \equiv \frac{i}{2}\left[\left(\frac{\bar{A}_i}{A_i}\right)^{*} - \frac{\bar{A}_j}{A_j}\right] \end{aligned} \tag{2.50}$$

and

$$C_{\mathrm{mix}} \equiv \frac{1-\left|\frac{q}{p}\right|^2}{1+\left|\frac{q}{p}\right|^2} \qquad D_{\mathrm{mix}} \equiv -\frac{2\,\Re\left(\frac{q}{p}\right)}{1+\left|\frac{q}{p}\right|^2} \qquad S_{\mathrm{mix}} \equiv \frac{2\,\Im\left(\frac{q}{p}\right)}{1+\left|\frac{q}{p}\right|^2}\,, \tag{2.51}$$

the product $A_i^{*}A_j\,\mathcal{T}_{i,q_{\mathrm{f}}}^{*}\mathcal{T}_{j,q_{\mathrm{f}}}$ can be expressed as

$$\begin{aligned} A_i^{*}A_j\,\mathcal{T}_{i,q_{\mathrm{f}}}^{*}\mathcal{T}_{j,q_{\mathrm{f}}} = A_i^{*}A_j\,\frac{1-q_{\mathrm{f}}\,C_{\mathrm{mix}}}{1-C_{\mathrm{mix}}^2}\,e^{-\Gamma_{\mathrm{s}}\,t} \\ \times\Big[\left(P_{ij}^{+}+P_{ij}^{-}\,C_{\mathrm{mix}}\right)\,\cosh\!\left(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\,t\right) \\ +\,q_{\mathrm{f}}\left(P_{ij}^{-}+P_{ij}^{+}\,C_{\mathrm{mix}}\right)\,\cos(\Delta m_{\mathrm{s}}\,t) \\ +\left(P_{ij}^{\Re}\,D_{\mathrm{mix}}+P_{ij}^{\Im}\,S_{\mathrm{mix}}\right)\,\sinh\!\left(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\,t\right) \\ +\,q_{\mathrm{f}}\left(P_{ij}^{\Im}\,D_{\mathrm{mix}}-P_{ij}^{\Re}\,S_{\mathrm{mix}}\right)\,\sin(\Delta m_{\mathrm{s}}\,t)\Big]\,. \end{aligned} \tag{2.52}$$

Notice that the amplitudes A_i and A_j occur in this expression and not the CP-average amplitudes A_i^{CP} and A_j^{CP}.

The fact that the phases of $\frac{q}{p}$ and $\frac{\bar{A}_i}{A_i}$ cannot be observed independently is reflected by the observation that only the combinations $P_{ij}^{\Re}\,D_{\mathrm{mix}} + P_{ij}^{\Im}\,S_{\mathrm{mix}}$ and $P_{ij}^{\Im}\,D_{\mathrm{mix}} - P_{ij}^{\Re}\,S_{\mathrm{mix}}$ are observable and not the $P_{ij}^{\Re}$, P_{ij}^{-}, D_{mix}, and S_{mix}

parameters individually. A phase convention can be chosen by, for example, setting $\arg(\frac{q}{p}) \equiv 0$, which makes the phase of $\frac{\bar{A}_i}{A_i}$ equal to the phase of λ_s^i. Another choice is $\arg(\frac{\bar{A}_0}{A_0}) \equiv 0$, which results in the equations $\arg(\frac{q}{p}) = \arg(\lambda_s^0)$ and $\arg(\frac{\bar{A}_i}{A_i}) = \arg(\lambda_s^i) - \arg(\lambda_s^0)$ for the remaining parallel, perpendicular, and S-wave amplitudes. This convention directly gives a measurement of phase differences.

2.4 Decay-Angle Distributions

Expanding Equation 2.41 into terms with the real and imaginary parts of the angular dependence gives

$$\begin{aligned} |\mathcal{A}(B_s^0 \to f)|^2 \propto & \sum_i |c_i|^2 \, |\mathcal{H}_i|^2 \\ & + \sum_{i \neq j} \Re(c_i^* \, c_j) \, \Re(\mathcal{H}_i^* \, \mathcal{H}_j) - \Im(c_i^* \, c_j) \, \Im(\mathcal{H}_i^* \, \mathcal{H}_j) \, , \end{aligned} \tag{2.53}$$

where the amplitude coefficients c_i are defined as $c_i \equiv A_i \, \mathcal{T}_i \, \mathcal{R}_i$. The angular dependence of this squared amplitude is described by the functions $|\mathcal{H}_i|^2$, $2\,\Re(\mathcal{H}_i^* \, \mathcal{H}_j)$, and $-2\,\Im(\mathcal{H}_i^* \, \mathcal{H}_j)$. The factors two in the functions of the interference terms originate from adding the terms with indices ij and ji, which give identical contributions.

Expressions for the products $\mathcal{H}_i^* \, \mathcal{H}_j$ are derived in Appendix A using the *helicity formalism* [60, 61]. In this description the angular dependence for an intermediate state in the $B_s^0 \to J/\psi \, K^+K^-$ decay with definite particle helicities is given by the product of two Wigner D-matrices (see Equation A.5).

Squaring the magnitude of the sum over D-matrices for the different helicity states yields angular functions for each of the states and for their interferences (Equations A.9, A.10, A.14, and A.15). Because helicity states are not CP eigenstates, the functions are combined into functions in the transversity basis by substituting their coefficients (*helicity amplitudes*) by combinations of the transversity amplitudes, A_i (Equation A.16). Finally, this gives the expressions for the functions $|\mathcal{H}_i|^2$, $2\,\Re(\mathcal{H}_i^* \, \mathcal{H}_j)$, and $-2\,\Im(\mathcal{H}_i^* \, \mathcal{H}_j)$.

The resulting angular functions are shown for each combination of intermediate states i and j in Tables A.1–A.4. A summary is given in Table 2.1. While in the appendix the coefficients of the amplitudes are given by only

the helicity/transversity amplitudes, the coefficients c_i are used in Table 2.1 to also introduce the dependence on decay time ($\mathcal{T}_i$) and K^+K^- mass ($\mathcal{R}_i$).

Table 2.1: *Angular functions for the $B_s^0 \to J/\psi K^+K^-$ decay in helicity angles. The coefficients, which depend on the decay time and the K^+K^- mass, are given by $c_i \equiv A_i\,\mathcal{T}_i\,\mathcal{R}_i$.*

coefficient	$f(\Omega) \times \frac{32\pi}{9}$
$\lvert c_0\rvert^2$	$2\,\cos^2\theta_K\,\sin^2\theta_\mu$
$\lvert c_\parallel\rvert^2$	$\sin^2\theta_K\,(1-\sin^2\theta_\mu\,\cos^2\varphi_h)$
$\lvert c_\perp\rvert^2$	$\sin^2\theta_K\,(1-\sin^2\theta_\mu\,\sin^2\varphi_h)$
$\Re(c_0^* c_\parallel)$	$+\frac{1}{\sqrt{2}}\,\sin 2\theta_K\,\sin 2\theta_\mu\,\cos\varphi_h$
$\Im(c_0^* c_\perp)$	$-\frac{1}{\sqrt{2}}\,\sin 2\theta_K\,\sin 2\theta_\mu\,\sin\varphi_h$
$\Im(c_\parallel^* c_\perp)$	$+\sin^2\theta_K\,\sin^2\theta_\mu\,\sin 2\varphi_h$
$\lvert c_S\rvert^2$	$\frac{2}{3}\,\sin^2\theta_\mu$
$\Re(c_0^* c_S)$	$\frac{4}{3}\sqrt{3}\,\cos\theta_K\,\sin^2\theta_\mu$
$\Re(c_\parallel^* c_S)$	$\frac{1}{3}\sqrt{6}\,\sin\theta_K\,\sin 2\theta_\mu\,\cos\varphi_h$
$\Im(c_\perp^* c_S)$	$\frac{1}{3}\sqrt{6}\,\sin\theta_K\,\sin 2\theta_\mu\,\sin\varphi_h$

After adding the results for the two possible helicity configurations of the muons in the J/ψ decay, only ten terms in Equation 2.53 give non-zero contributions. As a result, each combination of states i and j only appears once in Table 2.1. For the interference terms either the real or the imaginary part of the product $c_i^* c_j$ appears.

As indicated in Equation 2.35, the functions $\cos\theta_K$ and $\cos\theta_\mu$ are used as variables in the decay model rather than the corresponding angles θ_K and θ_μ. These angles are polar angles in a spherical coordinate system, which are defined between 0 and π. The corresponding cosines, with ranges $[-1,+1]$, completely specify the values of these two angles and are the natural variables to use. The angle φ_h is an azimuthal angle, defined between $-\pi$ and $+\pi$.

The angular functions in Table 2.1 are normalized such that the integrals over all three angular variables are equal to one for the diagonal terms ($|c_0|^2$, $|c_\parallel|^2$, $|c_\perp|^2$, and $|c_S|^2$). All interference terms vanish when integrated over all three variables. Also the one-angle distributions, given in Table 2.2, are dominated by the contributions from the diagonal terms. Interference terms that survive the integration over two angular variables are the $\Re(c_0^* c_S)$ term

in $\cos\theta_K$ and the $\Im(c_\parallel^* c_\perp)$ term in φ_h, but these contributions almost vanish when integrated over decay time (see Table 2.3 in Section 2.6.1).

Table 2.2: *Angular functions for the $B_s^0 \to J/\psi K^+K^-$ decay integrated over two of the three angular variables. The coefficients, which depend on the decay time and the K^+K^- mass, are given by $c_i \equiv A_i\, \mathcal{T}_i\, \mathcal{R}_i$.*

coefficient	$f(\cos\theta_K)$	$f(\cos\theta_\mu)$	$f(\varphi_h)$
$\lvert c_0\rvert^2$	$\frac{3}{2}\,\cos^2\theta_K$	$\frac{3}{4}\,(1-\cos^2\theta_\mu)$	$\frac{1}{2\pi}$
$\lvert c_\parallel\rvert^2$	$\frac{3}{4}\,(1-\cos^2\theta_K)$	$\frac{3}{8}\,(1+\cos^2\theta_\mu)$	$\frac{1}{4\pi}\,(2-\cos 2\varphi_h)$
$\lvert c_\perp\rvert^2$	$\frac{3}{4}\,(1-\cos^2\theta_K)$	$\frac{3}{8}\,(1+\cos^2\theta_\mu)$	$\frac{1}{4\pi}\,(2+\cos 2\varphi_h)$
$\lvert c_S\rvert^2$	$\frac{1}{2}$	$\frac{3}{4}\,(1-\cos^2\theta_\mu)$	$\frac{1}{2\pi}$
$\Im(c_\parallel^* c_\perp)$	–	–	$\frac{1}{2\pi}\,\sin 2\varphi_h$
$\Re(c_0^* c_S)$	$\sqrt{3}\,\cos\theta_K$	–	–

2.5 Invariant K^+K^--Mass Distribution

Besides the dependence on decay-time via $\mathcal{T}_i$, the angular coefficients in Equation 2.53 also depend on K^+K^- mass via $\mathcal{R}_i$. This dependence is described by a phenomenological model. To simplify the analysis of the time and angular distributions, the model for the $B_s^0 \to J/\psi K^+K^-$ decay is integrated over the K^+K^- mass. As will be discussed in Sections 2.6.2 and 3.5, the model is split into mass intervals, for which separate integrals are calculated.

As discussed in Section 1.3.3, the K^+K^- pair is required to have an invariant mass between 990 and 1050 MeV/c^2. The plot of the K^+K^--mass spectrum in Figure 1.9 shows that the contribution of the $\phi(1020)$ dominates in this region, but that there are also $f_0(980)$ and non-resonant contributions.

The K^+K^--mass integral gives a factor for each term in Equation 2.41, which is, in principle, different for each combination of intermediate states. However, because the dependence on K^+K^- mass is approximately equal for the three $B_s^0 \to J/\psi\,\phi$ states, this contribution can be factored out as an overall normalization and only relative factors between the K^+K^- S-wave and the $B_s^0 \to J/\psi\,\phi$ contributions remain (see also [3, 62]).

Before integrating the expressions for $\mathcal{R}_i^*\,\mathcal{R}_j$ over m_{KK} they are multiplied by the factor $|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|$ from Equation 2.35. The dependence of this

factor on m_{KK} can be derived from the kinematic relations in the "two-body decays" of the B_s^0 and the K^+K^- pair. The expressions for the magnitudes of the J/ψ momentum in the B_s^0 rest frame and the K^+ momentum in the K^+K^- rest frame depend only on the B_s^0, J/ψ, K^+K^-, and K^+ masses. Working out the kinematic relations, the factor $|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|$ is given by

$$|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}| = \frac{\sqrt{\lambda(m_{B_s^0}, m_{J/\psi}, m_{KK})\,\lambda(m_{KK}, m_{K^+}, m_{K^+})}}{4\,m_{B_s^0}\,m_{KK}}\,, \tag{2.54}$$

with

$$\lambda(M, m_1, m_2) \equiv M^4 + m_1^4 + m_2^4 - 2\,M^2 m_1^2 - 2\,M^2 m_2^2 - 2\,m_1^2 m_2^2\,. \tag{2.55}$$

The functions $\mathcal{R}_i(m_{KK})$ can be normalized by absorbing a constant factor into the corresponding transversity amplitude A_i, such that the integrals for diagonal terms are given by $\int \mathrm{d}m_{KK}\,|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|\,|\mathcal{R}_i|^2 \equiv 1$. This procedure moves the issue of overall normalization to the transversity amplitudes and leaves only non-trivial m_{KK} integrals in the interference terms. These integrals are given by (the real and imaginary parts of)

$$\int_{m_{KK}^-}^{m_{KK}^+} \mathrm{d}m_{KK}\,|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|\,\mathcal{R}_i^*\,\mathcal{R}_j \equiv K_{ij}\,e^{-i\kappa_{ij}}\,, \tag{2.56}$$

where $m_{KK}^- = 990\,\mathrm{MeV}/c^2$, $m_{KK}^+ = 1050\,\mathrm{MeV}/c^2$. The real-valued parameters K_{ij} and κ_{ij} form the K^+K^--mass factor for interference term ij. By construction, the conditions $K_{ii} = 1$ and $\kappa_{ii} = 0$ apply. With the Cauchy–Schwarz inequality, $|\int \mathrm{d}x\,A^*B\,|^2 \le \int \mathrm{d}x\,|A|^2 \cdot \int \mathrm{d}x\,|B|^2$, it follows that $0 \le K_{ij} \le 1$.

Because only the magnitude of the $B_s^0 \to J/\psi\,K^+K^-$ decay amplitude is observable, the overall phase of the transversity amplitudes is arbitrary and only phase differences between the individual amplitudes can be observed. Integrating over m_{KK}, the phases κ_{ij} can be absorbed into these phase differences:

$$\begin{aligned}\int_{m_{KK}^-}^{m_{KK}^+} \mathrm{d}m_{KK}\,|\boldsymbol{p}_{J/\psi}|\,|\boldsymbol{p}_{K^+}|\,A_i^*A_j\,\mathcal{R}_i^*\,\mathcal{R}_j &= A_i^*A_j\,K_{ij}\,e^{-i\kappa_{ij}} \\ &= K_{ij}\,|A_i|\,|A_j|\,e^{i[\arg(A_j)-\arg(A_i)-\kappa_{ij}]}\,.\end{aligned} \tag{2.57}$$

The interference terms become proportional to the *coupling factors* between states i and j, K_{ij}, which are treated separately from the transversity

amplitudes. Although these factors can in principle be determined from the angular distributions in the $B_s^0 \to J/\psi K^+K^-$ data, their statistical uncertainties would be large. For this reason the K_{ij} values are estimated by assuming phenomenological models for the K^+K^--mass shapes of the different components of the decay and performing the integrals of Equation 2.56.

It is assumed that the K^+K^--mass models only depend on the resonant state of the K^+K^- system and not on the angular momentum of the $J/\psi K^+K^-$ system. That is, the same mass shape is used for all three angular-momentum states of the $B_s^0 \to J/\psi\phi$ decay, which is assumed to be a good approximation. See references [59] and [46] for a discussion of the small dependence on the orbital angular momentum of the $J/\psi K^+K^-$ system.

Because the K^+K^--mass shapes for the three $B_s^0 \to J/\psi\phi$ intermediate states are thus taken to be identical, only one non-trivial K_{ij} factor remains. This factor is indicated by "K_{PS}", where "P" stands for "P-wave" and "S" for "S-wave", indicating the K^+K^- angular-momentum states for the $B_s^0 \to J/\psi\phi$ and K^+K^- S-wave contributions, respectively.

Similarly, the only remaining contribution to the phases of the transversity amplitudes is κ_{PS}. The following phase differences are defined:

$$\begin{aligned} \delta_\parallel - \delta_0 &\equiv \arg(A_\parallel) - \arg(A_0) \\ \delta_\perp - \delta_0 &\equiv \arg(A_\perp) - \arg(A_0) \\ \delta_S - \delta_\perp &\equiv \arg(A_S) - \arg(A_\perp) - \kappa_{\mathrm{PS}} \,. \end{aligned} \tag{2.58}$$

The remaining phase differences can be expressed in terms of the above three.

A relativistic Breit-Wigner function (see e.g. reference [45]) is used to model the m_{KK} dependence of the $\phi(1020)$ component of the K^+K^- system. The $f_0(980)$ contribution is modelled by a Flatté function [63]. The contribution from non-resonant K^+K^- pairs, which is small compared to the $\phi(1020)$ and $f_0(980)$ contributions in the 990–1050 MeV/c^2 mass window, is neglected in this measurement. See Section 3.5 for the result of the K_{PS} calculation, where also the experimental m_{KK} resolution is taken into account.

2.6 Decay-Rate Equations

Combining the expressions for the decay-time and K^+K^--mass dependence from Equations 2.45 and 2.57 yields the coefficients in the sum of the angular

dependence in Equation 2.53:

$$\begin{aligned}\int_{m_{\mathrm{KK}}^-}^{m_{\mathrm{KK}}^+} \mathrm{d}m_{\mathrm{KK}}\, |\boldsymbol{p}_{\mathrm{J}/\psi}|\, |\boldsymbol{p}_{\mathrm{K}^+}|\, c_i^* c_j &= \int_{m_{\mathrm{KK}}^-}^{m_{\mathrm{KK}}^+} \mathrm{d}m_{\mathrm{KK}}\, |\boldsymbol{p}_{\mathrm{J}/\psi}|\, |\boldsymbol{p}_{\mathrm{K}^+}|\, A_i^* A_j\, \mathcal{T}_i^*\, \mathcal{T}_j\, \mathcal{R}_i^*\, \mathcal{R}_j \\ &= K_{ij}\, e^{-i\kappa_{ij}}\, A_i^{\mathrm{CP}*} A_j^{\mathrm{CP}}\, e^{-\Gamma_{\mathrm{s}} t} \Big[C_{ij}^+\, \cosh\!\big(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\big) + C_{ij}^-\, \cos(\Delta m_{\mathrm{s}}\, t) \\ &\qquad + D_{ij}\, \sinh\!\big(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\big) - S_{ij}\, \sin(\Delta m_{\mathrm{s}}\, t) \Big] \,. \end{aligned} \tag{2.59}$$

The phase of the integral over the $\mathrm{K}^+\mathrm{K}^-$-mass functions can be absorbed in the phase difference between the amplitudes, as in Equation 2.57:

$$\begin{aligned} e^{-i\kappa_{ij}}\, A_i^{\mathrm{CP}*} A_j^{\mathrm{CP}} &= |A_i^{\mathrm{CP}}|\, |A_j^{\mathrm{CP}}|\, e^{i(\delta_j - \delta_i)} \\ &= |A_i^{\mathrm{CP}}|\, |A_j^{\mathrm{CP}}|\, [\cos(\delta_j - \delta_i) + i\, \sin(\delta_j - \delta_i)] \,, \end{aligned} \tag{2.60}$$

where

$$\delta_j - \delta_i \equiv \arg(A_j^{\mathrm{CP}}) - \arg(A_i^{\mathrm{CP}}) - \kappa_{ij} \,. \tag{2.61}$$

Depending on whether the real or imaginary part of the coefficient is required in Equation 2.53, the real and imaginary parts of the time-dependence of Equation 2.59 are multiplied by either $\cos(\delta_j - \delta_i)$ or $\sin(\delta_j - \delta_i)$. Taking the $\cosh\big(\frac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\big)$ term as an example, the coefficients of the angular distribution are proportional to

$$\Re\big(e^{i(\delta_j - \delta_i)}\, C_{ij}^+\big) = \cos(\delta_j - \delta_i)\, \Re(C_{ij}^+) - \sin(\delta_j - \delta_i)\, \Im(C_{ij}^+) \tag{2.62a}$$

$$\Im\big(e^{i(\delta_j - \delta_i)}\, C_{ij}^+\big) = \sin(\delta_j - \delta_i)\, \Re(C_{ij}^+) + \cos(\delta_j - \delta_i)\, \Im(C_{ij}^+) \,, \tag{2.62b}$$

where the first equation applies to the diagonal terms and the "0$\|$", "0S", and "$\|$S" terms and the second equation to the "0$\perp$", "$\|\perp$", and "$\perp$S" terms.

The full expression for the differential decay rate in decay time and decay angles is obtained by summing the products of the angular functions (Table 2.1) and their time-dependent coefficients (real or imaginary part of Equation 2.59). This expression is used to build the probability density function that models the $\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{J}/\psi\, \mathrm{K}^+\mathrm{K}^-$ decay, as will be discussed in Chapter 3.

2.6.1 Approximate Equations

To indicate where the sensitivity to the different parameters in the decay model comes from, two approximations of the terms in the differential rate

are given in Tables 2.3, 2.4, and 2.5. The first table shows the time-dependent functions in the differential decay rate in the case of no CP violation. Since CP violation in the $B_s^0 \to J/\psi\, K^+K^-$ decay is small, this gives a good indication of how the remaining parameters appear in the decay-rate equations. Without CP violation, the parameters C_s and S_s are equal to zero and D_s is equal to minus one. With Equation 2.49, this gives

$$C_{ij}^{+} \to 1 \quad C_{ij}^{-} \to 0 \quad D_{ij} \to -\eta_i \quad S_{ij} \to 0 \quad (\eta_i = +\eta_j) \tag{2.63a}$$

$$C_{ij}^{+} \to 0 \quad C_{ij}^{-} \to 1 \quad D_{ij} \to 0 \quad S_{ij} \to i\,\eta_i \quad (\eta_i = -\eta_j)\,. \tag{2.63b}$$

Because of the differences between coefficients for states with equal CP eigenvalues ($\eta_i = +\eta_j$) and states with opposite eigenvalues ($\eta_i = -\eta_j$) in Equation 2.49, different coefficients are non-zero for these two cases without CP violation. The coefficients in Equation 2.63a correspond to the diagonal terms ($i = j$) in the decay-rate equations and to interference terms with equal CP eigenvalues. The coefficients in Equation 2.63b correspond to interference of a CP-even and a CP-odd state, which have opposite eigenvalues.

Table 2.3: *Functions of decay time without CP violation.*

ij	$f(t) \times e^{+\Gamma_s t}$
ii	$\lvert A_i^{CP}\rvert^2 \left[\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right) - \eta_i\, \sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)\right]$
$0\lVert$	$\lvert A_0^{CP}\rvert\lvert A_\lVert^{CP}\rvert \cos(\delta_\lVert - \delta_0) \left[\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right) - \sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)\right]$
$0\perp$	$\lvert A_0^{CP}\rvert\lvert A_\perp^{CP}\rvert \left[\sin(\delta_\perp - \delta_0)\, \cos(\Delta m_s\, t) - \cos(\delta_\perp - \delta_0)\, \sin(\Delta m_s\, t)\right]$
$\lVert\perp$	$\lvert A_\lVert^{CP}\rvert\lvert A_\perp^{CP}\rvert \left[\sin(\delta_\perp - \delta_\lVert)\, \cos(\Delta m_s\, t) - \cos(\delta_\perp - \delta_\lVert)\, \sin(\Delta m_s\, t)\right]$
$0S$	$K_{PS}\lvert A_0^{CP}\rvert\lvert A_S^{CP}\rvert \left[\cos(\delta_S - \delta_0)\, \cos(\Delta m_s\, t) + \sin(\delta_S - \delta_0)\, \sin(\Delta m_s\, t)\right]$
$\lVert S$	$K_{PS}\lvert A_\lVert^{CP}\rvert\lvert A_S^{CP}\rvert \left[\cos(\delta_S - \delta_\lVert)\, \cos(\Delta m_s\, t) + \sin(\delta_S - \delta_\lVert)\, \sin(\Delta m_s\, t)\right]$
$\perp S$	$K_{PS}\lvert A_\perp^{CP}\rvert\lvert A_S^{CP}\rvert \sin(\delta_S - \delta_\perp) \left[\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right) + \sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)\right]$

Equation 2.63 and Table 2.3 show that for terms with $\eta_i = \eta_j$ only the $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ functions remain and for $\eta_i = -\eta_j$ terms only the $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ functions. In the experiment, flavour tagging is required to determine the coefficients of the latter, since these terms have opposite sign for B_s^0 and $\overline{B}_s^0$ decays (see Equation 2.47) and cancel in the sum of the two differential decay rates. Because it is not possible to tag all decay candidates (correctly), the statistical uncertainties of the associated pa-

rameters will generally be larger than the uncertainties of parameters in the $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ terms.

Without CP violation, light and heavy eigenstates of the B_s^0–$\overline{B}_s^0$ system coincide with the CP-even and CP-odd states, respectively. The time dependence of the diagonal terms, which form the decay-time distribution integrated over the decay angles, can consequently be expressed as

$$\begin{aligned} &|A_i^{\mathrm{CP}}|^2\, e^{-\Gamma_s t}\left[\cosh\left(\tfrac{1}{2}\Delta\Gamma_s\, t\right) - \eta_i\, \sinh\left(\tfrac{1}{2}\Delta\Gamma_s\, t\right)\right] \\ &\qquad = \begin{cases} |A_i^{\mathrm{CP}}|^2\, e^{-\Gamma_L t} & \text{if } \eta_i = +1 \quad \text{(CP even)} \\ |A_i^{\mathrm{CP}}|^2\, e^{-\Gamma_H t} & \text{if } \eta_i = -1 \quad \text{(CP odd)} \end{cases} . \end{aligned} \tag{2.64}$$

The values of the parameters Γ_s, $\Delta\Gamma_s$, and $|A_i^{\mathrm{CP}}|$ can be estimated relatively precisely from untagged decay candidates, but these estimates are correlated. The mean of the decay-time distribution is controlled by Γ_s, but also by the relative contributions of CP-even and CP-odd states ($|A_i^{\mathrm{CP}}|$), which have different lifetimes. The impact of changing the even and odd contributions depends on the value of $\Delta\Gamma_s$, which controls the difference in lifetime between the two states.

The phases of the transversity amplitudes can only be determined from interference terms. The cosine of $\delta_\parallel - \delta_0$ and the sine of $\delta_S - \delta_\perp$ appear in the untagged distribution, but tagged decay candidates are required to measure the sine of $\delta_\parallel - \delta_0$ and the cosine of $\delta_S - \delta_\perp$ from combinations of the remaining interference terms. As a result, the measurements of $\sin(\delta_\parallel - \delta_0)$ and $\cos(\delta_S - \delta_\perp)$ are less precise than the measurements of $\cos(\delta_\parallel - \delta_0)$ and $\sin(\delta_S - \delta_\perp)$ and an approximate symmetry in the estimates of these parameters arises for $\delta_\parallel - \delta_0 \longrightarrow 2\pi - (\delta_\parallel - \delta_0)$ and $\delta_S - \delta_\perp \longrightarrow \pi - (\delta_S - \delta_\perp)$, for which the values of $\cos(\delta_\parallel - \delta_0)$ and $\sin(\delta_S - \delta_\perp)$ do not change. The parameter $\delta_\perp - \delta_0$ is only determined with $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ terms in the approximation of Table 2.3 and does not show a similar symmetry.

The interference term of the parallel and perpendicular states depends on both the $\delta_\parallel - \delta_0$ and the $\delta_\perp - \delta_0$ parameters:

$$\begin{aligned} \sin(\delta_\perp - \delta_\parallel) &= \cos(\delta_\parallel - \delta_0)\, \sin(\delta_\perp - \delta_0) - \sin(\delta_\parallel - \delta_0)\, \cos(\delta_\perp - \delta_0) \\ &\approx \sin(\delta_\parallel - \delta_0) - \sin(\delta_\perp - \delta_0) \end{aligned} \tag{2.65a}$$

$$\begin{aligned} \cos(\delta_\perp - \delta_\parallel) &= \cos(\delta_\parallel - \delta_0)\, \cos(\delta_\perp - \delta_0) + \sin(\delta_\parallel - \delta_0)\, \sin(\delta_\perp - \delta_0) \\ &\approx 1\,. \end{aligned} \tag{2.65b}$$

In these equations an approximation was used where the values of $\delta_{\parallel} - \delta_0$ and $\delta_{\perp} - \delta_0$ are approximately equal to π, which is justified given previous measurements of these phase differences (see for example [3]). The combination of $\sin(\delta_{\parallel} - \delta_0)$ and $\sin(\delta_{\perp} - \delta_0)$ in the $\sin(\Delta m_s\, t)$ coefficient of this term introduces a correlation between the parameters $\delta_{\parallel} - \delta_0$ and $\delta_{\perp} - \delta_0$.

Since the fraction of S-wave is small, the sensitivity for the parameter Δm_s is expected to mainly come from the "$0\perp$" and "$\parallel\perp$" interference terms. The $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ functions in these terms can be combined into a single sine function with a phase:

$$\sin\delta\, \cos(\Delta m_s\, t) - \cos\delta\, \sin(\Delta m_s\, t) = -\sin(\Delta m_s\, t - \delta)\ . \tag{2.66}$$

Because a change in the frequency of a sine function in a limited range can be partially compensated by a phase shift, these terms introduce correlations between Δm_s and the phases of the transversity amplitudes. Since practically all information on $\delta_{\perp} - \delta_0$ comes from the "$0\perp$" and "$\parallel\perp$" terms, this parameter is affected most.

Table 2.4 shows the time dependence of the differential decay rate including CP violation, but with several approximations for small parameters. Given previous measurements, it is known that the CP violation in $B_s^0 \to J/\psi\, K^+K^-$ is small, which leads to $|\lambda_s^i| \approx 1$ and $\phi_s^i \approx 0$. To parameterize also CP violation in decay and CP violation in mixing with a small parameter, $|\lambda_s^i|$ is replaced with

$$C_s^i \equiv C_{ii}^- = \frac{1 - |\lambda_s^i|^2}{1 + |\lambda_s^i|^2}\ . \tag{2.67}$$

An expansion in C_s^i and ϕ_s^i is used for the approximation in Table 2.4. Expanding the coefficients of the decay-rate equations (Equation 2.43) in terms of these parameters at first order yields

$$\begin{aligned}
C_{ij}^+ &\approx 1 + i \cdot \tfrac{1}{2}(\phi_s^i - \phi_s^j) \\
C_{ij}^- &\approx \tfrac{1}{2}(C_s^i + C_s^j) - i \cdot \tfrac{1}{2}(\phi_s^i - \phi_s^j) \\
D_{ij} &\approx -\eta_i \left[1 + i \cdot \tfrac{1}{2}(\phi_s^i - \phi_s^j)\right] \\
S_{ij} &\approx -\eta_i \left[i \cdot \tfrac{1}{2}(C_s^i - C_s^j) + \tfrac{1}{2}(\phi_s^i + \phi_s^j)\right]
\end{aligned} \tag{2.68a}$$

for $\eta_i = \eta_j$ and

$$\begin{aligned} C_{ij}^{+} &\approx \tfrac{1}{2}(C_{\mathrm{s}}^{i} + C_{\mathrm{s}}^{j}) - i \cdot \tfrac{1}{2}(\phi_{\mathrm{s}}^{i} - \phi_{\mathrm{s}}^{j}) \\ C_{ij}^{-} &\approx 1 + i \cdot \tfrac{1}{2}(\phi_{\mathrm{s}}^{i} - \phi_{\mathrm{s}}^{j}) \\ D_{ij} &\approx +\eta_i \left[\tfrac{1}{2}(C_{\mathrm{s}}^{i} - C_{\mathrm{s}}^{j}) - i \cdot \tfrac{1}{2}(\phi_{\mathrm{s}}^{i} + \phi_{\mathrm{s}}^{j})\right] \\ S_{ij} &\approx +\eta_i \left[i - \tfrac{1}{2}(\phi_{\mathrm{s}}^{i} - \phi_{\mathrm{s}}^{j})\right] \end{aligned} \tag{2.68b}$$

for $\eta_i = -\eta_j$.

Also the value of $\Delta\Gamma_{\mathrm{s}}$ is small enough to allow an approximation of the $\cosh\left(\frac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\right)$ functions. A first-order expansion in $\Delta\Gamma_{\mathrm{s}}$ gives

$$\cosh\left(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\right) \approx 1 \qquad \sinh\left(\tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\right) \approx \tfrac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\ . \tag{2.69}$$

The sines and cosines of the phase differences $\delta_{\parallel} - \delta_0$ and $\delta_{\perp} - \delta_0$ are expanded around π, as in Equation 2.65:

$$\begin{aligned} \sin(\delta_{\parallel} - \delta_0) &\approx \pi - (\delta_{\parallel} - \delta_0) & \cos(\delta_{\parallel} - \delta_0) &\approx -1 \\ \sin(\delta_{\perp} - \delta_0) &\approx \pi - (\delta_{\perp} - \delta_0) & \cos(\delta_{\perp} - \delta_0) &\approx -1 \\ \sin(\delta_{\perp} - \delta_{\parallel}) &\approx \delta_{\perp} - \delta_{\parallel} & \cos(\delta_{\perp} - \delta_{\parallel}) &\approx +1\ . \end{aligned} \tag{2.70}$$

In the table, the remaining non-trivial sine and cosine functions are denoted by

$$s_{ij} \equiv \sin(\delta_j - \delta_i) \qquad c_{ij} \equiv \cos(\delta_j - \delta_i)\ . \tag{2.71}$$

Integrating over the decay angles, the shape of the distribution of decay time is given by the sum of the diagonal terms in the differential rate. As can be seen in Table 2.4, CP violation only affects the oscillatory terms in this distribution in the first-order approximation. These terms have opposite sign for $\mathrm{B}_{\mathrm{s}}^{0}$ and $\overline{\mathrm{B}}_{\mathrm{s}}^{0}$ and can be isolated by taking the difference between the two differential rates with $q_{\mathrm{f}} = +1$ and $q_{\mathrm{f}} = -1$ (see also Equation 2.47). For a single term, this difference is approximately given by

$$\mathcal{T}_{i,+1}^{*}\mathcal{T}_{i,+1} - \mathcal{T}_{i,-1}^{*}\mathcal{T}_{i,-1} \propto e^{-\Gamma_{\mathrm{s}} t}\left[C_{\mathrm{s}}^{i}\, \cos(\Delta m_{\mathrm{s}}\, t) + \eta_i\, \phi_{\mathrm{s}}^{i}\, \sin(\Delta m_{\mathrm{s}}\, t)\right]\ . \tag{2.72}$$

The CP violation parameters also appear in the interference terms, which provides additional information on their values. This is particularly noticeable in cases where the parameters appear in the $\cosh\left(\frac{1}{2}\Delta\Gamma_{\mathrm{s}}\, t\right)$ coefficient, which is approximately equal to one and not suppressed by the small value

Table 2.4: *Functions of decay time with approximations for small CP violation, small $\Delta\Gamma_s$, and phase differences close to 0 or π.*

ij	$f(t) \times e^{+\Gamma_s t}$
ii	$\lvert A_i^{\rm CP}\rvert^2 \Big[1 - \frac{1}{2}\,\eta_i\,\Delta\Gamma_s\,t + C_s^i\,\cos(\Delta m_s\,t) + \eta_i\,\phi_s^i\,\sin(\Delta m_s\,t)\Big]$
$0\lVert$	$-\lvert A_0^{\rm CP}\rvert\lvert A_\lVert^{\rm CP}\rvert\Big[1 - \frac{1}{2}\,\Delta\Gamma_s\,t$ $+\frac{1}{2}(C_s^0 + C_s^\lVert)\cos(\Delta m_s\,t) + \frac{1}{2}(\phi_s^0 + \phi_s^\lVert)\sin(\Delta m_s\,t)\Big]$
$0\perp$	$-\lvert A_0^{\rm CP}\rvert\lvert A_\perp^{\rm CP}\rvert\Big[\frac{1}{2}(\phi_s^\perp - \phi_s^0) - \frac{1}{4}(\phi_s^0 + \phi_s^\perp)\,\Delta\Gamma_s\,t$ $-(s_{0\perp} + \frac{1}{2}\phi_s^\perp - \frac{1}{2}\phi_s^0)\cos(\Delta m_s\,t) - \sin(\Delta m_s\,t)\Big]$
$\lVert\perp$	$+\lvert A_\lVert^{\rm CP}\rvert\lvert A_\perp^{\rm CP}\rvert\Big[\frac{1}{2}(\phi_s^\perp - \phi_s^\lVert) - \frac{1}{4}(\phi_s^\lVert + \phi_s^\perp)\,\Delta\Gamma_s\,t$ $+(s_{\lVert\perp} - \frac{1}{2}\phi_s^\perp + \frac{1}{2}\phi_s^\lVert)\cos(\Delta m_s\,t) - \sin(\Delta m_s\,t)\Big]$
$0S$	$-K_{\rm PS}\lvert A_0^{\rm CP}\rvert\lvert A_S^{\rm CP}\rvert\Big[\frac{1}{2}(C_s^0 + C_s^S)c_{\perp S} - \frac{1}{2}(\phi_s^S - \phi_s^0)s_{\perp S}$ $-\big[\frac{1}{2}(C_s^S - C_s^0)c_{\perp S} - \frac{1}{2}(\phi_s^0 + \phi_s^S)s_{\perp S}\big]\cdot\frac{1}{2}\Delta\Gamma_s\,t$ $+\big[c_{\perp S} + (s_{0\perp} + \frac{1}{2}\phi_s^S - \frac{1}{2}\phi_s^0)s_{\perp S}\big]\cos(\Delta m_s\,t)$ $-\big[(s_{0\perp} + \frac{1}{2}\phi_s^S - \frac{1}{2}\phi_s^0)c_{\perp S} - s_{\perp S}\big]\sin(\Delta m_s\,t)\Big]$
$\lVert S$	$+K_{\rm PS}\lvert A_\lVert^{\rm CP}\rvert\lvert A_S^{\rm CP}\rvert\Big[\frac{1}{2}(C_s^\lVert + C_s^S)c_{\perp S} - \frac{1}{2}(\phi_s^S - \phi_s^\lVert)s_{\perp S}$ $-\big[\frac{1}{2}(C_s^S - C_s^\lVert)c_{\perp S} - \frac{1}{2}(\phi_s^\lVert + \phi_s^S)s_{\perp S}\big]\cdot\frac{1}{2}\Delta\Gamma_s\,t$ $+\big[c_{\perp S} - (s_{\lVert\perp} - \frac{1}{2}\phi_s^S + \frac{1}{2}\phi_s^\lVert)s_{\perp S}\big]\cos(\Delta m_s\,t)$ $+\big[(s_{\lVert\perp} - \frac{1}{2}\phi_s^S + \frac{1}{2}\phi_s^\lVert)c_{\perp S} + s_{\perp S}\big]\sin(\Delta m_s\,t)\Big]$
$\perp S$	$-K_{\rm PS}\lvert A_\perp^{\rm CP}\rvert\lvert A_S^{\rm CP}\rvert\Big[\big[\frac{1}{2}(\phi_s^S - \phi_s^\perp)c_{\perp S} - s_{\perp S}\big](1 + \frac{1}{2}\Delta\Gamma_s\,t)$ $-\big[\frac{1}{2}(\phi_s^S - \phi_s^\perp)c_{\perp S} + \frac{1}{2}(C_s^\perp + C_s^S)s_{\perp S}\big]\cos(\Delta m_s\,t)$ $-\big[\frac{1}{2}(C_s^S - C_s^\perp)c_{\perp S} - \frac{1}{2}(\phi_s^\perp + \phi_s^S)s_{\perp S}\big]\sin(\Delta m_s\,t)\Big]$

Table 2.5: *Functions of decay time with approximations for small CP violation, $\phi_s^i \equiv \phi_s$, $C_s^i \equiv C_s$, small $\Delta\Gamma_s$, and phase differences close to 0 or π.*

ij	$f(t) \times e^{+\Gamma_s t}$
ii	$\lvert A_i^{\rm CP}\rvert^2 \Big[1 - \frac{1}{2}\,\eta_i\,\Delta\Gamma_s\,t + C_s\,\cos(\Delta m_s\,t) + \eta_i\,\phi_s\,\sin(\Delta m_s\,t)\Big]$
$0\lVert$	$-\lvert A_0^{\rm CP}\rvert\lvert A_\lVert^{\rm CP}\rvert \Big[1 - \frac{1}{2}\,\Delta\Gamma_s\,t + C_s\,\cos(\Delta m_s\,t) + \phi_s\,\sin(\Delta m_s\,t)\Big]$
$0\perp$	$+\lvert A_0^{\rm CP}\rvert\lvert A_\perp^{\rm CP}\rvert \Big[\frac{1}{2}\,\phi_s\,\Delta\Gamma_s\,t + s_{0\perp}\,\cos(\Delta m_s\,t) + \sin(\Delta m_s\,t)\Big]$
$\lVert\perp$	$-\lvert A_\lVert^{\rm CP}\rvert\lvert A_\perp^{\rm CP}\rvert \Big[\frac{1}{2}\,\phi_s\,\Delta\Gamma_s\,t - s_{\lVert\perp}\,\cos(\Delta m_s\,t) + \sin(\Delta m_s\,t)\Big]$
0S	$-K_{\rm PS}\lvert A_0^{\rm CP}\rvert\lvert A_{\rm S}^{\rm CP}\rvert \Big[C_s\,c_{\perp{\rm S}} + \frac{1}{2}\,\phi_s\,s_{\perp{\rm S}}\,\Delta\Gamma_s\,t$ $+(c_{\perp{\rm S}} + s_{0\perp}\,s_{\perp{\rm S}})\cos(\Delta m_s\,t) - (s_{0\perp}\,c_{\perp{\rm S}} - s_{\perp{\rm S}})\sin(\Delta m_s\,t)\Big]$
$\lVert$S	$+K_{\rm PS}\lvert A_\lVert^{\rm CP}\rvert\lvert A_{\rm S}^{\rm CP}\rvert \Big[C_s\,c_{\perp{\rm S}} + \frac{1}{2}\,\phi_s\,s_{\perp{\rm S}}\,\Delta\Gamma_s\,t$ $+(c_{\perp{\rm S}} - s_{\lVert\perp}\,s_{\perp{\rm S}})\cos(\Delta m_s\,t) + (s_{\lVert\perp}\,c_{\perp{\rm S}} + s_{\perp{\rm S}})\sin(\Delta m_s\,t)\Big]$
$\perp$S	$+K_{\rm PS}\lvert A_\perp^{\rm CP}\rvert\lvert A_{\rm S}^{\rm CP}\rvert \Big[s_{\perp{\rm S}}(1 + \frac{1}{2}\Delta\Gamma_s\,t)$ $+C_s\,s_{\perp{\rm S}}\,\cos(\Delta m_s\,t) - \phi_s\,s_{\perp{\rm S}}\,\sin(\Delta m_s\,t)\Big]$

of $\Delta\Gamma_s$ or by the effects of imperfect flavour tagging. In most cases this contributes to the estimates of the differences between the ϕ_s^i parameters, but in the "0S" and "||S" terms also the values of $C_s^0 + C_s^S$ and $C_s^{\|} + C_s^S$.

Table 2.5 shows the time dependence of the differential decay rate in the same approximation, but with the additional requirement that CP violation is identical for all intermediate states. In this case the ϕ_s^i differences vanish, but $C_s^0 + C_s^S$ and $C_s^{\|} + C_s^S$ reduce to $2C_s$. As a result, a significant part of the sensitivity for C_s (or $|\lambda_s|$) comes from the "0S" and "||S" terms, despite the small value of the S-wave amplitude.

2.6.2 A Symmetry in the Equations

In the limit of equal CP violation for all intermediate states (or small differences), Equation 2.62 reveals an (approximate) symmetry in the decay-rate equations. As explained in Section 2.3.1, simultaneously applying the operations $\phi_s \longrightarrow \pi - \phi_s$ and $\Delta\Gamma_s \longrightarrow -\Delta\Gamma_s$ is equivalent to taking the complex conjugates of the coefficients $C_{ij}^{\pm}$, D_{ij}, and S_{ij} in this limit. The corresponding sign flip in the imaginary parts of the coefficients can be cancelled by flipping the sign of either $\sin(\delta_j - \delta_i)$ for Equation 2.62a or $\cos(\delta_j - \delta_i)$ for Equation 2.62b.

The specific structure of the appearance of real and imaginary parts of the products $c_i^* c_j$ in the angular coefficients in Table 2.1 enables the required sign flips in $\sin(\delta_j - \delta_i)$ and $\cos(\delta_j - \delta_i)$. Simultaneously applying the operations $\phi_s \longrightarrow \pi - \phi_s$, $\Delta\Gamma_s \longrightarrow -\Delta\Gamma_s$, $\delta_{\|} - \delta_0 \longrightarrow -(\delta_{\|} - \delta_0)$, $\delta_{\perp} - \delta_0 \longrightarrow \pi - (\delta_{\perp} - \delta_0)$, and $\delta_S - \delta_{\perp} \longrightarrow \pi - (\delta_S - \delta_{\perp})$ does not change the value of the $B_s^0 \to J/\psi K^+K^-$ differential decay rate in the limit of equal CP violation for all intermediate states. As a result, in this limit there is an alternative set of parameter values for each given set, for which the decay model fits the experimental data equally well.

To resolve this discrete ambiguity, the measurement is performed in multiple intervals of K^+K^- mass. Integrating over m_{KK} within each interval gives different values for the phase κ_{PS}, which affects the value $\delta_S - \delta_{\perp}$ that is measured for each interval. Comparing the behaviour of $\delta_S - \delta_{\perp}$ across the intervals to what is expected from the K^+K^--mass models determines which of the two ambiguous sets of parameter values corresponds to the physical situation [64].

Moving through the $\phi(1020)$ resonance from low m_{KK} to high m_{KK}, the

phase of the $\phi(1020) \to K^+K^-$ contribution roughly increases by π, while the phase of the $f_0(980) \to K^+K^-$ contribution remains approximately constant. As a result, the phase κ_{PS} is expected to increase from intervals at low m_{KK} to intervals at high m_{KK} and the phase difference $\delta_S - \delta_\perp$ is expected to decrease. This behaviour has been shown to correspond to a value of ϕ_s close to zero and a positive value of $\Delta\Gamma_s$, as opposed to ϕ_s close to π and negative $\Delta\Gamma_s$ [3, 65].

2.6.3 Parameterization

As mentioned in Section 2.6.1, the small parameter C_s^i (Equation 2.67) is used instead of $|\lambda_s|$, in addition to ϕ_s^i. As can be seen from Table 2.4, the actual CP-violation observables in the (angles-integrated) decay-time distribution are the sums of C_s^i and ϕ_s^i, weighted by the squared magnitudes of the corresponding transversity amplitudes. These observables form the coefficients of the oscillatory functions in the diagonal terms of the differential decay rate. Ignoring contributions from interference terms, these are also the observables that are measured with the assumption that CP violation is identical for all intermediate states.

To parameterize in terms of these CP-violation observables as much as possible, the following linear combinations are defined:

$$\begin{aligned}
C_s^{av} &\equiv \tfrac{1}{2}C_s^0 + \tfrac{1}{4}C_s^{\parallel} + \tfrac{1}{4}C_s^{\perp} & \phi_s^{av} &\equiv \phi_s^0 + \tfrac{1}{2}\phi_s^{\parallel} - \tfrac{1}{2}\phi_s^{\perp} \\
\Delta C_s^{\parallel} &\equiv C_s^{\parallel} - C_s^0 & \Delta\phi_s^{\parallel} &\equiv \phi_s^{\parallel} - \phi_s^0 \\
\Delta C_s^{\perp} &\equiv C_s^{\perp} - C_s^0 & \Delta\phi_s^{\perp\prime} &\equiv \phi_s^{\perp} - \tfrac{1}{2}\phi_s^0 - \tfrac{1}{2}\phi_s^{\parallel} \\
C_s^{avS} &\equiv \tfrac{1}{2}C_s^0 + \tfrac{1}{2}C_s^S & \Delta\phi_s^S &\equiv \phi_s^S - \phi_s^0
\end{aligned} \tag{2.73}$$

The parameters C_s^{av} and ϕ_s^{av} are the observables measured in the angles-integrated time distribution, ignoring interference terms, also ignoring the S-wave, and approximating the ratio $|A_0^{CP}|^2 : |A_\parallel^{CP}|^2 : |A_\perp^{CP}|^2$ by $2:1:1$. This last approximation is justified by the values reported in reference [3], which are given by $|A_0^{CP}|^2 \approx 0.52$, $|A_\parallel^{CP}|^2 \approx 0.23$, and $|A_\perp^{CP}|^2 \approx 0.25$. The normalization factors are chosen such that $C_s^{av} \to C_s$ and $\phi_s^{av} \to \phi_s$ in the case of equal CP violation for all states.

The remaining parameters are the differences between the parameters of the different states and the parameters of the longitudinal state, with the exceptions of $\Delta\phi_s^{\perp\prime}$ and C_s^{avS}. The parameter $\Delta\phi_s^{\perp\prime}$ is a combination of $\phi_s^\perp - \phi_s^0$ and $\phi_s^\perp - \phi_s^\parallel$, which both appear as coefficients of $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ in the interference terms. C_s^{avS} appears in the $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ coefficient of the "0S" term.

Using these parameters instead of $\phi_s^\perp - \phi_s^0$ and $C_s^S - C_s^0$ reduces correlations between the ϕ_s^i parameters and between the C_s^i parameters, respectively.

In the case of equal CP violation for all intermediate states, the coefficients of the decay-time functions are given by (see also Equation 2.48)

$$\begin{aligned} C_s &\equiv \frac{1-|\lambda_s|^2}{1+|\lambda_s|^2} \\ D_s &\equiv -\frac{2\,\Re(\lambda_s)}{1+|\lambda_s|^2} = -\sqrt{1-C_s^2}\,\cos\phi_s \approx -1 \\ S_s &\equiv +\frac{2\,\Im(\lambda_s)}{1+|\lambda_s|^2} = -\sqrt{1-C_s^2}\,\sin\phi_s \approx -\phi_s\,, \end{aligned} \tag{2.74}$$

where a first-order expansion in the parameters C_s and ϕ_s was used for the approximation. For historical reasons, the parameter $|\lambda_s|$ is used in this case instead of C_s.

The magnitudes of the transversity amplitudes are parameterized by their squares, $|A_i^{CP}|^2$. Since only the shape of the time and angular distributions are measured and not the absolute scale of the differential decay rate, the overall scale of the transversity amplitudes is arbitrary. To fix the scale, the magnitudes of the amplitudes are multiplied by a common factor, such that the sum of the squares for the $J/\psi\,\phi$ polarization states is equal to one:

$$|A_0^{CP}|^2 + |A_\parallel^{CP}|^2 + |A_\perp^{CP}|^2 \equiv 1\,. \tag{2.75}$$

The parameters $|A_0^{CP}|^2$ and $|A_\perp^{CP}|^2$ are used in the decay model and the value of $|A_\parallel^{CP}|^2$ follows from Equation 2.75. This procedure makes the squared magnitudes of the $J/\psi\,\phi$ amplitudes essentially polarization fractions, although one should keep in mind that these amplitudes are combinations of the B_s^0 and $\overline{B}_s^0$ values, which also contain mixing parameters (see Equation 2.44).

The magnitude of the S-wave amplitude is parameterized by a fraction, given by

$$F_S^{CP} \equiv \frac{|A_S^{CP}|^2}{|A_0^{CP}|^2 + |A_\parallel^{CP}|^2 + |A_\perp^{CP}|^2 + |A_S^{CP}|^2} = \frac{|A_S^{CP}|^2}{1+|A_S^{CP}|^2}\,. \tag{2.76}$$

As discussed in Section 2.5, differences between the phases of the transversity amplitudes are parameterized by $\delta_\parallel - \delta_0$, $\delta_\perp - \delta_0$, and $\delta_S - \delta_\perp$. Both the S-wave fraction and the phase difference between the S-wave and the $J/\psi\,\phi$

polarizations are measured in intervals of K^+K^- mass to resolve the discrete ambiguity discussed in Section 2.6.2: $F_{S,b}^{CP}$ and $\delta_{S,b} - \delta_{\perp,b}$.

Values of the parameters in the model are estimated by fitting the shape of the differential decay-rate equation to the $B_s^0 \to J/\psi K^+K^-$ data. To be able to describe the experimental data several experimental effects have to be added to the model that was discussed in this chapter. The final experimental model will be discussed in Chapter 3.

Chapter 3

Data Analysis

To extract the parameters that describe the $B_s^0 \to J/\psi K^+K^-$ decay from the experimental data, the model of decay time and decay angles, as discussed in Chapter 2, is fitted to the distribution of decays in the data. Several experimental effects have to be taken into account in this fit. Some of these effects, for instance the uncertainty in the measurement of the decay time, are included by augmenting the theoretical model. Backgrounds, on the other hand, are dealt with by selecting signal-like decay candidates in the data and by subtracting remaining background from the data after selection.

This chapter deals with the preparation of experimental data (Section 3.2) and the model that can be used to fit these data (Sections 3.3–3.6). Also the use of the decay model in the fit (Section 3.1) and in simulation (Section 3.7) are discussed.

3.1 Maximum-Likelihood Fit

The fit of the decay model to the data is an *unbinned maximum-likelihood fit*. A probability density function (PDF) is constructed by normalizing the expression for the differential decay rate, including experimental effects, by dividing by its integral over decay time and decay angles. A likelihood function of the PDF parameters for a single B_s^0 decay is given by the PDF at the values of the variables, e.g. time and angles, for that decay. The likelihood function for the full sample of decays is given by the product of all individual likelihoods.

Parameter values are estimated by maximizing the likelihood function

for the sample. In practice, the negative logarithm of the likelihood function (NLL) is minimized to find the maximum likelihood. Instead of a product, the NLL is a sum of the contributions from individual decays.

The shape of the NLL around the minimum can be approximated by a second order Taylor series, i.e. a parabola. Since, by construction, the first derivative of this function vanishes at the point where the NLL reaches its minimum, the approximation for a given parameter μ can be written in the form

$$\mathrm{NLL}(\mu) \approx \frac{1}{2\,\sigma_\mu^2}\,(\mu - \hat{\mu})^2 + C\,, \tag{3.1}$$

where $\hat{\mu}$ is the value of parameter μ in the minimum, σ_μ determines the width of the NLL shape around the minimum, and C is the NLL value in the minimum. The width, σ_μ, is (related to) the statistical uncertainty of the estimate of parameter μ. The wider the NLL shape around the minimum, the larger the value of σ_μ and the uncertainty.

It depends on the actual shape of the NLL how well it is approximated by a parabola away from the minimum. An important factor is the number of decays that is used to build the NLL. With more data the statistical uncertainties of the parameter estimates become smaller and, in general, the NLL becomes more parabolic within an interval of a few times the value of σ_μ around the minimum.

In case the distribution of parameter estimates $\hat{\mu}$ from different measurements is described by a Gaussian shape, the shape of the NLL will be truly parabolic. In this case the parameter σ_μ is an estimate of the standard deviation of the $\hat{\mu}$ distribution, which can be used as a measure of the statistical uncertainty of the parameter estimate. The parameter interval between the values $\hat{\mu} - \sigma_\mu$ and $\hat{\mu} + \sigma_\mu$ contains the true value of μ in 68% of the measurements in the Gaussian case. This is an example of a *confidence interval* with, in this case, a *coverage* of 68%. The value of σ_μ is determined from the second derivative of the NLL, which is given by $\frac{1}{\sigma_\mu^2}$.

If the shape of the NLL around its minimum is not sufficiently parabolic, the value of σ_μ and the interval defined by the values $\hat{\mu} \pm \sigma_\mu$ are still well defined, but may not provide the desired measure of the uncertainty in the parameter estimate. In general, the value of σ_μ is not a good estimate of the $\hat{\mu}$ standard deviation and the corresponding confidence interval does not have a coverage of 68%.

A more general method, which constructs a confidence interval with a

coverage of approximately 68%, defines the interval as the range of μ values between the points where the NLL reaches a value of $\frac{1}{2} + C$. In the parabolic case the NLL reaches this point at $\hat{\mu} \pm \sigma_\mu$, which results in the expected Gaussian confidence interval with corresponding uncertainty σ_μ. In a more general case this NLL value may be reached at different distances below and above the value of $\hat{\mu}$, which results in an asymmetric uncertainty.

If the NLL has multiple minima or a very broad, shallow minimum, a point estimate of the parameter value is not representative for the $\hat{\mu}$ distribution and the estimate is usually represented by confidence intervals only. A straightforward way of constructing these intervals is to determine the points between which the NLL is smaller than a certain value, as was done above with $\frac{1}{2} + C$. Common NLL values to use are the values that a parabola would reach at $n \cdot \sigma_\mu$ from the minimum, which are given by $\frac{1}{2}\, n^2 + C$. Hence the intervals are referred to as *n-sigma intervals*.

In general the NLL is a function of multiple parameters, which are estimated by the values that minimize this multivariate function. The uncertainty for an individual parameter, μ, can be obtained by minimizing the NLL with respect to all other parameters for each value of μ and applying the techniques discussed above. The likelihood function corresponding to the resulting NLL for μ is called a *profile likelihood*.

In the Gaussian case, the distribution of parameter estimates is given by a multivariate Gaussian shape. A covariance matrix containing the standard deviations and correlation coefficients of the parameter estimates takes the place of σ_μ^2 in Equation 3.1. The elements of the (inverse) covariance matrix are estimated by the second-order partial derivatives of the NLL in the minimum for all pairs of parameters. Constructing the profile likelihood for parameter μ, the minimum of the NLL in all other parameters, ν_i, for a given μ value is reached at

$$\frac{1}{\sigma_{\nu_i}} \left(\nu_i - \hat{\nu}_i\right) = \rho_{\mu\nu_i} \frac{1}{\sigma_\mu} \left(\mu - \hat{\mu}\right) , \tag{3.2}$$

where $\rho_{\mu\nu_i}$ is the correlation coefficient between the parameters μ and ν_i.

3.1.1 Fit with Weighted Decay Candidates

As will be described in Section 3.2, the time and angular distribution for $B_s^0 \to J/\psi\, K^+K^-$ signal decays is obtained by subtracting the background distribution from the distribution that is observed in the data. This is accom-

plished by adding background decay-candidates to the sample with negative weights. In the NLL, the contribution of each decay candidate is then multiplied with the value of its weight.

Although the position of the minimum of a weighted NLL still gives a good estimate of the values of the NLL parameters, the parameter uncertainties cannot be estimated directly from the shape of the NLL at its minimum any more. This can be seen by considering a fit in intervals of a given variable, where the observed number of decays in each interval is compared to the prediction of this number by a model. The uncertainties in the estimates of the model parameters are now related to the uncertainties in the observed number of decays in each interval.

For unweighted decays the distribution of the observed number of decays (N) is a Poisson distribution, for which both the mean and the variance are given by the expected number in the interval (ν). An estimate of the expected number of decays, $\hat{\nu}$, is given by N and an estimate of the corresponding uncertainty by the square root of the estimated variance, $\sqrt{N}$.

In a fit with weighted decay candidates, where each candidate counts with a weight w_c, the observed number of decays is replaced by the sum of weights, $W \equiv \sum_c w_c$. The estimate of ν is now also given by W, as expected, but the uncertainty is estimated by $\sqrt{W}$, which cannot be correct. If all weights are multiplied by a constant number, n, the relative uncertainty in the estimate of ν should not change, since no information was added to the data sample. This means that the absolute uncertainty should increase by a factor n, as $\hat{\nu} = W$ does. If the uncertainty is estimated by $\sqrt{W}$, it only increases by a factor $\sqrt{n}$.

The correct uncertainty estimate for the expected number of decays is given by the square root of $W' \equiv \sum_c w_c^2$. Unlike $\sqrt{W}$, $\sqrt{W'}$ increases by a factor n if all weights are multiplied by this common factor. To obtain this estimate of the uncertainty, the original estimate from the Poisson distribution should be divided by a factor $\sqrt{\alpha}$, where α is given by

$$\alpha \equiv \frac{W}{W'} = \frac{\sum_c w_c}{\sum_c w_c^2} \,. \tag{3.3}$$

In an unbinned maximum likelihood fit the correction factor α can be used to modify the shape of the NLL. Since the uncertainties are estimated from the second derivatives of the NLL, which are given by $\frac{1}{\sigma_\mu^2}$ in the parabolic case, the NLL is multiplied by a factor α:

$$\mathrm{NLL}'(\mu) \approx \frac{\alpha}{2\,\sigma_\mu^2}\,(\mu - \hat{\mu})^2 + C' \,. \tag{3.4}$$

This will make the shape of the parabola around the NLL minimum wider (if $\alpha < 1$) or narrower (if $\alpha > 1$). Notice that this does not affect the position of the minimum, from which the parameter values are estimated.

Note that this procedure was designed to correct the uncertainty estimates for parameter distributions with a variance that scales with the observed number of decays. In other cases the correctness of the obtained results is not guaranteed. The resulting uncertainties are verified by evaluating the shape of the distribution of parameter estimates in simulated experiments, as will be discussed in Sections 3.7 and 4.1.

3.2 Decay-Candidate Selection and Background

Whereas the decay model discussed in Chapter 2 describes the time and angular distribution of $B_s^0 \to J/\psi K^+K^-$ signal decays, the experimental data contains both signal and background decay candidates. The background distribution is subtracted from the total distribution in the data to be able to fit the model to the distribution of signal decays. Since there are statistical uncertainties associated to both the total distribution and the subtracted background distribution, the resulting uncertainties in the signal-parameter estimates become larger with an increasing amount of background decay candidates. Therefore, a selection procedure is applied, designed to reject as many background candidates as possible without removing too many signal candidates.

3.2.1 Selection

The decay-candidate selection procedure was introduced in Section 1.4.2 and is briefly described in the context of the $B_s^0 \to J/\psi K^+K^-$ decay here. See reference [66] for a detailed discussion.

As described in Section 1.4.2, the first selection requirements are applied in the L0 trigger, which only selects events with hits in the muon stations and particles with a sufficiently high (transverse) momentum. From the events that remain, the HLT reconstructs and selects $\mu^+\mu^-$ pairs that are likely to originate from a $J/\psi \to \mu^+\mu^-$ decay. Offline, the resulting $J/\psi \to \mu^+\mu^-$ candidates are matched to K^+K^- pairs to form $B_s^0 \to J/\psi K^+K^-$ decay candidates. The four tracks these candidates consist of are required to be compatible with a B_s^0 that was produced in the associated primary vertex and decayed at a

decay time above a given threshold.

In both HLT1 and HLT2 there are two different types of selections applied. The first type does not use any information on the distance that the B_s^0 travelled before it decayed and the second type does. Requiring a minimum flight distance reduces the fraction of background candidates significantly, since all four tracks originate from the primary vertex for most combinatorial background, while the tracks in a B_s^0 decay originate from a secondary vertex at some distance from the primary vertex. However, the flight distance variable is correlated with the decay time variable and requiring a minimum flight distance introduces a non-trivial selection efficiency as a function of decay time. Therefore, the sets of selection criteria that use information on the flight distance are called *decay-time biasing* or *biased*.

The unbiased HLT1 selection requires two oppositely charged muon candidates that are close enough to originate from one decay vertex and have a $\mu^+\mu^-$ invariant mass greater than 2.7 GeV/c^2. The biased HLT1 selection does not require a J/ψ candidate, but selects single tracks with a perpendicular distance to any primary vertex greater than 0.1 mm. Both of these selections reduce the number of selected events to a manageable level. Approximately 68% of the decay candidates that are finally used in the fit is selected by both the unbiased and biased selections, approximately 14% exclusively by the unbiased selection, and approximately 19% exclusively by the biased selection.

Both the unbiased and the biased HLT2 selections require $J/\psi \to \mu^+\mu^-$ candidates with an invariant $\mu^+\mu^-$-mass within a window of 0.24 GeV/c^2 centred at the J/ψ mass of 3.10 GeV/c^2. In addition, the biased selection requires a minimum *decay-length significance* (DLS) of three for the J/ψ candidate. The DLS is defined as the distance between the $\mu^+\mu^-$ vertex and the associated primary vertex (*decay length*) divided by the uncertainty on this distance. Since the decay length is a measure of the flight distance of the B_s^0, this requirement is decay-time biasing.

Without the minimum-DLS requirement the rate of events that pass the HLT2 selection would be too high to handle online. Therefore, the rate of the unbiased selection is reduced by processing only a fraction of the events that pass the HLT1 selection. To keep the selection unbiased, the processed events are randomly selected, without considering any information on the reconstructed particles. In the final data sample that is used in the fit, the number of decay candidates that is exclusively selected by the unbiased HLT2 selection is three per cent of the number of candidates selected by the biased

selection. Considering only signal candidates, the unbiased HLT2 selection adds one per cent to the total.

The method of modelling the non-trivial decay-time efficiency shape introduced by the biased HLT selections is discussed in Section 3.3.2. Because the unbiased HLT2 selection adds only few signal candidates to the final data sample and including these data complicates modelling of the efficiency significantly, only the time and angular distributions of HLT2-biased candidates are fitted. However, candidates selected by the unbiased HLT2 selection are used to extract the efficiency shapes from the $B_s^0 \to J/\psi K^+K^-$ data. The efficiencies of the biased selections are determined relative to the uniform efficiencies of the unbiased selections.

In the offline reconstruction process the muon tracks and the $\mu^+\mu^-$ vertex of $J/\psi \to \mu^+\mu^-$ candidates selected by HLT2 are combined into $B_s^0 \to J/\psi K^+K^-$ candidates with two oppositely charged kaons that form a K^+K^- vertex. The subsequent stripping selection imposes requirements on how well detector hits form tracks within experimental uncertainties, how well tracks form $\mu^+\mu^-$, K^+K^-, and $\mu^+\mu^- K^+K^-$ vertices, the particle transverse momenta, the likelihood that the particles are identified correctly as muons and kaons, the invariant masses of the reconstructed $\mu^+\mu^-$, K^+K^-, and $J/\psi K^+K^-$ combinations, and the decay time of the candidate. About twelve million $B_s^0 \to J/\psi K^+K^-$ candidates remain after this selection stage.

The selection of decay candidates is refined in the final offline stage. To visualize the effect of the selection, the distribution in $J/\psi K^+K^-$ mass of remaining decay candidates is plotted for different selection requirements in Figure 3.1. Notice that the figure shows only a part of the mass range.

For $B_s^0 \to J/\psi K^+K^-$ signal candidates this distribution is a peak around the value of the B_s^0 mass of approximately 5367 MeV/c^2. The width of this peak is determined by the experimental resolution on the $J/\psi K^+K^-$-mass measurement. The background is mainly combinatorial and follows an exponential distribution, which decreases slowly across the considered mass range. Because of the obvious difference in distributions of $J/\psi K^+K^-$ mass for the signal and the background, this variable can be used to statistically separate the two contributions.

The distribution of black data points in Figure 3.1 is for candidates that pass the stripping selection without further requirements. A peak of signal candidates is visible on top of a large background distribution. To determine the numbers of signal and background candidates, the surface areas under-

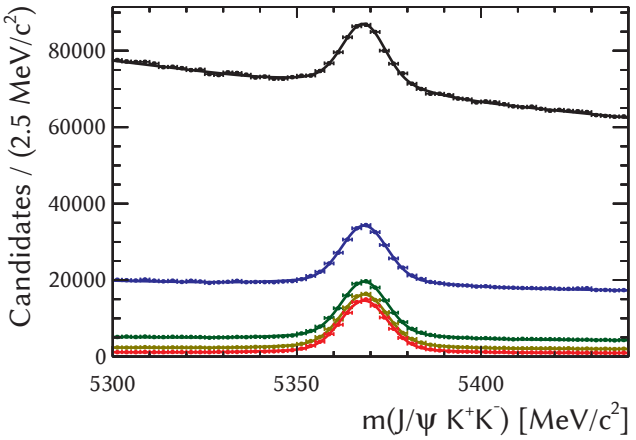


Figure 3.1: *Distribution of $B_s^0 \to J/\psi K^+K^-$ decays in $J/\psi K^+K^-$ mass for different selection criteria. The data are shown as points, while the lines represent a model that consists of the sum of two Gaussian shapes (signal) and an exponential shape (combinatorial background). Subsequent selection criteria are applied in addition to previous criteria: black: stripping selection; blue: minimum quality of B_s^0 decay-vertex reconstruction; green: minimum K^+K^- transverse momentum; yellow: minimum decay time; red: full offline selection. Notice that only a part of the total mass range is shown.*

neath the peak and the exponential distribution are determined with a fit. The signal peak is modelled with the sum of a narrow ($\sigma \approx 6\,\mathrm{MeV}/c^2$) and wide ($\sigma \approx 16\,\mathrm{MeV}/c^2$) Gaussian shape and the background with and exponential shape. The fit yields approximately 123 thousand signal candidates, which corresponds to a signal fraction of approximately one per cent.

In the rest frame of the B_s^0, the momentum of the K^+K^- pair is approximately 1.6 GeV/c for signal decays. Boosting into the frame of the detector, approximately along the beam axis, this translates into a typical transverse momentum above 1 GeV/c. The sum of the transverse momentum components of two kaons that accidentally form a suitable K^+K^- vertex in combinatorial background is often not sufficient to make a transverse momentum of 1 GeV/c. As a result, requiring this value as a minimum in the offline selection removes about 74% of the background, but only 6% of the signal, with respect to the stripping selection where already a minimum of 0.5 GeV/c was required. The $J/\psi K^+K^-$-mass distribution with the K^+K^- transverse momen-

tum requirement is shown by the blue points in Figure 3.1.

To estimate the position of the B_s^0 decay vertex and the kinematics of the particles in the decay as well as possible, these variables are determined from a fit. The fit takes the position of the primary vertex, the muon and kaon tracks, and the corresponding experimental uncertainties as inputs and minimizes a χ^2 function for the position of the decay vertex and its perpendicular distance to the flight path of the B_s^0, which should be equal to zero. The value of the χ^2 function in its minimum is a measure of the quality of the fit, i.e. of how well the four reconstructed particles form a B_s^0 decay vertex. The fit will, in general, be good for signal candidates, which results in small values of the χ^2 function.

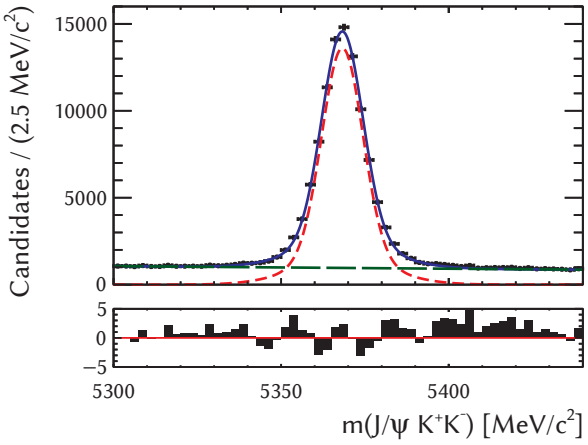
Requiring a χ^2 smaller than five times the number of degrees of freedom in the vertex fit removes about 75% of the background that remains after requiring a minimum K^+K^- transverse momentum of 1 GeV/c. Approximately 8% of the remaining signal is lost. The distribution of decay candidates after the K^+K^- transverse momentum and vertex fit quality requirements is shown by the green points in Figure 3.1.

A third requirement that removes a significant part of the background is a minimum on the reconstructed value of the decay time. For most of the background candidates the four particles originate directly from the primary vertex, which corresponds to vanishing decay time within experimental resolution. The stripping selection already requires a minimum decay time of 0.2 ps. Requiring a minimum of 0.3 ps in addition to the stripping selection and the K^+K^- transverse momentum and vertex fit χ^2 requirements removes about 54% of the remaining background and 6% of the signal. The distribution for remaining candidates is shown by the yellow points in Figure 3.1.

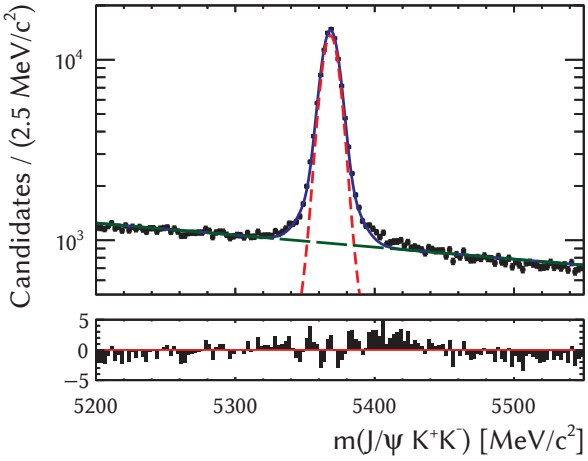
Finally applying the full offline selection removes another 62% of the background and 6% of the signal, leaving approximately 94 thousand signal candidates and 135 thousand background candidates for further analysis. This corresponds to a signal fraction of 41% in the full $J/\psi K^+K^-$-mass range of 5200–5550 MeV/c^2 after the full selection.

The final $J/\psi K^+K^-$-mass distribution is shown by the red points in Figure 3.1 and also in Figure 3.2. The latter figure shows the distribution both in the 5300–5440 MeV/c^2 mass range on a linear vertical scale (left) and in the full mass range on a logarithmic vertical scale. The signal and combinatorial background contributions as determined in the fit are shown separately.

In the areas below the $J/\psi K^+K^-$-mass distributions in Figure 3.2 the *pull*



(a)



(b)

Figure 3.2: *Distribution of $B_s^0 \to J/\psi K^+K^-$ decay candidates in $J/\psi K^+K^-$ mass in (a) the signal mass range and (b) the full mass range on a logarithmic vertical scale. The black points show the distribution of the data and the blue, solid line shows a model that was fitted to the data. The model is the sum of two Gaussian shapes for the signal (shown by the red, short-dashed line) and an exponential shape for combinatorial background (shown by the green, long-dashed line). The panels below the distributions show the pulls of the data with respect to the model (see also text).*

is shown for each mass interval. The pull is defined as the difference between the numbers of decay candidates that are observed (data) and expected (PDF), divided by the estimated statistical uncertainty of the observed value. If the PDF describes the data correctly, this quantity is expected to be distributed according to the standard normal distribution in the limit of a large number of decay candidates. Therefore, the pull quantifies how well the model describes the data in each interval.

The pulls in Figure 3.2b seem to indicate a problem with the mass model, since the PDF generally underestimates the number of decay candidates in the region of the signal peak, while it overestimates the number of candidates towards the edges of the mass range at 5200 MeV/c^2 and 5550 MeV/c^2. This issue is mainly caused by misidentified backgrounds and will be addressed in Section 3.2.2.

Notice from Figures 3.1 and 3.2 that the signal candidates are concentrated in a mass window of approximately 60 MeV/c^2, centred at the B_s^0 mass, where the signal fraction is approximately 80%. The regions to the left and the right of this *signal window* are called mass *side bands* and contain mainly background candidates. The candidates in the side bands are used to estimate the time and angular distribution of the combinatorial background and to subtract this from the total distribution in the signal window to obtain the signal distribution. This procedure is described in the following section.

3.2.2 Background Subtraction

The distribution of background decay candidates that remain after selection is subtracted from the total distribution by including background candidates with negative weights in the data sample. For combinatorial background the most straightforward method of doing this would be to give candidates in the signal region a weight equal to plus one and candidates in the $J/\psi K^+K^-$-mass side bands a small negative weight, such that the contributions of background in the signal and side-band regions exactly cancel. Note that this assumes that the background distribution in all variables of interest does not depend on the $J/\psi K^+K^-$ mass.

Because the definitions of the signal and side-band regions are rather arbitrary, a more sophisticated technique is applied. Although the bulk of the signal decay candidates have a reconstructed $J/\psi K^+K^-$ mass in a region of approximately 60 MeV/c^2 around the B_s^0 mass, the tails of the signal mass distribution stretch beyond this mass window. Moreover, at the edges of the

signal window the contribution of background is larger than in the centre. To take these features of the mass distribution into account and estimate the contribution of combinatorial background in the most optimal way, the candidate weights are not binary, but take a different value as a function of the $J/\psi K^+K^-$ mass.

The weights for subtracting combinatorial background are computed with the ${}_s\mathcal{P}lot$/sFit technique [67], which takes the $J/\psi K^+K^-$-mass distributions of the signal and the background as inputs. The resulting candidate weights are called *sWeights*, which are larger than one in the centre of the signal peak and gradually become smaller away from the peak and finally become negative in the side bands.

The ${}_s\mathcal{P}lot$/sFit method assumes that, in addition to the background distribution, also the signal distribution in the variables of interest does not depend on the $J/\psi K^+K^-$ mass. In other words, there can be no correlations between the $J/\psi K^+K^-$ mass and any other variable that is to be analysed for the signal and background distributions separately. Note that, in general, there are correlations for the sum of signal and background.

Before fitting the distribution of signal and combinatorial-background candidates and computing sWeights, contributions from other backgrounds are subtracted by adding decay candidates with negative weights to the data sample. Two additional backgrounds are considered, both originating from the misidentification of particles in the detector: $B^0 \to J/\psi K^{*0}(\to K^+\pi^-)$ and $\Lambda_b^0 \to J/\psi p K^-$. Because the detected particles for these processes originate from decays of unstable particles (resonances), these contributions are also called *resonant backgrounds*.

Figure 3.3 shows the sum of the reconstructed $J/\psi K^+K^-$-mass distributions for the two resonant backgrounds. They form a broad structure, without features that clearly distinguish these backgrounds from the signal and from combinatorial background. Therefore, the resonant backgrounds are subtracted by adding simulated $B^0 \to J/\psi K^{*0}(\to K^+\pi^-)$ and $\Lambda_b^0 \to J/\psi p K^-$ decay candidates to the data sample with negative weights that represent the estimated amount of real resonant-background data, in total approximately seven thousand candidates. Uncertainties in the numbers of background candidates and their distributions lead to systematic uncertainties in the final parameter estimates, as will be discussed in Section 4.3.

The $J/\psi K^+K^-$-mass distribution after subtracting the two resonant backgrounds is shown in Figure 3.4. Comparing to Figure 3.2, it can be seen from

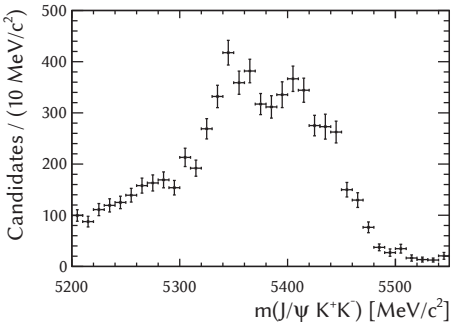


Figure 3.3: *Distribution of weighted simulated resonant-background $B_s^0 \to J/\psi K^+K^-$ decay candidates in $J/\psi K^+K^-$ mass in the full mass range.*

the pulls that the PDF follows the data more closely. The systematic underestimates of the numbers of decay candidates in the signal region and the overestimates in the side-band regions have been reduced. It appears, however, that this effect is still partially present. Moreover, the pulls in the signal region seem to be large than one would expect.

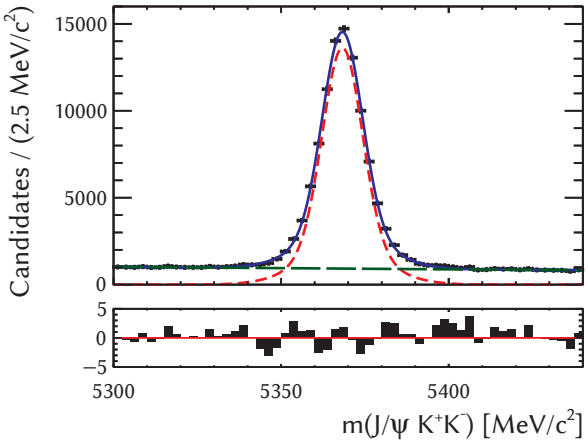
To address these issues, the double-Gaussian signal model is replaced by a more sophisticated model for the mass resolution. Figure 3.5 shows the distribution of the estimated uncertainty in the $J/\psi\,K^+K^-$-mass measurement for each decay. This distribution does not feature a structure with two sharp peaks, as would be expected for a double-Gaussian mass model, but is instead described by one broad peak. To describe the resulting mass distribution a double-sided Hypatia function with a symmetric core [68] is used, which is designed to describe the effects of an experimental resolution that varies from decay candidate to decay candidate.

The Hypatia function is defined by

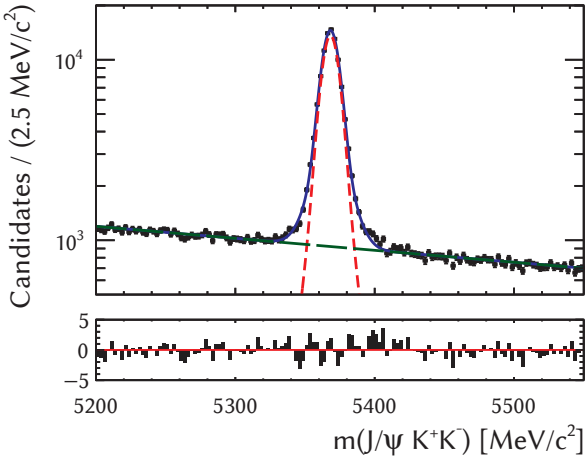
$$I(m) \equiv \begin{cases} \left[(m-\mu)^2+\delta^2\right]^{\frac{1}{2}\lambda-\frac{1}{4}} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{(m-\mu)^2+\delta^2}\right) & \\ \qquad\qquad \text{if } -a_L\,\sigma < m-\mu < +a_R\,\sigma & \\ \frac{A_L}{(B_L-m+\mu)^{n_L}} \quad \text{if } m-\mu < -a_L\,\sigma & \\ \frac{A_R}{(B_R+m-\mu)^{n_R}} \quad \text{if } m-\mu > +a_R\,\sigma & \end{cases} \tag{3.5a}$$

with

$$\alpha \equiv \frac{1}{\sigma}\sqrt{\frac{\zeta\,K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}} \qquad \text{and} \qquad \delta \equiv \sigma\sqrt{\frac{\zeta\,K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}}\,, \tag{3.5b}$$



(a)



(b)

Figure 3.4: *Distribution of $B_s^0 \to J/\psi K^+K^-$ decay candidates in $J/\psi K^+K^-$ mass after subtracting resonant backgrounds in (a) the signal mass range and (b) the full mass range on a logarithmic vertical scale. The black points show the distribution of the data and the blue, solid line shows a model that was fitted to the data. The model is the sum of two Gaussian shapes for the signal (shown by the red, short-dashed line) and an exponential shape for combinatorial background (shown by the green, long-dashed line). The panels below the distributions show the pulls of the data with respect to the model.*

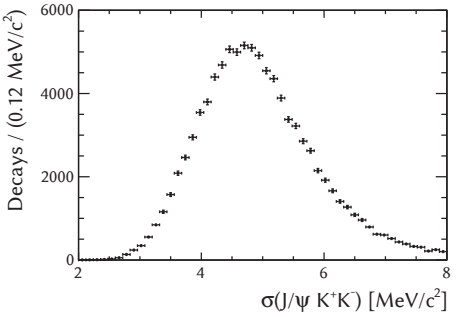


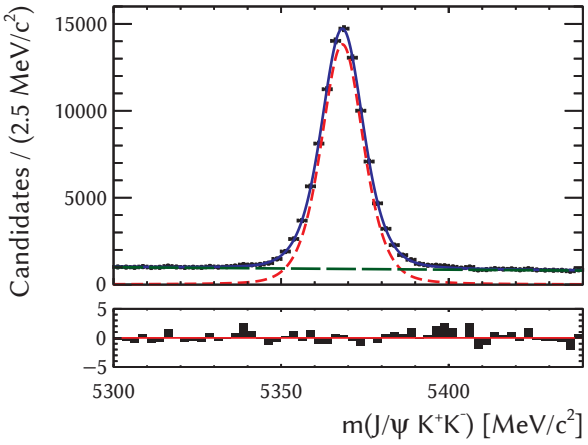
Figure 3.5: *Distribution of $B_s^0 \to J/\psi K^+K^-$ signal decays in the estimated $J/\psi K^+K^-$-mass uncertainty.*

where m is the mass variable and $K_\nu(z)$ the modified Bessel function of the second kind. The parameter μ controls the position of the mass peak, σ controls the width of its core, and λ and ζ control the shape of its core. The double-sided Hypatia function has an enhanced tail on both the left and the right side of the mass peak, the positions and shapes of which are controlled by the parameters $a_{L/R}$ and $n_{L/R}$, respectively. The parameters $A_{L/R}$ and $B_{L/R}$ are obtained by imposing continuity and differentiability at the points of transition between the core of the function and its tails.

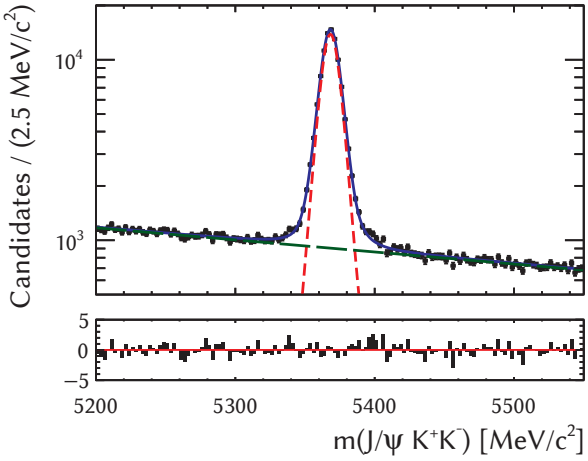
The Hypatia parameters $\lambda \approx -2.5$, $0 < \zeta < 0.5$, $a_{L/R} \approx 2.5$, and $0 < n_{L/R} < 3$ are determined from simulated $B_s^0 \to J/\psi\phi$ signal decays. Only the mean ($\mu \approx 5368\,\mathrm{MeV}/c^2$) and the width ($\sigma \approx 8\,\mathrm{MeV}/c^2$) of the core are determined in a fit to the real data. Consequently, the number of free parameters in the Hypatia model is smaller than in the double-Gaussian model, for which the mean, the two widths, and the relative contribution of the two functions were free parameters.

The result of a fit with the Hypatia signal model and an exponential shape for the combinatorial background is shown in Figure 3.6. The improvement with the Hypatia model can be seen from the pulls, which are smaller in the signal region and more symmetrically distributed around zero in comparison to Figure 3.4.

Studies of the $J/\psi K^+K^-$-mass shape of signal decays have shown that it depends on the K^+K^- invariant mass. Therefore, the parameters of the mass model are determined separately in the K^+K^--mass intervals that are also used in the final time and angular model to determine the trend in the difference



(a)



(b)

Figure 3.6: *Distribution of $B_s^0 \to J/\psi K^+K^-$ decay candidates in $J/\psi K^+K^-$ mass after subtracting resonant backgrounds in (a) the signal mass range and (b) the full mass range on a logarithmic vertical scale. The black points show the distribution of the data and the blue, solid line shows a model that was fitted to the data. The model is the sum of an Hypatia shape for the signal (shown by the red, short-dashed line) and an exponential shape for combinatorial background (shown by the green, long-dashed line). The panels below the distributions show the pulls of the data with respect to the model.*

between the $J/\psi\,\phi$ and K^+K^- S-wave phases (see Section 2.6.2). Combinatorial background is subtracted separately for each interval.

The K^+K^--mass intervals are defined in Table 3.1. In addition to this binning, the table also lists the final numbers of signal and combinatorial-background decays and the weighted-likelihood correction factors (see Section 3.1.1) for each interval.

There is no evidence for the $J/\psi\,K^+K^-$-mass distribution of the background to depend on the K^+K^- mass, but it does depend on two other variables in the data. The background distribution is found to be different for the 2011 and 2012 run periods and for decay candidates that are selected by the biased HLT2 trigger and candidates that are exclusively selected by the unbiased HLT2 trigger. Because these two variables are used in the measurement and implementation of the decay-time efficiency function (see Section 3.3.2), background subtraction is also performed separately in the four categories associated to these variables. In addition, the width of the signal peak is determined separately for HLT2-biased and not-HLT2-biased candidates.

The final parameters of the $J/\psi\,K^+K^-$-mass model that were determined from the real data are listed in Table 3.2, where the exponential function for the background mass distribution is described by $e^{-\gamma m}$, where γ is a free parameter that is different for the 2011 and 2012 runs and for HLT2-biased and not-HLT2-biased candidates. The number of signal decays determined with this model is 96 thousand and the number of combinatorial-background decays 127 thousand.

To check whether the decay time and decay angles are correlated with the $J/\psi\,K^+K^-$ mass, the mass distribution is plotted for different intervals in time and angles. A correlation appears when considering $\cos\theta_\mu$, for which the distribution is shown in the intervals $|\cos\theta_\mu| < 0.25$, $0.25 \leq |\cos\theta_\mu| < 0.7$, and $|\cos\theta_\mu| \geq 0.7$ in Figure 3.7. In all three cases the nominal mass model from Figure 3.6 is shown.

From Figure 3.7 it is clear that the $J/\psi\,K^+K^-$-mass resolution for the signal depends on the value of $\cos\theta_\mu$. This is also supported by Figure 3.8, which shows that the distribution of the estimated $J/\psi\,K^+K^-$-mass uncertainty varies for the three $\cos\theta_\mu$ regions. Formally this implies that the procedure of background subtraction does not fully work for the $\cos\theta_\mu$ distribution. The impact of this effect is studied by subtracting the background with the different mass models obtained in the $\cos\theta_\mu$ intervals and a corresponding systematic uncertainty is evaluated (see Section 4.3).

Table 3.1: *Definition of the intervals in K^+K^- invariant mass. For each interval, the numbers of signal and background candidates, and the correction factors for a likelihood with weighted candidates for the 2011 and 2012 runs (see Section 3.1.1) are shown.*

int.	range [MeV/c^2]	signal [10^3]	comb. bkg. [10^3]	α_{2011}	α_{2012}
1	990–1008	2	23	0.42	0.39
2	1008–1016	10	15	0.79	0.76
3	1016–1020	35	11	0.94	0.93
4	1020–1024	28	12	0.92	0.92
5	1024–1032	13	19	0.77	0.77
6	1032–1050	8	46	0.51	0.51

Table 3.2: *Parameters of the $J/\psi K^+K^-$-mass model that is used for background subtraction. The subscripts on the parameters refer to the interval in K^+K^- mass (1–6), the HLT2-biased/not-HLT2-biased category (b and nb, respectively), and the run period (2011 and 2012).*

par.	value [MeV/c^2]	stat. uncert. [MeV/c^2]
μ_1	5368.47	0.25
μ_2	5367.79	0.09
μ_3	5367.93	0.04
μ_4	5368.65	0.05
μ_5	5368.69	0.08
μ_6	5368.39	0.13
$\sigma_{1,\mathrm{b}}$	8.9	0.4
$\sigma_{2,\mathrm{b}}$	8.23	0.11
$\sigma_{3,\mathrm{b}}$	7.98	0.05
$\sigma_{4,\mathrm{b}}$	7.87	0.05
$\sigma_{5,\mathrm{b}}$	8.40	0.10
$\sigma_{6,\mathrm{b}}$	8.90	0.18
$\sigma_{1,\mathrm{nb}}$	14.7	3.4
$\sigma_{2,\mathrm{nb}}$	6.2	0.9
$\sigma_{3,\mathrm{nb}}$	7.0	0.4
$\sigma_{4,\mathrm{nb}}$	7.0	0.5
$\sigma_{5,\mathrm{nb}}$	7.1	0.9
$\sigma_{6,\mathrm{nb}}$	7.6	1.8

par.	value [c^2/MeV]	stat. uncert. [c^2/MeV]
$\gamma_{\mathrm{b},2011}$	0.00176	0.00006
$\gamma_{\mathrm{b},2012}$	0.00153	0.00003
$\gamma_{\mathrm{nb},2011}$	0.00032	0.00024
$\gamma_{\mathrm{nb},2012}$	0.00054	0.00015

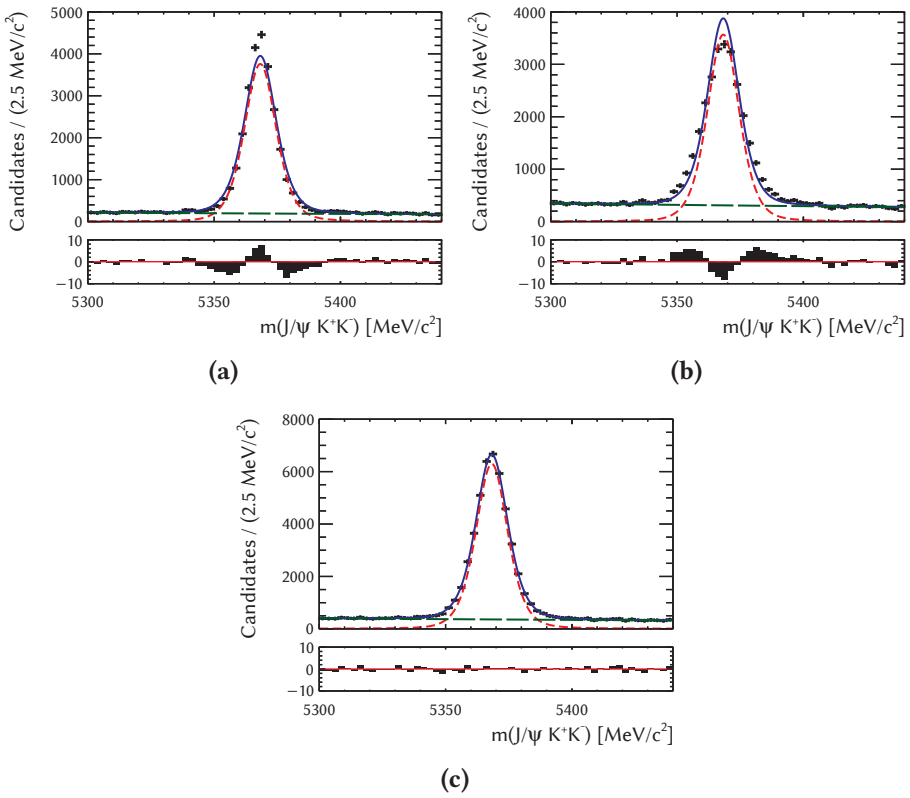


Figure 3.7: *Distribution of $B_s^0 \to J/\psi K^+K^-$ decay candidates in $J/\psi K^+K^-$ mass after subtracting resonant backgrounds in intervals of* $\cos\theta_\mu$*, where the shape of the model is set to the nominal shape from Figure 3.6. (a)* $|\cos\theta_\mu| < 0.25$*, (b)* $|\cos\theta_\mu| \geq 0.7$*, (c)* $0.25 \leq |\cos\theta_\mu| < 0.7$*. (Compare to the nominal mass distribution in Figure 3.6)*

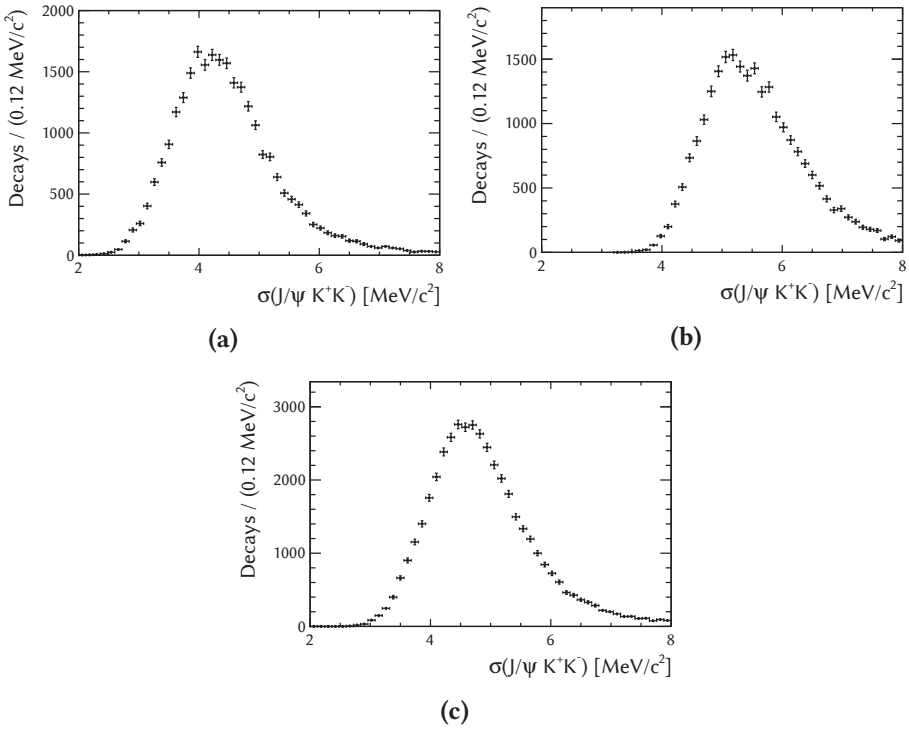


Figure 3.8: *Distribution of $B_s^0 \to J/\psi K^+K^-$ signal decays in the estimated $J/\psi K^+K^-$-mass uncertainty in different intervals of* $\cos\theta_\mu$*: (a)* $|\cos\theta_\mu| < 0.25$*, (b)* $|\cos\theta_\mu| \geq 0.7$*, and (c)* $0.25 \leq |\cos\theta_\mu| < 0.7$*.*

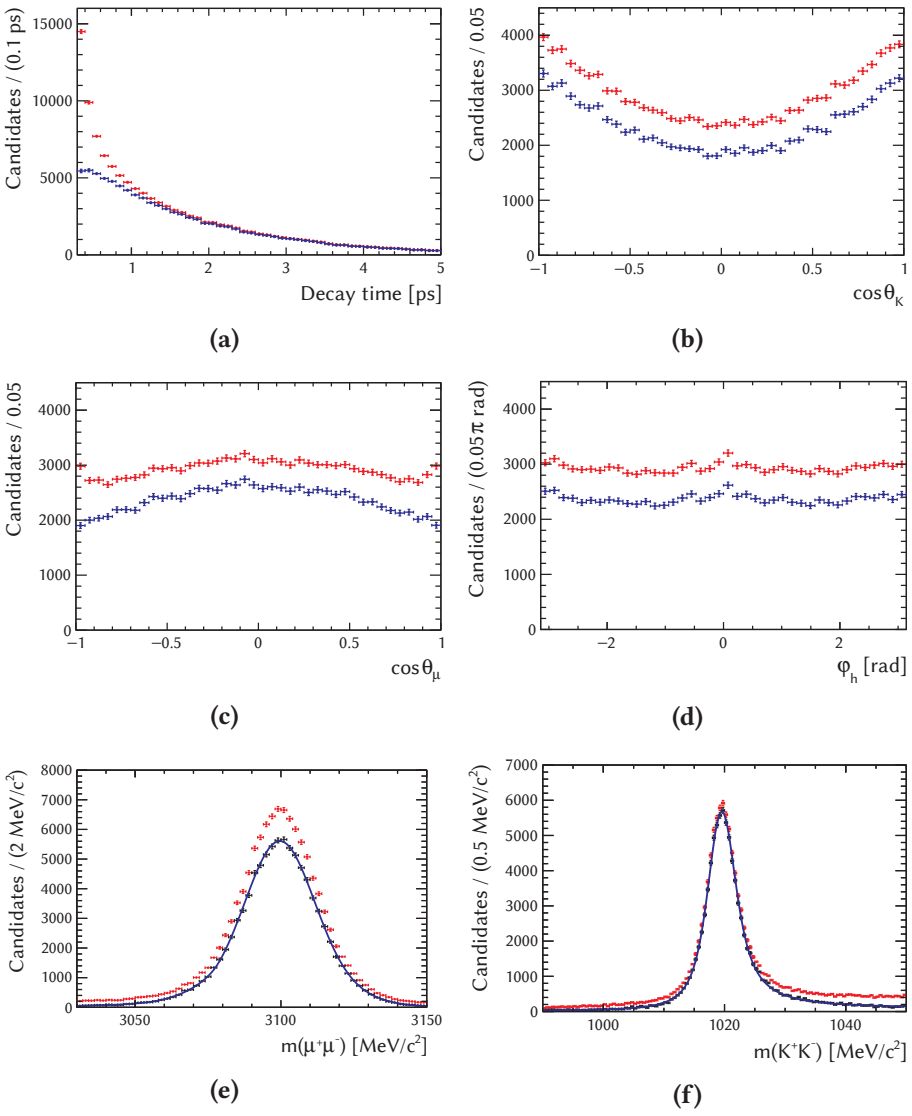


Figure 3.9: *Distributions of $B_s^0 \to J/\psi K^+K^-$ decay candidates before (red) and after (blue) background subtraction in (a) decay time, (b) $\cos\theta_K$, (c) $\cos\theta_\mu$, (d) φ_h, (e) $\mu^+\mu^-$ mass, and (f) K^+K^- mass. The distributions before background subtraction are for decay candidates in a $J/\psi K^+K^-$-mass window of 60 MeV/c^2 around the B_s^0 mass.*

Figure 3.9 shows the distributions of decay time, decay angles, $\mu^+\mu^-$ invariant mass, and K^+K^- invariant mass before and after background subtraction. The distributions before background subtraction are only shown for a $J/\psi K^+K^-$-mass window of 60 MeV/c^2 around the B_s^0 mass.

From the decay-time distribution in Figure 3.9a it can be seen that most background candidates have a small decay time. Notice that the decay-time range in this plot starts at 0.3 ps. Even though the first of the oscillations in the decay time distribution of the signal is lost, this requirement improves the measurement, because it removes a significant amount of background, as shown in Section 3.2.1.

The model that was fitted to the $\mu^+\mu^-$-mass distribution in Figure 3.9e consists of the sum of two Gaussian shapes, with a polynomial that describes the enhanced tail on the left of the mass peak, which is due to photon radiation after the decay of the J/ψ. The model for the K^+K^--mass distribution consists of a relativistic Breit-Wigner shape for the $J/\psi\,\phi$ contribution and a polynomial for the K^+K^- S-wave. From this fit, an S-wave fraction of approximately 5% is found.

Note that the models for the $\mu^+\mu^-$ and K^+K^- mass that were used to fit the distributions in Figures 3.9e and 3.9f are not used in the decay model, except for the latter in the calculation of the K_{PS} factors (see Section 3.5). The distributions of time and angles in Figures 3.9a–d are described by the model discussed in Chapter 2 with the experimental effects that will be discussed in the following sections.

3.3 Decay Time

The theoretical model of the decay time of $B_s^0 \to J/\psi K^+K^-$ decays is distorted by two experimental effects; the uncertainty in the time measurement (*resolution*) and the efficiency of the measurement (*acceptance*) as a function of time. As discussed in Section 1.4.2, the resolution is roughly 0.05 ps. The resolution model is presented in Section 3.3.1. The decay-time measurement is affected by non-trivial acceptance effects from both the trigger and reconstruction processes, which will be discussed in Section 3.3.2.

3.3.1 Resolution

The uncertainty in the decay-time measurement causes a difference between the measured decay time and the true decay time. This difference is a random variable and the resulting measured time distribution is a smeared version of the underlying true distribution. For the oscillatory $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ functions in the differential decay rate this smearing causes a decrease, or *dilution*, of the oscillation amplitude.

Not accounting for the effect of a finite resolution would lead to biases in the estimates of the parameters that describe the oscillation, towards a vanishing amplitude. As described in Section 2.6.1, the main sensitivity for the CP-violation parameters originates from the oscillation terms and hence a diluted oscillation amplitude reduces the statistical precision of the estimates for these parameters if resolution effects are modelled.

For each decay the decay-time uncertainty is estimated by propagating the uncertainties in the positions of the primary and secondary vertices and particle momenta, as discussed in Section 1.4.2. The distribution of this estimate (σ_t) is shown in Figure 3.10. For each decay, the probability for the measured decay time to deviate by a given amount from the true decay time is approximately described by a Gaussian distribution with width σ_t.

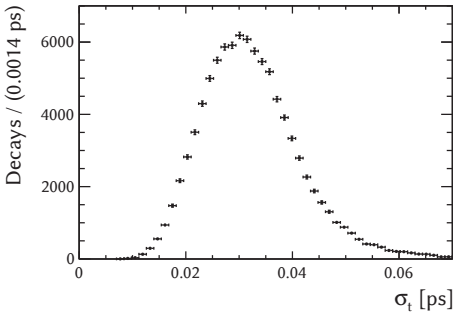


Figure 3.10: *Distribution of $B_s^0 \to J/\psi K^+K^-$ signal decays in the estimated decay-time uncertainty.*

Decay-time resolution is included in the model for the signal decay by convolving the theoretical model with the model for the difference between the measured time and the true time. This convolution is the integral of the two-dimensional PDF of the true and measured decay times over all possible

values of the former:

$$P_{\text{meas}}(t_{\text{meas}}, \Omega) \equiv \int_0^\infty \mathrm{d}t_{\text{true}}\, R(t_{\text{meas}} - t_{\text{true}}|\sigma_t)\, P_{\text{true}}(t_{\text{true}}, \Omega)\ , \tag{3.6}$$

where R is the PDF for the difference between the measured and true times, or the *resolution model*. The resolution model is conditional on σ_t., i.e. normalized with respect to decay time for each individual value of σ_t.

Although the resolution model is approximated by a Gaussian PDF with width σ_t, a more sophisticated model is required to describe the resolution in the PDF with sufficient precision. The model is a sum of two Gaussian PDFs with a common, non-zero mean and a width that depends quadratically on σ_t. The width of the first Gaussian PDF is approximately equal to σ_t, while the width of the second Gaussian PDF is roughly a factor two larger.

The double-Gaussian model is based on and validated with both simulated $\mathrm{B_s^0} \to \mathrm{J/\psi\, K^+K^-}$ decays and prompt background decay candidates. Since all tracks of prompt background candidates can be assumed to originate from the primary vertex, their "true" decay time is equal to zero and the distribution of measured decay times is essentially the resolution model. Studies of the model and its parameters are described in detail in references [66] and [2]. The values of the resolution parameters are fixed in the fit of decay time and angles. Uncertainties in the parameter values are propagated after the fit and accounted for as systematic uncertainties (see Section 4.3).

3.3.2 Acceptance

To account for the efficiencies of the decay-candidate reconstruction and selection processes, the PDF is multiplied by an acceptance function and renormalized to create a new PDF. The acceptance is modelled as a product of two functions of decay time and one function of decay angles. The angular part will be discussed in Section 3.4. This section describes the two decay-time functions.

Track-Reconstruction Acceptance

The first part of the non-trivial acceptance in decay time originates from an inefficiency in the reconstruction of particle tracks. This efficiency decreases for increasing perpendicular distance between the track and the symmetry axis of the VELO detector, which coincides with the proton beams. Since the

distance between the primary and secondary vertex and, therefore, the decay time are correlated with the distance between the four $B_s^0 \to J/\psi K^+K^-$ tracks and the beams, the efficiency also decreases with increasing decay time.

In the measurement presented here the track-reconstruction acceptance is modelled by an exponential function in *true* decay time. The advantage of this model is that its implementation in the model of the decay-time distribution is straightforward. An exponential function can be absorbed in the $e^{-\Gamma_s t}$ factor of the differential decay rate (Equation 2.59):

$$e^{-\Gamma_s t} \longrightarrow e^{\beta t}\, e^{-\Gamma_s t} = e^{-(\Gamma_s - \beta)\, t} \equiv e^{-\Gamma_s^{\text{eff}} t}\,, \tag{3.7}$$

where β is the parameter that quantifies the rate at which the efficiency changes as a function of decay time. The parameter $\Gamma_s^{\text{eff}} \equiv \Gamma_s - \beta$ can now be included in the model in the place of the parameter Γ_s.

The parameter β has been determined by a combination of studies with real and simulated data [2]. Because of changes in the online reconstruction algorithms between the 2011 and 2012 runs and the different proton-collision energies in these periods, the corresponding values of β are evaluated separately: $\beta_{2011} = -0.0090 \pm 0.0022\,\text{ps}^{-1}$ and $\beta_{2012} = -0.0124 \pm 0.0019\,\text{ps}^{-1}$. These values have to be compared with $\Gamma_s \approx 0.66\,\text{ps}^{-1}$.

Although there is no sensitivity in the $B_s^0 \to J/\psi K^+K^-$ data to the parameters Γ_s, β_{2011}, and β_{2012} separately, the value of Γ_s^{eff} can be determined in the fit of decay time separately for the 2011 and 2012 periods. Reparameterizing, there is sensitivity for the combinations $\Gamma_s - \frac{1}{2}(\beta_{2011} + \beta_{2012})$ and $\beta_{2012} - \beta_{2011}$.

To include the external information on the β values, the likelihood function of the $B_s^0 \to J/\psi K^+K^-$ model and data is multiplied by two Gaussian likelihoods that represent the external measurements. For each parameter this gives an additional parabolic term in the NLL of the form $\frac{1}{2\hat{\sigma}^2}(\beta - \hat{\beta})^2$, where the external estimate of β is denoted by $\hat{\beta}$ and the estimated uncertainty by $\hat{\sigma}$.

This procedure pulls the β estimates towards the externally determined values, effectively constraining them. Because the external constraints are much tighter than the constraints from the $B_s^0 \to J/\psi K^+K^-$ data, the values that are finally estimated in the fit are comparable to the external values:

$$\beta_{2011} = -0.0086 \pm 0.0021\,\text{ps}^{-1} \quad \text{and} \quad \beta_{2012} = -0.0127 \pm 0.0018\,\text{ps}^{-1}\,.$$

There are some limitations to this model of the track-reconstruction acceptance. The model is implemented for true decay time, whereas the external

values of the β parameters were evaluated with the measured decay time. In this measurement, this is assumed to be a good approximation, because the time scale of the variations in the model is much larger than the resolution. That is, $\beta^{-1} \approx 10^2\,\text{ps} \gg 0.05\,\text{ps}$.

In addition, the model itself is an approximation. As was shown in reference [2], the shape of the acceptance is better described by the function $1 + \beta\,t + \beta'\,t^2$ than by $e^{\beta\,t} \approx 1 + \beta\,t$. A systematic uncertainty in the parameter estimates corresponding to the assumption of the shape $e^{\beta\,t}$ is estimated in Section 4.3.

Trigger Acceptance

As described in Section 3.2.1, there are also trigger requirements that introduce non-trivial acceptance effects in decay time. The shapes of the acceptance functions corresponding to the decay-time biasing trigger categories are determined relative to the shapes of the unbiased categories, which have a uniform acceptance function.

The shape of the trigger acceptance is described in intervals of decay time. For each interval, numbers of decays are counted in the different trigger categories to determine the relative efficiency in the interval. To also include information on the exponential shape of the decay in the resulting binned acceptance function, the decay counts are varied in the fit of decay time and angles.

Since normally only the shape of the differential decay rate in time and angles is determined in the fit, additional terms need to be included in the NLL to count decays in the different trigger categories. For unweighted decays the PDF for the number of decays in a category would be a Poisson distribution, which is proportional to $\nu^n\,e^{-\nu}$, where n is the observed number of decays and ν the parameter for the expected number of decays. This PDF would give an additional NLL term of $\nu - n\,\ln\nu$. However, because each decay candidate is counted with its signal weight, this term is modified to obtain the correct uncertainty on the parameter ν from a maximum-likelihood fit.

The variable n in the Poissonian NLL term is replaced by the sum of the decay-candidate weights and the NLL term is multiplied by a factor that also depends on the weights:

$$\frac{\sum w}{\sum w^2}\left(\nu - \ln\nu \sum w\right)\,, \tag{3.8}$$

where $\sum w$ is the sum of the decay weights and $\sum w^2$ the sum of the squared decay weights. This function reaches its minimum at $\nu = \sum w$, so the estimated value for ν in a maximum-likelihood fit with only this function would be the sum of the decay weights. The inverse of the second derivative of the function, from which the uncertainty in ν is estimated, is given by $\sum w^2$ instead of $\sum w$ for an unmodified Poisson term. The former number is equal to the variance that is expected when counting weighted decay candidates.

Decays are categorized with the sets of HLT selection criteria discussed in Section 3.2.1. For the purpose of determining selection efficiencies, the combination of the HLT1 requirements and the part of the selection requirements that the two HLT2 selections have in common is termed *phase 1*. The subsequent *phase 2* then only consists of the minimum-DLS requirement in the biased HLT2 selection and the uniform reduction of decays in the unbiased HLT2 selection.

Decays that pass the selection requirements of phase 1 are selected by only the biased HLT1 requirements, by only the unbiased HLT1 requirements, or by both sets. Only the shape of the acceptance in decay time is required for the fit, which is assumed to be uniform for the unbiased phase-1 (HLT1) selection. The shape of the biased phase-1 selection is determined relative to the unbiased shape by calculating the ratio of the numbers of biased and unbiased decays in each decay-time interval.

Since the uniform selection of decays in the unbiased HLT2 selection does not depend on any of the properties of a decay, the biased and unbiased phase-2 selections are independent. That is, the efficiency of the minimum-DLS requirement does not depend on whether a decay was selected by the uniform selection and vice versa. As a result, decays selected by the unbiased phase-2 requirements are representative for the full sample of events and the biased phase-2 efficiency is given by the fraction of unbiased decays that was also selected by the minimum-DLS requirement.

Implementing these concepts in an acceptance model for the fit, six trigger categories are defined based on whether of not decays were selected by the unbiased set of requirements in phase 1, by the minimum-DLS requirement in phase 2, and by the uniform selection in phase 2. In principle this gives a total of eight categories, but in two of these no decays are selected by construction, because they are neither selected by the minimum-DLS requirement nor by the uniform phase-2 selection. If a decay is not selected by the unbiased phase-1 (HLT1) selection it can still be selected (exclusively) by

Table 3.3: *Definition of the six trigger categories that are used to determine the shape of the decay-time acceptance introduced by the HLT. For each category it is indicated if the decays it contains are selected by the unbiased phase-1 selection ("p1 unb."), by the unbiased phase-2 selection ("p2 unb."), and by the biased phase-2 selection ("p2 bias."). The ratio of the exclusively-biased and unbiased phase-1 efficiencies is given by r_1, the efficiency of the unbiased phase-2 selection by ε_{2U}, and the efficiency of the biased phase-2 selection by ε_{2B}. The number of decays that passes the unbiased phase-1 selection is given by ν_{1U}.*

category	p1 unb.	p2 unb.	p2 bias.	number of decays
1	*	*	*	$\varepsilon_{2U}\,\varepsilon_{2B}\,\nu_{1U}$
2	*		*	$(1-\varepsilon_{2U})\,\varepsilon_{2B}\,\nu_{1U}$
3	*	*		$\varepsilon_{2U}\,(1-\varepsilon_{2B})\,\nu_{1U}$
4		*	*	$r_1\,\varepsilon_{2U}\,\varepsilon_{2B}\,\nu_{1U}$
5			*	$r_1\,(1-\varepsilon_{2U})\,\varepsilon_{2B}\,\nu_{1U}$
6		*		$r_1\,\varepsilon_{2U}\,(1-\varepsilon_{2B})\,\nu_{1U}$

biased phase-1 (HLT1) selection, which leaves six non-empty categories. The categories are listed in Table 3.3.

By counting decays in the fit in each of the six trigger categories the parameters in the last column of Table 3.3 can be determined. The expression for the number of decays in the category takes the place of ν in Equation 3.8 and the parameters in the expression are varied in the fit.

The ratio of the phase-1 exclusively-biased and unbiased efficiencies, r_1, is effectively determined from the respective ratios of the numbers of decays in categories 4, 5, and 6 and the categories 1, 2, and 3. The distributions of decays over categories 1, 2, and 3 and over categories 4, 5, and 6 determine the value of the number of decays the pass the unbiased phase-1 selection, ν_{1U}, the efficiency of the phase-2 unbiased selection, ε_{2U}, and the efficiency of the phase-2 biased selection, ε_{2B}.

Because the efficiency of the biased phase-2 selection is high, 96–98%, the not-biased phase-2 categories 3 and 6 contain few decays. For this reason these two categories are merged and only the sum of weights of both categories is parameterized. This reduces the sensitivity to the parameter r_1, which is only estimated with the ratio of categories 4 and 1 and the ratio of categories 5 and 2.

Although the sums of weights are evaluated for all available decays, the

NLL for the fit of decay time and angles is built with only biased phase-2 decays. The PDF is multiplied by a factor ε_{2B} for unbiased phase-1 decays (the sum of categories 1 and 2) and a factor $r_1\,\varepsilon_{2U}$ for exclusively-biased phase-1 decays (the sum of categories 4 and 5). This simplifies the measurement, because no NLL is constructed for the remaining biased phase-2 decays. Almost no sensitivity to parameter estimates is lost, because of the small fraction of decays in these remaining categories.

The above description of the HLT acceptance function assumes that the efficiencies for the phase-2 selection are independent of the phase-1 category. That is, the values of the parameters ε_{2U} and ε_{2B} are equal for categories 1–3 and 4–6. This is based on the assumption that, at each point in decay time, the minimum-DLS requirement in the biased phase-2 selection is independent of any of the phase-1 requirements.

To test these assumptions the fit is repeated with an alternative acceptance model. In the alternative model, the unbiased phase-1 categories (1–3) are unchanged and are give estimates of the phase-2 efficiencies for unbiased phase-1 decays, $\varepsilon_{2U|1U}$ and $\varepsilon_{2B|1U}$. Instead of counting decays in all exclusively-biased phase-1 categories (4–6), only the total sum of weights for phase-2 biased decays is used in the fit (the sum of categories 4 and 5). This sum is parameterized by $r_1\,\varepsilon_{2B|1exclB}\,\nu_{1U}$.

Only the product $r_1\,\varepsilon_{2B|1exclB}$ can be estimated in this model and not these parameters separately, but this suffices, since it is exactly this product that is required in the time and angular PDF. Although the removal of assumptions in the alternative model leads to slightly larger statistical uncertainties on the efficiency parameters, the results with the two acceptance models are compatible.

The shapes of the HLT-acceptance functions that are used in the time and angular model are shown in Figure 3.11 for both the 2011 and 2012 runs. The scale of the vertical axes was divided by the efficiency of the unbiased phase-1 selection, which is undetermined. Whereas 40 decay-time intervals are used for the nominal measurement, only 11 intervals are used for these plots to reduce statistical fluctuations.

A sharp rise in efficiency, or *turn-on*, is visible in all acceptance curves just above the minimum decay time of 0.3 ps. After the turn-on, a dip in the efficiency occurs around 2 ps for the unbiased phase-1 curve for 2011 (Figure 3.11a). This dip was caused by an inefficiency in the vertex reconstruction, which was reduced in 2012 (Figure 3.11c). The exclusively-biased

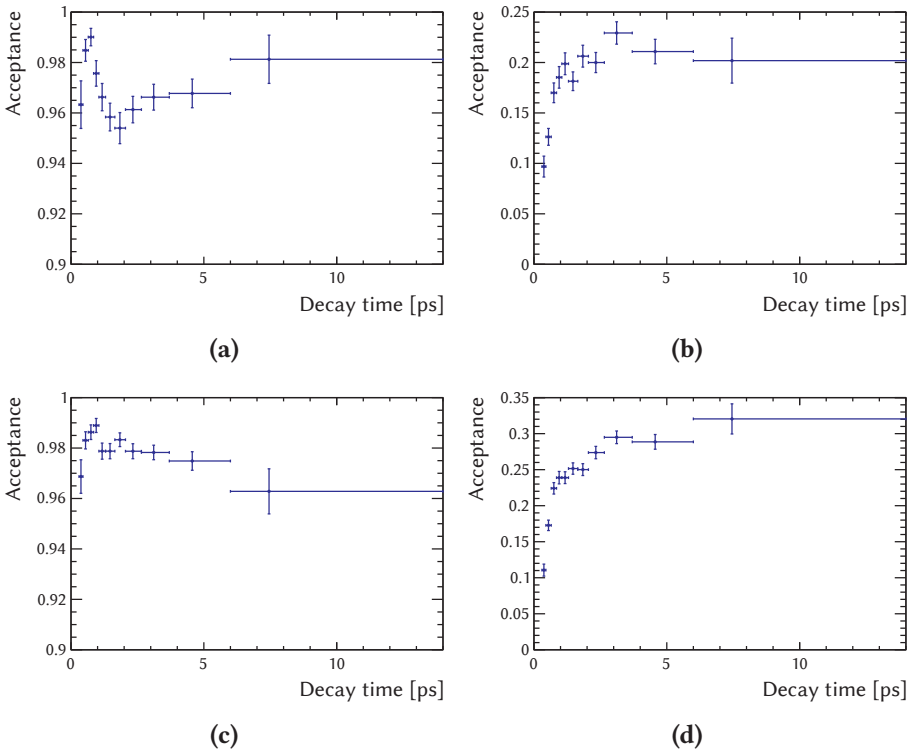


Figure 3.11: *Trigger acceptance in intervals of decay time for (a, b) the 2011 run and (c, d) the 2012 run and for (a, c) the unbiased phase-1/biased phase-2 selection and (b, d) the exclusively-biased phase-1/biased phase-2 selection. Notice that the efficiency range of the unbiased phase-1/biased phase-2 graphs is 90–100%. Because the absolute efficiency of the phase-1 selections is unknown, the scales of the efficiencies in the graphs are given with respect to the efficiency of the unbiased phase-1 selection.*

phase-1 curves (Figures 3.11b and 3.11d) have a slower turn-on than the unbiased phase-1 curves, because of the decay-time biasing requirements on muon tracks in HLT1.

3.4 Decay Angles

Also the measurement of the decay angles has a finite precision and is affected by a non-trivial acceptance shape. Whereas resolution effects in decay time directly affect the amplitude of the measured decay-time oscillation and, therefore, the estimates of the CP-violation parameters, the effect of angular resolution are indirect and expected to be smaller. Because a convolution of the angular functions in the decay model with a resolution function would be far from trivial, angular resolution is not included in the model of the decay. However, resolution effects cannot be entirely neglected and do introduce systematic uncertainties in the final parameter estimates (see Section 4.3).

The acceptance as a function of decay angles is included in the decay model. It can either be modelled as a function or by so-called *normalization weights*, as will be described in Sections 3.4.1 and 3.4.2, respectively. The shape of the acceptance function is determined from simulated decays, for which the observed angular distribution including acceptance effects can be compared to the original distribution that was generated.

The shape of the angular acceptance function is shown in Figure 3.12. The figure shows the acceptance function for each of three angles, integrated over the two remaining angles. The data points are sums of simulated decays, weighted by the inverse of the PDF that was used to generate the decays at each point in decay angles. This results in the ratio of the observed distribution including acceptance effects and the generated distribution. The shape of this ratio in the decay angles is given by the shape of the angular acceptance function. In essence this is also how the acceptance function for the decay model, represented by the blue line, is determined.

Following the above reasoning, the observed angular shape of simulated decays, after detector simulation and selection, can be expressed as the generated shape multiplied by the acceptance function. Normalizing these two shapes to obtain the observed and generated PDFs, the ratio of the two dis-

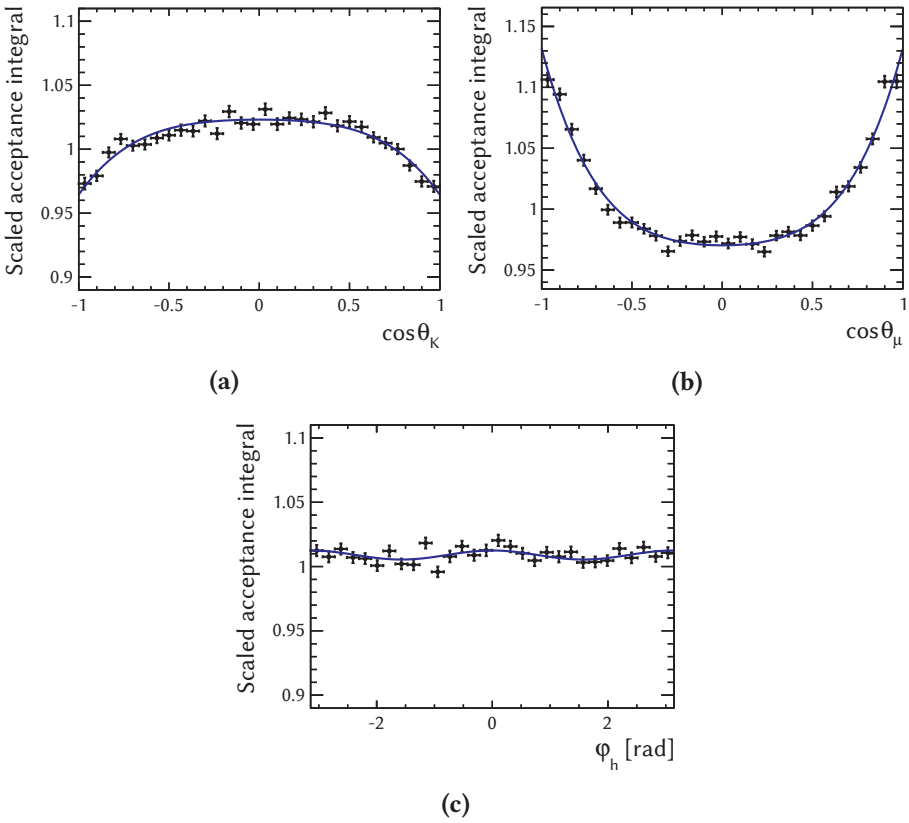


Figure 3.12: *Shape of the acceptance function in each of the three decay angles: (a)* $\cos\theta_K$*, (b)* $\cos\theta_\mu$*, (c)* φ_h*. The angular acceptance function is integrated over the two remaining angles in each of the figures. The blue line represents a parameterization of the function in terms of Legendre polynomials for* $\cos\theta_K$ *and real-valued spherical harmonics for* $\cos\theta_\mu$ *and* φ_h*. The data points are obtained by a sum over simulated decays, which are weighted by the inverse value of the PDF that was used to generate the decays at each point in decay angles. Notice that the vertical scale of these figures does not start at zero.*

tributions can be written as

$$\begin{aligned}\frac{P^{\text{obs}}(\Omega|t)}{P^{\text{gen}}(\Omega|t)} &= \frac{\varepsilon(t,\Omega)\,p(t,\Omega)}{\int \mathrm{d}\Omega\,\varepsilon(t,\Omega)\,p(t,\Omega)}\;\frac{\int \mathrm{d}\Omega\,p(t,\Omega)}{p(t,\Omega)} \\ &= \frac{\varepsilon(t,\Omega)\;\int \mathrm{d}\Omega\,p(t,\Omega)}{\int \mathrm{d}\Omega\,\varepsilon(t,\Omega)\,p(t,\Omega)} = \frac{\varepsilon(t,\Omega)}{\int \mathrm{d}\Omega\,\varepsilon(t,\Omega)\,P^{\text{gen}}(\Omega|t)} = \frac{\varepsilon(t,\Omega)}{\langle\varepsilon\rangle(t)} \\ \varepsilon(t,\Omega) &= \langle\varepsilon\rangle(t)\;\frac{P^{\text{obs}}(\Omega|t)}{P^{\text{gen}}(\Omega|t)}\,,\end{aligned} \tag{3.9}$$

where P^{obs} is the observed PDF, P^{gen} the generated PDF, p the function of time and angles from which the PDFs are built, and ε the time and angular acceptance function. Both angular PDFs are conditional on decay time. The angular mean of the acceptance function, which depends on decay time, is given by

$$\langle\varepsilon\rangle(t) \equiv \int \mathrm{d}\Omega\,\varepsilon(t,\Omega)\,P^{\text{gen}}(\Omega|t)\,. \tag{3.10}$$

Assuming the time and angular acceptance functions factorize, that is $\varepsilon(t,\Omega) = \varepsilon_t(t) \times \varepsilon_a(\Omega)$, the mean acceptance can be expressed as

$$\langle\varepsilon\rangle(t) = \varepsilon_t(t) \times \int \mathrm{d}\Omega\,\varepsilon_a(\Omega)\,P^{\text{gen}}(\Omega|t) = \varepsilon_t(t) \times \langle\varepsilon_a\rangle(t)\,. \tag{3.11}$$

Inserting this expression into the last line of Equation 3.9 yields for the acceptance function

$$\varepsilon(t,\Omega) = \varepsilon_t(t) \times \langle\varepsilon_a\rangle(t)\;\frac{P^{\text{obs}}(\Omega|t)}{P^{\text{gen}}(\Omega|t)}\,, \tag{3.12}$$

which means that the angular acceptance function is given by

$$\varepsilon_a(\Omega) = \langle\varepsilon_a\rangle(t)\;\frac{P^{\text{obs}}(\Omega|t)}{P^{\text{gen}}(\Omega|t)}\,. \tag{3.13}$$

The factor $\langle\varepsilon_a\rangle(t)$ is the mean of the angular acceptance function, which generally depends on decay time. Since only the shape of the angular acceptance function is relevant in this model, it is not required to determine the value of this angular constant factor. The ratio of the observed and generated PDFs suffices to determine the shape of the function in the decay angles.

3.4.1 Acceptance Parameterization

The function $\varepsilon_a(\Omega)$ is represented in Figure 3.12 by the blue line. For this figure, it is parameterized in terms of Legendre polynomials and real-valued spherical harmonics, $P_j(\cos\theta_\mathrm{K})\, Y_{l,\,m}(\cos\theta_\mu, \varphi_\mathrm{h})$ (see also Section A.2, Equations A.11 and A.13). These functions form an orthogonal basis of functions of the three decay angles.

In principle an arbitrary function is expressed as an infinite sum of basis functions, but in practice the relatively uniform acceptance function can be described by a limited number of contributions. The function $P_0\, Y_{0,\,0}$ is a constant and would be the only contribution for a truly uniform acceptance function. A few higher-order functions up to $j = 4$, $l = 4$, and $m = 2$ are included to describe the shape shown in Figure 3.12. Even higher orders represent faster changes in the acceptance function and can be omitted for this slowly-changing shape.

The coefficients of the basis functions, which specify the shape of the acceptance function, are determined from the simulated $\mathrm{B_s^0} \to \mathrm{J/\psi\, K^+K^-}$ decays that were used for Figure 3.12. Expressing the acceptance function as

$$\varepsilon_a(\Omega) = \sum_{j,l,m} c_{jlm}\, b_{jlm}(\Omega)\ , \tag{3.14}$$

the coefficients are defined as

$$c_{jlm} \equiv (j + \tfrac{1}{2}) \int \mathrm{d}\Omega\, b_{jlm}(\Omega)\, \varepsilon_a(\Omega)\ , \tag{3.15}$$

where b_{jlm} is a basis function and c_{jlm} the corresponding coefficient. The normalization factor $j + \frac{1}{2}$ arises from the fact that Legendre polynomials are orthogonal, but not orthonormal:

$$\int \mathrm{d}\cos\theta_\mathrm{K}\, P_j(\cos\theta_\mathrm{K}) \cdot P_j(\cos\theta_\mathrm{K}) = \frac{1}{j + \frac{1}{2}}\ . \tag{3.16}$$

The integral in Equation 3.15 is calculated by means of Monte Carlo integration, using the simulated decays. A constant is defined to absorb the factor $\langle\varepsilon_a\rangle(t)$ in Equation 3.13:

$$E_a \equiv \left[\int \mathrm{d}t\, \frac{P^\mathrm{obs}(t)}{\langle\varepsilon_a\rangle(t)}\right]^{-1}\ . \tag{3.17}$$

Notice that E_a reduces to the mean angular acceptance, $\langle\varepsilon_a\rangle$, if the PDFs for time and angles factorize, $P^{\text{gen}}(t,\Omega) = P^{\text{gen}}(t) \times P^{\text{gen}}(\Omega)$. In this case $P^{\text{gen}}(\Omega|t)$ is equal to $P^{\text{gen}}(\Omega)$ and $\langle\varepsilon_a\rangle$ is independent of time and the integral in Equation 3.17 evaluates to $\langle\varepsilon_a\rangle^{-1}$.

With the constant E_a and the definition of the angular acceptance in Equation 3.13, the coefficients of Equation 3.15 can be expressed as

$$\begin{aligned}\tfrac{1}{E_a}\, c_{jlm} &= \int \mathrm{d}t\, \frac{P^{\text{obs}}(t)}{\langle\varepsilon_a\rangle(t)} \cdot (j+\tfrac{1}{2}) \int \mathrm{d}\Omega\, b_{jlm}(\Omega)\, \varepsilon_a(\Omega) \\ &= (j+\tfrac{1}{2}) \int \mathrm{d}t\, \mathrm{d}\Omega\, \frac{P^{\text{obs}}(t)}{\langle\varepsilon_a\rangle(t)}\, b_{jlm}(\Omega)\, \langle\varepsilon_a\rangle(t)\, \frac{P^{\text{obs}}(\Omega|t)}{P^{\text{gen}}(\Omega|t)} \\ &= (j+\tfrac{1}{2}) \int \mathrm{d}t\, \mathrm{d}\Omega\, P^{\text{obs}}(t,\Omega)\, \frac{b_{jlm}(\Omega)}{P^{\text{gen}}(\Omega|t)}\,.\end{aligned} \tag{3.18}$$

The value of this integral is estimated with a sum over simulated decays:

$$E\left(\tfrac{1}{E_a}\, c_{jlm}\right) = (j+\tfrac{1}{2})\, \frac{1}{N^{\text{obs}}} \sum_{e=1}^{N^{\text{obs}}} \frac{b_{jlm}(\Omega_e)}{P^{\text{gen}}(\Omega_e|t_e)} \tag{3.19}$$

Since the "mean acceptance" is a constant equal for all coefficients, this factor can be ignored for the shape of the acceptance function. The values of the coefficients that were used for the function in Figure 3.12 are specified in Table 3.4. Note that, although not shown in the table, there are correlations between these estimates.

Table 3.4: *Values of the coefficients of the angular acceptance function that were used for the function shown in Figure 3.12, obtained from simulated $B_s^0 \to J/\psi K^+K^-$ decays. The specified values are estimates of the quantity $\frac{1}{E_a}\, c_{jlm}$ (Equation 3.19).*

coefficient (jlm)	value	statistical uncertainty
000	3.5767	0.0011
020	+0.1535	0.0030
040	+0.0305	0.0031
022	+0.0096	0.0026
200	-0.125	0.006
400	-0.033	0.008

For a uniform acceptance function, $\varepsilon_a(\Omega) = \langle\varepsilon_a\rangle = \text{constant}$, and Equation 3.18 reduces to

$$\tfrac{1}{E_a}\, c_{jlm} = (j + \tfrac{1}{2}) \int \mathrm{d}\Omega\, b_{jlm}(\Omega)\ . \tag{3.20}$$

This integral is only non-zero for $j = l = m = 0$, for which $\frac{1}{E_a}\, c_{000} = 2\sqrt{\pi} \approx 3.54$. Table 3.4 shows that the acceptance function is not too far from being uniform.

3.4.2 Acceptance Normalization Weights

In principle the parameterization of the acceptance function in terms of Legendre polynomials and spherical harmonics could be used to describe the acceptance function in the PDF for the time and angular fit. However, this would require criteria that specify which set of basis functions to include in the description and a study of the effect of not including functions that are not in this set. A slightly different approach circumvents this problem and includes all relevant acceptance information.

Using the notation of Equation 3.9, the PDF including acceptance effects is given by

$$P^{\text{obs}}(t, \Omega) = \frac{\varepsilon(t, \Omega)\, p(t, \Omega)}{\int \mathrm{d}t\, \mathrm{d}\Omega\ \varepsilon(t, \Omega)\, p(t, \Omega)} \tag{3.21}$$

Assuming the time and angular acceptance functions factorize and writing the function $p(t, \Omega)$ in the normalization integral as a sum of angular terms (see Sections 2.4 and 2.6), the PDF can be expressed as

$$P^{\text{obs}}(t, \Omega) = \varepsilon_t(t)\, \varepsilon_a(\Omega)\, \frac{p(t, \Omega)}{\sum_k \int \mathrm{d}t\, \varepsilon_t(t) f_{t,k}(t) \int \mathrm{d}\Omega\, \varepsilon_a(\Omega) f_{a,k}(\Omega)}\ , \tag{3.22}$$

where $f_{t,k}(t)$ and $f_{a,k}(\Omega)$ are the time and angular functions of a term k in the differential decay rate, respectively (see Equation 2.47 and Table 2.1).

If the angular acceptance function does not contain any free parameters, the factor $\varepsilon_a(\Omega)$ in Equation 3.22 becomes a constant term in the minimization of the NLL (see also Section 3.1) and can be ignored. The angular acceptance function then only remains in the normalization integral $\int \mathrm{d}\Omega\, \varepsilon_a(\Omega) f_{a,k}(\Omega)$. This integral is very similar to the integral in Equation 3.15 and can be estimated in the same way (Equations 3.18 and 3.19).

This procedure leads to the definition of *normalization weights*, as described in [69]. The weights are determined for each term in the differential decay rate and are defined by

$$\xi_k \equiv \int \mathrm{d}\Omega\, \varepsilon_a(\Omega) f_{a,k}(\Omega) \;, \tag{3.23}$$

with a corresponding estimate from simulated events

$$E\left(\tfrac{1}{E_a}\,\xi_k\right) = \frac{1}{N^{\mathrm{obs}}} \sum_{e=1}^{N^{\mathrm{obs}}} \frac{f_{a,k}(\Omega_e)}{P^{\mathrm{gen}}(\Omega_e|t_e)} \;. \tag{3.24}$$

The estimated values of the normalization weights are given in Table 3.5. The largest correlations between these estimates are found between the "00" weight and the "$\parallel\parallel$" and "$\perp\perp$" weights; -68% and -69%, respectively.

Table 3.5: *Values of the normalization weights of the angular acceptance function. The specified values are estimates of the quantity* $\frac{1}{E_a}\,\xi_k$ *(Equation 3.24).*

weight (k)	value	statistical uncertainty
00	0.9744	0.0005
$\parallel\parallel$	1.0245	0.0006
$\perp\perp$	1.0263	0.0006
$0\parallel$	+0.0001	0.0005
$0\perp$	-0.0003	0.0004
$\parallel\perp$	-0.0006	0.0007
SS	0.9896	0.0004
0S	+0.0007	0.0014
$\parallel$S	+0.0007	0.0006
$\perp$S	+0.0003	0.0006

For a uniform angular acceptance function, the expression for the acceptance weights reduces to

$$\tfrac{1}{E_a}\,\xi_k = \int \mathrm{d}\Omega\, f_{a,k}(\Omega) \;. \tag{3.25}$$

This integral evaluates to 1 for the "diagonal" terms "00", "$\parallel\parallel$", "$\perp\perp$", and "SS" and vanishes for the remaining interference terms. The values of the

acceptance weights confirm that the acceptance function is close to being uniform.

Since the full set of basis-function coefficients contains all the information on the angular acceptance function, the acceptance weights can be expressed in terms of the coefficients. Because the weights only contain the information that is required for the PDF normalization integral, the set of weights does not completely specify the acceptance function and the transformation from weights to coefficients is not unique. That is, there is an infinitely large set of acceptance functions that yield a given set of normalization weights.

Expressions for the acceptance weights in terms of basis coefficients are derived by expressing the decay-rate functions $f_{a,k}$ in terms of basis functions. The integral of Equation 3.23 can then be expressed as a sum of the basis-function integrals in Equation 3.15. The resulting expressions for the acceptance weights are given by

$$
\begin{aligned}
2\sqrt{\pi}\,\xi_{00} &= c_{000} - \tfrac{1}{\sqrt{5}}c_{020} + \tfrac{2}{5}c_{200} - \tfrac{2}{5\sqrt{5}}c_{220} \\
2\sqrt{\pi}\,\xi_{\parallel\parallel} &= \tfrac{1}{2}\left(2c_{000} + \tfrac{1}{\sqrt{5}}c_{020} - \sqrt{\tfrac{3}{5}}c_{022} - \tfrac{2}{5}c_{200} - \tfrac{1}{5\sqrt{5}}c_{220} + \tfrac{1}{5}\sqrt{\tfrac{3}{5}}c_{222}\right) \\
2\sqrt{\pi}\,\xi_{\perp\perp} &= \tfrac{1}{2}\left(2c_{000} + \tfrac{1}{\sqrt{5}}c_{020} + \sqrt{\tfrac{3}{5}}c_{022} - \tfrac{2}{5}c_{200} - \tfrac{1}{5\sqrt{5}}c_{220} - \tfrac{1}{5}\sqrt{\tfrac{3}{5}}c_{222}\right) \\
2\sqrt{\pi}\,\xi_{0\parallel} &= -\tfrac{3}{32}\sqrt{\tfrac{6}{5}}\pi\left(c_{121} - \tfrac{1}{4}c_{321} - \tfrac{5}{128}c_{521} - \ldots\right) \\
2\sqrt{\pi}\,\xi_{0\perp} &= +\tfrac{3}{32}\sqrt{\tfrac{6}{5}}\pi\left(c_{12\,-1} - \tfrac{1}{4}c_{32\,-1} - \tfrac{5}{128}c_{52\,-1} - \ldots\right) \\
2\sqrt{\pi}\,\xi_{\parallel\perp} &= \sqrt{\tfrac{3}{5}}\left(c_{02\,-2} - \tfrac{1}{5}c_{22\,-2}\right) \\
2\sqrt{\pi}\,\xi_{\mathrm{SS}} &= c_{000} - \tfrac{1}{\sqrt{5}}\,c_{020} \\
2\sqrt{\pi}\,\xi_{0\mathrm{S}} &= \tfrac{2}{\sqrt{3}}\left(c_{100} - \tfrac{1}{\sqrt{5}}c_{120}\right) \\
2\sqrt{\pi}\,\xi_{\parallel\mathrm{S}} &= -\tfrac{3}{8}\sqrt{\tfrac{2}{5}}\pi\left(c_{021} - \tfrac{1}{8}c_{221} - \tfrac{1}{64}c_{421} - \ldots\right) \\
2\sqrt{\pi}\,\xi_{\perp\mathrm{S}} &= -\tfrac{3}{8}\sqrt{\tfrac{2}{5}}\pi\left(c_{02\,-1} - \tfrac{1}{8}c_{22\,-1} - \tfrac{1}{64}c_{42\,-1} - \ldots\right).
\end{aligned}
\tag{3.26}
$$

3.5 K^+K^--Mass Integrals

As discussed in Sections 2.5 and 2.6.2, the dependence of the differential decay rate on K^+K^- mass is summarized in the PDF by the quantities K_{PS} and

κ_{PS}, which represent the K^+K^--mass integral of the rate. A separate PDF is constructed for each interval in K^+K^- mass, with values of K_{PS} and κ_{PS} that correspond to the K^+K^--mass shapes in the interval.

The phase κ_{PS} is absorbed in the phase difference between the $J/\psi\,\phi$ and the S-wave amplitudes (see Equations 2.57 and 2.58), which is measured for each interval. In principle, the K_{PS} values could also be measured, but because these factors only modify the magnitudes of the small S-wave interference terms more data would be needed for a sensible measurement. For this measurement the K_{PS} factors are computed and fixed in the fit of time and angles.

The decay model of Chapter 2 is integrated over K^+K^- mass, where it is assumed the mass shape of the $B_s^0 \to J/\psi\,\phi$ decay is described by a relativistic Breit-Wigner function (see e.g. reference [45]) and the K^+K^- S-wave contribution to the $B_s^0 \to J/\psi\,K^+K^-$ decay by a Flatté function [63]. Because the K^+K^--mass measurement has a finite resolution of approximately $1\,\mathrm{MeV}/c^2$, these functions are convoluted with a resolution function, which is determined from the resolution in simulated decays (see Reference [2]). The resulting K_{PS} factors are listed in the third column of Table 3.6.

Table 3.6: *S-wave–$J/\psi\,\phi$ coupling factors in the six K^+K^--mass intervals (K^i_{PS}). The K_{PS} factors are computed for three different scenarios [2]: the nominal K^+K^--mass model with a Flatté mass shape for the S-wave and the mass resolution estimated with simulated decays, a scenario with a 20% worse K^+K^--mass resolution, and a scenario with a uniform K^+K^--mass shape for the S-wave.*

int.	KK mass [MeV/c^2]	nominal	+20% resolution	uniform S-wave
1	990 – 1008	0.9178	0.9152	0.9586
2	1008 – 1016	0.9022	0.8797	0.9110
3	1016 – 1020	0.8619	0.8357	0.8618
4	1020 – 1024	0.8875	0.8599	0.8828
5	1024 – 1032	0.9360	0.9207	0.9227
6	1032 – 1050	0.9641	0.9624	0.9110

Systematic uncertainties related to the assumed mass shapes and resolution are evaluated by varying these inputs (see Section 4.3). To this end the K_{PS} factors are recalculated with a 20% worse K^+K^--mass resolution (fourth column in Table 3.6) and with a uniform K^+K^--mass shape for the S-wave (fifth column in Table 3.6).

3.6 Flavour Tagging

In Chapter 2 the expressions for the differential decay rates of B_s^0 and $\overline{B}_s^0$ decays are distinguished by the variable q_f, which takes the value +1 for B_s^0 and −1 for $\overline{B}_s^0$. Equation 2.47 in Section 2.3 shows the dependence on this variable.

As explained in Section 1.4.1, it is experimentally not possible to unambiguously determine whether the produced meson was B_s^0 and $\overline{B}_s^0$ for each individual decay. In practice, the meson flavour is estimated for each decay, together with a probability that the estimate is wrong. This *flavour tag* is denoted by q_t, which takes the values +1 for a B_s^0 estimate and −1 for a $\overline{B}_s^0$ estimate. The estimate of the wrong-tag probability is denoted by η.

For either value of q_t the differential rate is a sum of the rates for true B_s^0 ($q_f = +1$) and true $\overline{B}_s^0$ ($q_f = -1$), where the relative contributions depend on the wrong-tag probability. For small wrong-tag probability the two contributions are well separated, which enables the measurement of the oscillation amplitude from which the main sensitivity to CP-violation parameters originates (see Section 2.6.1). If the wrong-tag probability becomes larger, the opposite oscillations of true B_s^0 and true $\overline{B}_s^0$ start to cancel, reducing the sensitivity to the CP-violation parameters.

Decay candidates are assigned to different *flavour-tagging categories* according to the estimate of the wrong-tag probability. For the main measurement the candidates are only classified as *tagged* ($0 \leq \eta < 0.5$) or *untagged* ($\eta = 0.5$). For tagged decay candidates the decay model depends on the value of η for each candidate. For untagged candidates the flavour-tagging algorithms are unable to estimate the B_s^0 flavour. For these candidates the wrong-tag probability is 50% by definition.

Although the implementation of flavour tagging that uses the value of η for each individual decay candidate is more optimal, the analysis can be simplified by using an average wrong-tag probability for tagged candidates. In this case the tagging information can be used more optimally by defining several categories of tagged events, such that candidates with small and large wrong-tag probabilities are separated. This is achieved by defining ranges in η.

In summing the contributions of true B_s^0 and true $\overline{B}_s^0$ any asymmetries in their production and detection should be taken into account. The LHC collides protons, which contain more matter than antimatter. This creates a small matter–antimatter asymmetry in the fragmentation and hadronization

processes and consequently a small asymmetry in the numbers of B_s^0 and $\overline{B}_s^0$ that are produced. In addition, the flavour-tagging process relies on the detection of charged kaons, which is asymmetric for K^+ and K^-. This creates a difference in the fractions of the B_s^0 and $\overline{B}_s^0$ decays in each tagging category.

3.6.1 Formalism

The effect of any of B_s^0–$\overline{B}_s^0$ normalization asymmetries on the differential decay rate can be written in the form $1 + q_f A$, where the asymmetry is denoted by A. Examining Equation 2.47, an additional normalization asymmetry arises from CP violation in mixing. The factor $1 - q_f C_{mix}$ is included as an asymmetry contribution, with $A = -C_{mix}$.

Exploiting the relation $q_f{}^2 = +1$, the product of all asymmetries can be expressed in a general form as

$$\prod_i (1 + q_f A_i) \equiv C_{\mathcal{E}}^{avg} + q_f C_{\mathcal{O}}^{avg} , \tag{3.27}$$

where $C_{\mathcal{E}}^{avg}$ and $C_{\mathcal{O}}^{avg}$ are factors that contain the asymmetries, but not the variable q_f. Denoting the $\cosh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_s\, t\right)$ terms as $\mathcal{E}$ (even under $B_s^0 \leftrightarrow \overline{B}_s^0$) and the $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ terms as $\mathcal{O}$ (odd under $B_s^0 \leftrightarrow \overline{B}_s^0$), the differential decay rate can be expressed as

$$\begin{aligned} \frac{d^4\Gamma}{dt\, d\Omega} &= (C_{\mathcal{E}}^{avg} + q_f C_{\mathcal{O}}^{avg})(\mathcal{E} + q_f \mathcal{O}) \\ &= (C_{\mathcal{E}}^{avg} + q_f C_{\mathcal{O}}^{avg})\, \mathcal{E} + (C_{\mathcal{O}}^{avg} + q_f C_{\mathcal{E}}^{avg})\, \mathcal{O} . \end{aligned} \tag{3.28}$$

The two tagging algorithms that are used for the $B_s^0 \to J/\psi\, K^+K^-$ measurement give separate estimates of the B_s^0 flavour and the corresponding wrong-tag probability. The true wrong-tag probability for B_s^0, which is a function of the estimated probability η, is denoted by w. Because of the B_s^0–$\overline{B}_s^0$ tagging asymmetries mentioned above, the probability for $\overline{B}_s^0$ decays has a different dependence on η and is denoted by $\overline{w}$.

Expressions for the differential decay rates of different combinations of opposite-side and same-side flavour tags can be derived by multiplying the rates for true B_s^0 and true $\overline{B}_s^0$ by the appropriate combinations of wrong-tag probabilities. Assuming the probabilities for the opposite-side and same-side

algorithms are uncorrelated and labelling the algorithms by "o" and "s", respectively, the resulting rates are given by

$$\begin{aligned}
q_{\mathrm{t}}^{\mathrm{o}} = +1;\, q_{\mathrm{t}}^{\mathrm{s}} = +1: \quad & (1-w^{\mathrm{o}})(1-w^{\mathrm{s}})\,(C_{\mathcal{E}}^{\mathrm{avg}} + C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}+\mathcal{O}) \\
& + \overline{w}^{\mathrm{o}}\,\overline{w}^{\mathrm{s}}\,(C_{\mathcal{E}}^{\mathrm{avg}} - C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}-\mathcal{O}) \\
q_{\mathrm{t}}^{\mathrm{o}} = +1;\, q_{\mathrm{t}}^{\mathrm{s}} = -1: \quad & (1-w^{\mathrm{o}})\,w^{\mathrm{s}}\,(C_{\mathcal{E}}^{\mathrm{avg}} + C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}+\mathcal{O}) \\
& + \overline{w}^{\mathrm{o}}\,(1-\overline{w}^{\mathrm{s}})\,(C_{\mathcal{E}}^{\mathrm{avg}} - C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}-\mathcal{O}) \\
q_{\mathrm{t}}^{\mathrm{o}} = -1;\, q_{\mathrm{t}}^{\mathrm{s}} = +1: \quad & w^{\mathrm{s}}\,(1-w^{\mathrm{o}})\,(C_{\mathcal{E}}^{\mathrm{avg}} + C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}+\mathcal{O}) \\
& + (1-\overline{w}^{\mathrm{s}})\,\overline{w}^{\mathrm{o}}\,(C_{\mathcal{E}}^{\mathrm{avg}} - C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}-\mathcal{O}) \\
q_{\mathrm{t}}^{\mathrm{o}} = -1;\, q_{\mathrm{t}}^{\mathrm{s}} = -1: \quad & w^{\mathrm{s}}\,w^{\mathrm{o}}\,(C_{\mathcal{E}}^{\mathrm{avg}} + C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}+\mathcal{O}) \\
& + (1-\overline{w}^{\mathrm{s}})(1-\overline{w}^{\mathrm{o}})\,(C_{\mathcal{E}}^{\mathrm{avg}} - C_{\mathcal{O}}^{\mathrm{avg}})(\mathcal{E}-\mathcal{O})\,.
\end{aligned} \tag{3.29}$$

Notice that the sum of the rates of the four cases is given by the expression in Equation 3.28, summing the rates of $q_{\mathrm{f}} = +1$ and $q_{\mathrm{f}} = -1$.

Rewriting the expressions in Equation 3.29, the observed differential decay rate in one of the tagging categories can be expressed in a form similar to Equation 3.28:

$$\left(\frac{\mathrm{d}^4\Gamma}{\mathrm{d}t\,\mathrm{d}\Omega}\right)_{c,\,q_{\mathrm{t}}^{\mathrm{o}},\,q_{\mathrm{t}}^{\mathrm{s}}} = \varepsilon_c\,(C_{\mathcal{E}}\,\mathcal{E} + C_{\mathcal{O}}\,\mathcal{O})\,, \tag{3.30}$$

where the coefficients $C_{\mathcal{E}}$ and $C_{\mathcal{O}}$ both depend on the tagging category, c, and on the flavour tags, $q_{\mathrm{t}}^{\mathrm{o}}$ and $q_{\mathrm{t}}^{\mathrm{s}}$. The parameter ε_c is the average of the fractions of true $\mathrm{B}_{\mathrm{s}}^0$ and true $\overline{\mathrm{B}}_{\mathrm{s}}^0$ decays in the category. To express the coefficients in terms of these quantities, a *tagging-dilution factor* and a corresponding "asymmetry" are defined as

$$\mathcal{D} = 1 - w - \overline{w} \qquad \text{and} \qquad \Delta = \frac{w - \overline{w}}{1 - w - \overline{w}}\,. \tag{3.31}$$

In general, all tagging parameters depend on the tagging category: $\mathcal{D}_c$, Δ_c, $C_{\mathcal{E}_c}^{\mathrm{avg}}$, and $C_{\mathcal{O}_c}^{\mathrm{avg}}$. In terms of these parameters the coefficients $C_{\mathcal{E}}$ and $C_{\mathcal{O}}$ are given by

$$\begin{aligned}
2\,C_{\mathcal{E}} \equiv\; & C_{\mathcal{E}_c}^{\mathrm{avg}} + q_{\mathrm{t}}^{\mathrm{o}}\,\mathcal{D}_c^{\mathrm{o}}\,(C_{\mathcal{O}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{o}}\,C_{\mathcal{E}_c}^{\mathrm{avg}}) + q_{\mathrm{t}}^{\mathrm{s}}\,\mathcal{D}_c^{\mathrm{s}}\,(C_{\mathcal{O}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{s}}\,C_{\mathcal{E}_c}^{\mathrm{avg}}) \\
& + q_{\mathrm{t}}^{\mathrm{o}}\,q_{\mathrm{t}}^{\mathrm{s}}\,\mathcal{D}_c^{\mathrm{o}}\,\mathcal{D}_c^{\mathrm{s}}\left[(1 + \Delta_c^{\mathrm{o}}\,\Delta_c^{\mathrm{s}})\,C_{\mathcal{E}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{o}}\,\Delta_c^{\mathrm{s}}\,C_{\mathcal{O}_c}^{\mathrm{avg}}\right] \qquad (3.32a) \\
2\,C_{\mathcal{O}} \equiv\; & C_{\mathcal{O}_c}^{\mathrm{avg}} + q_{\mathrm{t}}^{\mathrm{o}}\,\mathcal{D}_c^{\mathrm{o}}\,(C_{\mathcal{E}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{o}}\,C_{\mathcal{O}_c}^{\mathrm{avg}}) + q_{\mathrm{t}}^{\mathrm{s}}\,\mathcal{D}_c^{\mathrm{s}}\,(C_{\mathcal{E}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{s}}\,C_{\mathcal{O}_c}^{\mathrm{avg}})
\end{aligned}$$

$$+ q_{\mathrm{t}}^{\mathrm{o}}\, q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{s}} \left[(1 + \Delta_c^{\mathrm{o}}\, \Delta_c^{\mathrm{s}})\, C_{\mathcal{O}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{o}}\, \Delta_c^{\mathrm{s}}\, C_{\mathcal{E}_c}^{\mathrm{avg}} \right] . \tag{3.32b}$$

Various limits can be considered for the $C_{\mathcal{E}}$ and $C_{\mathcal{O}}$ coefficients in Equations 3.30 and 3.32. Without $\mathrm{B_s^0}$–$\overline{\mathrm{B}}{}_{\mathrm{s}}^{0}$ normalization asymmetries the coefficients $C_{\mathcal{E}}^{\mathrm{avg}}$ and $C_{\mathcal{O}}^{\mathrm{avg}}$ reduce to one and zero, respectively:

$$2\, C_{\mathcal{E}} = 1 - q_{\mathrm{t}}^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{o}}\, \Delta_c^{\mathrm{o}} - q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{s}}\, \Delta_c^{\mathrm{s}} + q_{\mathrm{t}}^{\mathrm{o}}\, q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{s}}\, (1 + \Delta_c^{\mathrm{o}}\, \Delta_c^{\mathrm{s}}) \tag{3.33a}$$

$$2\, C_{\mathcal{O}} = q_{\mathrm{t}}^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{o}} + q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{s}} - q_{\mathrm{t}}^{\mathrm{o}}\, q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{s}}\, \Delta_c^{\mathrm{o}}\, \Delta_c^{\mathrm{s}} . \tag{3.33b}$$

Without any asymmetries, the coefficients are given by

$$2\, C_{\mathcal{E}} = 1 + q_{\mathrm{t}}^{\mathrm{o}}\, q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{s}} \tag{3.34a}$$

$$2\, C_{\mathcal{O}} = q_{\mathrm{t}}^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{o}} + q_{\mathrm{t}}^{\mathrm{s}}\, \mathcal{D}_c^{\mathrm{s}} . \tag{3.34b}$$

In case only one of the two flavour tags is considered, the differential rate is given by the sum of the $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_{\mathrm{s}}^{0}$ rates of the other tag. Considering only opposite-side tagging gives

$$\sum_{q_{\mathrm{t}}^{\mathrm{s}}} C_{\mathcal{E}} = C_{\mathcal{E}_c}^{\mathrm{avg}} + q_{\mathrm{t}}^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{o}} \left(C_{\mathcal{O}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{o}}\, C_{\mathcal{E}_c}^{\mathrm{avg}} \right) \tag{3.35a}$$

$$\sum_{q_{\mathrm{t}}^{\mathrm{s}}} C_{\mathcal{O}} = C_{\mathcal{O}_c}^{\mathrm{avg}} + q_{\mathrm{t}}^{\mathrm{o}}\, \mathcal{D}_c^{\mathrm{o}} \left(C_{\mathcal{E}_c}^{\mathrm{avg}} - \Delta_c^{\mathrm{o}}\, C_{\mathcal{O}_c}^{\mathrm{avg}} \right) . \tag{3.35b}$$

Effectively this is the same as considering the candidates untagged for the same-side algorithm, which gives $w \equiv \overline{w} \equiv 0.5 \longrightarrow \mathcal{D} = 0$. Without any flavour tags the coefficients are given by

$$\sum_{q_{\mathrm{t}}^{\mathrm{o}}, q_{\mathrm{t}}^{\mathrm{s}}} C_{\mathcal{E}} = 2\, C_{\mathcal{E}_c}^{\mathrm{avg}} \tag{3.36a}$$

$$\sum_{q_{\mathrm{t}}^{\mathrm{o}}, q_{\mathrm{t}}^{\mathrm{s}}} C_{\mathcal{O}} = 2\, C_{\mathcal{O}_c}^{\mathrm{avg}} . \tag{3.36b}$$

Note that without flavour tagging, one is generally not interested in an expression of the differential decay rate that depends on the tagging category. The sum of all rates cannot depend on flavour-tagging variables or parameters and should be given by

$$\sum_{c,\, q_{\mathrm{t}}^{\mathrm{o}},\, q_{\mathrm{t}}^{\mathrm{s}}} \left(\frac{\mathrm{d}^4\Gamma}{\mathrm{d}t\, \mathrm{d}\Omega} \right)_{c, q_{\mathrm{t}}^{\mathrm{o}}, q_{\mathrm{t}}^{\mathrm{s}}} = 2\, C_{\mathcal{E}_{\mathrm{S}}}^{\mathrm{avg}}\, \mathcal{E} + 2\, C_{\mathcal{O}_{\mathrm{S}}}^{\mathrm{avg}}\, \mathcal{O} , \tag{3.37}$$

where the coefficients $C_{\mathcal{E}_S}^{\text{avg}}$ and $C_{\mathcal{O}_S}^{\text{avg}}$ are defined as in Equation 3.27, but including only asymmetries that do not depend on tagging. As a result, the weighted sums of asymmetry coefficients over tagging categories are given by

$$\sum_c \varepsilon_c \, C_{\mathcal{E}_c}^{\text{avg}} = C_{\mathcal{E}_S}^{\text{avg}} \qquad \text{and} \qquad \sum_c \varepsilon_c \, C_{\mathcal{O}_c}^{\text{avg}} = C_{\mathcal{O}_S}^{\text{avg}} \, . \tag{3.38}$$

The above relations follow from the requirement that the fractions of true B_s^0 and true $\overline{\mathrm{B}}_s^0$ decays in the tagging categories both add up to one:

$$\sum_c \varepsilon_c \prod_l (1 + q_f \, A_{c,l}) \equiv 1 \tag{3.39}$$

for both values of q_f, where l only iterates over category-dependent asymmetries, $A_{c,l}$. With this requirement the sum of coefficients is given by

$$\begin{aligned} \sum_c \varepsilon_c \, (C_{\mathcal{E}_c}^{\text{avg}} + q_f \, C_{\mathcal{O}_c}^{\text{avg}}) &= \sum_c \varepsilon_c \prod_k (1 + q_f \, A_k) \prod_l (1 + q_f \, A_{c,l}) \\ &= \prod_k (1 + q_f \, A_k) \sum_c \varepsilon_c \prod_l (1 + q_f \, A_{c,l}) \\ &= (C_{\mathcal{E}_S}^{\text{avg}} + q_f \, C_{\mathcal{O}_S}^{\text{avg}}) \sum_c \varepsilon_c \prod_l (1 + q_f \, A_{c,l}) \\ &= C_{\mathcal{E}_S}^{\text{avg}} + q_f \, C_{\mathcal{O}_S}^{\text{avg}} \, , \end{aligned} \tag{3.40}$$

where k iterates over asymmetries that do not depend on the tagging category, A_k. The relations in Equation 3.38 follow from Equation 3.40.

3.6.2 Implementation

For the main measurement the data and the decay model are split into the four tagging categories that are listed in Table 3.7. The categories are combinations of the tagged and untagged categories for the opposite-side and same-side algorithms. The second column in the table gives the number of signal decays in each of the categories and the third column the effective fraction of perfectly tagged decays.

The effective fraction of perfectly tagged decays is obtained by multiplying the fraction of decays in a category by the mean of the squared dilution factor, $\langle {\mathcal{D}_c}^2 \rangle$. Ignoring normalization and wrong-tag asymmetries, the oscillatory terms in the differential decay rate, from which the main sensitivity

Table 3.7: *Definition of the flavour-tagging categories.*

category	decays [10^3]	effective fraction
untagged	31	0
OS tagged – SS untagged	13	1.2%
OS untagged – SS tagged	34	0.8%
OS tagged – SS tagged	16	1.7%

to CP-violation parameters originates, are proportional to the dilution factor (Equations 3.30 and 3.34). With a non-zero wrong-tag probability the value of this factor is between plus and minus one, which makes the measured oscillation amplitudes for B_s^0 and $\overline{B}_s^0$ tags smaller than the underlying amplitudes for true B_s^0 and $\overline{B}_s^0$ decays.

As a result, the uncertainties on the derived underlying amplitudes are a factor $\frac{1}{\mathcal{D}_c}$ larger than the uncertainties on the measured amplitudes, for a single value of the wrong-tag probability. Accounting for decays with different wrong-tag probabilities and assuming the statistical uncertainty of the measured oscillation amplitude is inversely proportional to square root of the number of decays in the category, the decrease in precision due to wrong tags can be effectively described by a decrease in number of perfectly tagged decays by a factor $\langle {\mathcal{D}_c}^2 \rangle$.

The η dependence of the wrong-tag probabilities for tagged decays, which enters the differential rates through the factors $\mathcal{D}$ and Δ, is described by a phenomenological model:

$$w = p_0 + p_1 \left(\eta - \langle\eta\rangle\right) \tag{3.41a}$$

$$\overline{w} = \bar{p}_0 + \bar{p}_1 \left(\eta - \langle\eta\rangle\right) \tag{3.41b}$$

where $\langle\eta\rangle$ is the mean value of η for tagged decays. The parameters p_0, $\bar{p}_0$, p_1, $\bar{p}_1$ are measured in a flavour-tagging calibration procedure and have different values for the opposite-side and same-side algorithms. In the calibration procedure, which is described in Reference [2], B-meson decays with known decay-time distributions are used to measure the a-priori unknown tagging parameters.

After calibration, the values of η, p_0 and p_1 are chosen such that for tagged decays $\frac{1}{2}(p_0 + \bar{p}_0) = \langle\eta\rangle$ and $\frac{1}{2}(p_1 + \bar{p}_1) = 1$, which implies $\frac{1}{2}(w + \overline{w}) = \eta$. The wrong-tag probabilities for untagged decays are given by $w \equiv \overline{w} \equiv 0.5$ by definition. The resulting distributions of η^{o} and η^{s} for tagged decays are shown

in Figure 3.13.

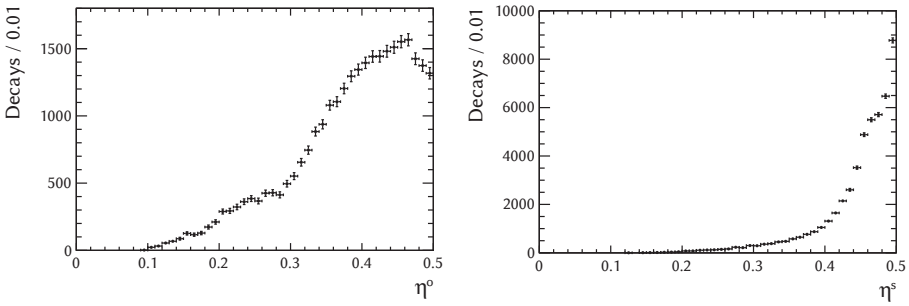


Figure 3.13: *Distribution of $B_s^0 \to J/\psi K^+K^-$ signal decays in the estimated wrong-tag probability for (a) opposite-side tagging and (b) same-side tagging. The contributions of untagged decays ($\eta = 0.5$) are not shown.*

From Figure 3.13 it can be seen that although the efficiency of same-side tagging is higher than the efficiency for opposite-side tagging, the mean wrong-tag probability of the former is closer to 0.5 ("untagged"). As a result, the effective same-side efficiency is lower than the effective opposite-side efficiency (Table 3.7): 1.2% for OS-only decays and 0.8% for SS-only decays.

Because the oscillatory functions in the "$0\perp$" and "$\|\perp$" interference terms are not proportional to unknown CP-violation parameters, there is some sensitivity for the value of the tagging dilution in the $B_s^0 \to J/\psi K^+K^-$ data. This information is combined with the value measured in the tagging-calibration process by applying the procedure that was also used for the parameters β of the reconstruction-acceptance function (Section 3.3.2). The calibration parameters are varied in the fit, representing the calibration measurements by additional parabolic terms in the NLL.

In practice the calibration parameters p_i and $\bar{p}_i$ are not measured directly, but rather their averages and differences, defined as $\frac{1}{2}(p_i + \bar{p}_i)$ and $p_i - \bar{p}_i$, respectively. The values measured after calibration [2] and the values from the fit are shown in Table 3.8. The fit values are only marginally different from the calibration values, which indicates that the estimates of these parameters in the flavour-tagging procedure are more precise than the estimates from the $B_s^0 \to J/\psi K^+K^-$ data.

An individually normalized PDF is used for each tagging category to be

Table 3.8: *Values of tagging-calibration parameters.*

parameter	calibration value	fit value
$\frac{1}{2}(p_0^{\mathrm{o}} + \overline{p}_0^{\mathrm{o}})$	0.379 ± 0.004	0.381 ± 0.004
$\frac{1}{2}(p_0^{\mathrm{s}} + \overline{p}_0^{\mathrm{s}})$	0.445 ± 0.005	0.446 ± 0.005
$\frac{1}{2}(p_1^{\mathrm{o}} + \overline{p}_1^{\mathrm{o}})$	1.00 ± 0.04	1.01 ± 0.03
$\frac{1}{2}(p_1^{\mathrm{s}} + \overline{p}_1^{\mathrm{s}})$	1.00 ± 0.09	0.97 ± 0.08
$p_0^{\mathrm{o}} - \overline{p}_0^{\mathrm{o}}$	+0.0140 ± 0.0012	+0.0140 ± 0.0012
$p_0^{\mathrm{s}} - \overline{p}_0^{\mathrm{s}}$	−0.0158 ± 0.0014	−0.0158 ± 0.0014
$p_1^{\mathrm{o}} - \overline{p}_1^{\mathrm{o}}$	+0.066 ± 0.012	+0.066 ± 0.012
$p_1^{\mathrm{s}} - \overline{p}_1^{\mathrm{s}}$	+0.008 ± 0.022	+0.008 ± 0.022

insensitive to the fractions of decays in the categories, for which there are no predictions. Since also the distributions in estimated wrong-tag probability are unpredicted, the PDFs are also made conditional on this variable and normalized individually with respect to decay time and decay angles for each combination of η^{o} and η^{s} values.

The PDFs are also normalized individually for each value of the estimates of the $\mathrm{B_s^0}$ flavour, $q_{\mathrm{t}}^{\mathrm{o}}$ and $q_{\mathrm{t}}^{\mathrm{s}}$. Although expressions for the differential decay rates do predict the relative amounts of $\mathrm{B_s^0}$ tags and $\overline{\mathrm{B}}{}_{\mathrm{s}}^{0}$ tags, a PDF with a common normalization would be very sensitive to the values of the normalization asymmetries that are used in the model. With individual normalizations this sensitivity is reduced, which justifies the assumption of no normalization asymmetries (see also Section 4.3).

To demonstrate the sensitivity to normalization asymmetries, the decay-rate equations with one flavour tag (Equation 3.35) are considered in the limit of small normalization asymmetries ($C_{\mathcal{E}}^{\mathrm{avg}} \approx 1$, $A \equiv C_{\mathcal{O}}^{\mathrm{avg}} \approx 0$), large wrong-tag probability ($\mathcal{D} \approx 0$), and no wrong-tag asymmetries ($\Delta \equiv 0$):

$$\left(\frac{\mathrm{d}^4\Gamma}{\mathrm{d}t\,\mathrm{d}\Omega}\right)_{q_{\mathrm{t}}} \approx (1 + q_{\mathrm{t}}\,A\,\mathcal{D})\,\mathcal{E} + (A + q_{\mathrm{t}}\,\mathcal{D})\,\mathcal{O}\,. \tag{3.42}$$

Using individually normalized PDFs for $q_{\mathrm{t}} = +1$ and $q_{\mathrm{t}} = -1$, an incorrect value of the normalization asymmetry A gives an incorrect estimate of the "effective dilution factor" $A \pm \mathcal{D}$. The effects of this incorrect dilution value on the estimates of the parameters contained in $\mathcal{E}$ and $\mathcal{O}$ are expected to roughly cancel between $q_{\mathrm{t}} = +1$ and $q_{\mathrm{t}} = -1$.

If instead a common normalization is used for $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$ tags, the distribution of decays in q_t is fitted, which is approximately given by

$$P_{q_\mathrm{t}} \equiv \frac{\int \mathrm{d}t\,\mathrm{d}\Omega \left(\frac{\mathrm{d}^4\Gamma}{\mathrm{d}t\,\mathrm{d}\Omega}\right)_{q_\mathrm{t}}}{\sum_{q_\mathrm{t}} \int \mathrm{d}t\,\mathrm{d}\Omega \left(\frac{\mathrm{d}^4\Gamma}{\mathrm{d}t\,\mathrm{d}\Omega}\right)_{q_\mathrm{t}}} \approx \frac{1}{2}\,\frac{1 + q_\mathrm{t}\,A\,\mathcal{D} + (A + q_\mathrm{t}\,\mathcal{D})\,R_{\mathcal{O}/\mathcal{E}}}{1 + A\,R_{\mathcal{O}/\mathcal{E}}}\,, \tag{3.43}$$

where $R_{\mathcal{O}/\mathcal{E}}$ is the ratio of the integrals over time and angles of the $\mathcal{E}$ and $\mathcal{O}$ terms. As a result, the analysis is sensitive to the asymmetry between the $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$ values, approximated by

$$\frac{P_+ - P_-}{P_+ + P_-} \approx \mathcal{D}\,\frac{A + R_{\mathcal{O}/\mathcal{E}}}{1 + A\,R_{\mathcal{O}/\mathcal{E}}}\,. \tag{3.44}$$

An incorrect estimate of the normalization asymmetry A now directly affects the estimated values of the parameters contained in the ratio $R_{\mathcal{O}/\mathcal{E}}$. Since the time integral over the many $\cos(\Delta m_\mathrm{s}\,t)$ and $\sin(\Delta m_\mathrm{s}\,t)$ periods in $\mathcal{O}$ almost vanishes, which makes $R_{\mathcal{O}/\mathcal{E}}$ very small by construction, this effect is large and will be avoided by using individually normalized PDFs.

3.7 Simulation

Datasets of simulated decays are used to study the statistical precision of parameter estimates and the effect of the detector on distributions of relevant variables. These data are produced using Monte Carlo methods, where random values are generated for the variables that describe a decay. This can be achieved either by directly generating values for the decay time and decay angles or by generating the particles in the decay and their interactions with the detector.

The latter option involves a full simulation of the proton–proton collisions of the LHC, the decays of particles that are produced, and the response of the LHCb detector to the resulting particles. These stages of the simulation are performed separately for each simulated LHC event, generating values for variables that describe particles and detector components. The resulting data are stored in the format of real detector data and are processed in the same way by applying trigger and selection requirements. This type of simulation is applied to determine the contribution of resonant backgrounds (see Section 3.2.2), the detector resolution (see Section 3.2.2 for $\mathrm{J/\psi\,K^+K^-}$ mass and

Section 3.3.1 for decay time), and the detector acceptance (see Section 3.3.2 for decay time and Section 3.4 for decay angles).

In the direct type of simulation the values of decay time and decay angles in $B_s^0 \to J/\psi K^+K^-$ decays are drawn from the PDF that is constructed with the model discussed in Chapter 2 and this chapter. All effects of B_s^0 production, $B_s^0 \to J/\psi K^+K^-$ decay, detection, reconstruction, and selection are then assumed to be described by this model, as in the fit of time and angles. This type of simulation is applied to generate so-called *pseudo experiments*, each of which consists of a dataset that is equivalent to the dataset obtained from the real experiment. Distributions of the parameter estimates in the final fit are obtained by performing the time and angular analysis on each of the pseudo experiments. These distributions are used to determine the statistical uncertainty in the parameter estimates (see Section 4.1) and the effects of variations in the assumed model.

To get a realistic estimate of the parameter distributions from the pseudo experiments not only signal decays, but also background decay candidates should be generated. Because no PDF for the time and angular distributions of these candidates is constructed for this measurement, these distributions are taken from real background data.

Background decay-candidates are selected from the real data by taking data from the $J/\psi K^+K^-$-mass side bands. As in the background-subtraction procedure, it is assumed that the signal contribution in the side bands is small (and in this case negligible) and the time and angular distributions of background candidates in the side bands are representative for the full mass range (see Section 3.2). The distribution of time and angles of the real candidates are used to generate the distribution of background in the pseudo experiments.

Values of the $J/\psi K^+K^-$ mass are generated with the mass PDF that is also used in the background-subtraction procedure. Since the mass distribution is not modelled in the final fit of decay time and angles, statistical fluctuations in this distribution are not propagated to the distributions of the fit parameters. A single set of $J/\psi K^+K^-$-mass values and corresponding signal weights is generated, which is used for all pseudo experiments. The number of decay candidates in this set is equal to the number of candidates in the real experiment.

Both the $J/\psi K^+K^-$-mass PDFs and PDFs for decay time and decay angles are conditional on a number of other variables in the decay, such as the run-period category (2011 or 2012), the K^+K^--mass category, the trigger category,

the estimated decay-time uncertainty, and the tagging variables. As for the time and angular distributions of background candidates, no PDFs are available for these conditional variables and their distributions are taken from the real data.

The distributions of conditional variables are not modelled in the time and angular fit, so a single set of values is used for all pseudo experiments, as for the $J/\psi K^+K^-$ mass. This set is "generated" by taking the values of the conditional variables from the decay candidates in the real data. Differences between the signal and the background distributions are taken into account by generating the $J/\psi K^+K^-$-mass values in intervals of the conditional variables, using the fractions of signal and background candidates that are found for each individual interval.

Pseudo experiments are produced by generating values for the decay time and the decay angles for each candidate in the set of mass and conditional-variable values. For each candidate it is first decided whether to draw values from the signal or from the background distribution for time and angles. The probabilities to get either one of the distributions are given by the signal and background fractions, which are determined by the values of the mass PDFs and the fractions in the conditional-variable interval of the decay candidate. As described above, the values of time and angles are generated with the signal PDF for "signal candidates" and with the side-band distributions for "background candidates".

The application of pseudo experiments is discussed in the next chapter, where the results of the $B_s^0 \to J/\psi K^+K^-$ measurement are presented. Parameter uncertainties as estimated from the shape of the NLL are verified with the distributions of parameter estimates in pseudo experiments. In addition to these statistical uncertainties, systematic uncertainties in the parameter estimates are evaluated.

Chapter 4

Results

The model of decay time and decay angles in the $B_s^0 \to J/\psi K^+K^-$ decay, as presented in Chapters 2 and 3, is fitted to the data from reconstructed and selected B_s^0 and $\overline{B}_s^0$ candidates, after background subtraction. Statistical uncertainties in the resulting parameter estimates are evaluated from the shape of the likelihood function, as described in Chapter 3. Systematic uncertainties are estimated by repeating the fit with variations of the model or the data.

Results are presented for three different parameterizations of CP violation. Different assumptions are made, based on the conjectures that the $B_s^0 \to J/\psi K^+K^-$ decay is dominated by a tree-level amplitude and that CP violation in mixing is small, as discussed in Section 1.3.2. If contributions from additional decay amplitudes are small, CP violation is induced by the B_s^0–$\overline{B}_s^0$ mixing process and does not depend on the intermediate angular-momentum state. Also CP violation in decay will be small in this case.

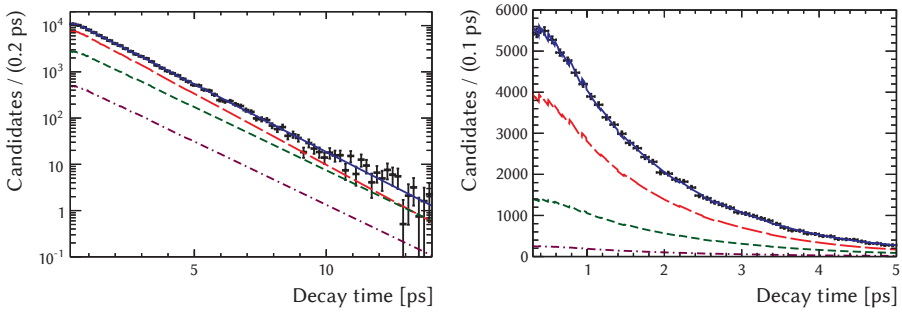
The first parameterization is the most general one and includes different parameters ϕ_s^i and C_s^i for the states $i \in \{0, \|, \perp, S\}$. See Section 2.6.3 for a description of the actual parameters that are used. The second parameterization is more restrictive and assumes that CP-violating effects are identical among the intermediate states. Finally, for the third parameterization the additional assumptions of no CP violation in mixing and no CP-violation in decay are made, i.e., only CP violation in the interference between mixed and unmixed decays is allowed. The following parameters describe CP violation for these three cases:

1. ϕ_s^i/C_s^i model:
   $\phi_s^{\text{av}}, \Delta\phi_s^{\|}, \Delta\phi_s^{\perp\prime}, \Delta\phi_s^{S}$ and $C_s^{\text{av}}, \Delta C_s^{\|}, \Delta C_s^{\perp}, C_s^{\text{avS}}$

2. $\phi_s/|\lambda_s|$ model:
   $\phi_s^0 = \phi_s^{\parallel} = \phi_s^{\perp} = \phi_s^S \equiv \phi_s$ and $|\lambda_s^0| = |\lambda_s^{\parallel}| = |\lambda_s^{\perp}| = |\lambda_s^S| \equiv |\lambda_s|$

3. $\phi_s/|\lambda_s| = 1$ model:
   $\phi_s^0 = \phi_s^{\parallel} = \phi_s^{\perp} = \phi_s^S \equiv \phi_s$ and $|\lambda_s^0| = |\lambda_s^{\parallel}| = |\lambda_s^{\perp}| = |\lambda_s^S| \equiv 1$

Figures 4.1 and 4.2 show the background-subtracted distributions of decays in time and angles and the corresponding one-dimensional PDF projections for case 2, the $\phi_s/|\lambda_s|$ model. The decay-time distribution is shown in both the full selected range of [0.3, 14] ps on a logarithmic vertical scale (Figure 4.1a) and in a reduced range of [0.3, 5] ps on a linear vertical scale (Figure 4.1b).



(a) **(b)**

Figure 4.1: *Background-subtracted distribution of decays in decay time (data points) and the corresponding one-dimensional projection of the PDF (blue line). Figure (a) shows the distribution in the full range between 0.3 and 14 ps on a logarithmic vertical scale, while Figure (b) shows the distribution between 0.3 and 5 ps on a linear vertical scale. Additional PDF projections are shown for the CP-even (long-dashed, red line) and CP-odd (short-dashed, green line) components of $B_s^0 \to J/\psi\phi$ and for the S-wave (dashed-dotted, magenta line).*

The solid, blue lines in Figures 4.1 and 4.2 represent the projections of the full PDF and are to be compared to the distributions of the data. In addition, the contributions of CP-even and CP-odd intermediate states to the PDF are shown separately. The long-dashed, red lines represent the sum of the $|A_0|^2$ and $|A_{\parallel}|^2$ terms in the PDF, which are CP even. The CP-odd $|A_{\perp}|^2$ and $|A_S|^2$ terms are shown as the short-dashed, green and dashed-dotted, magenta lines,

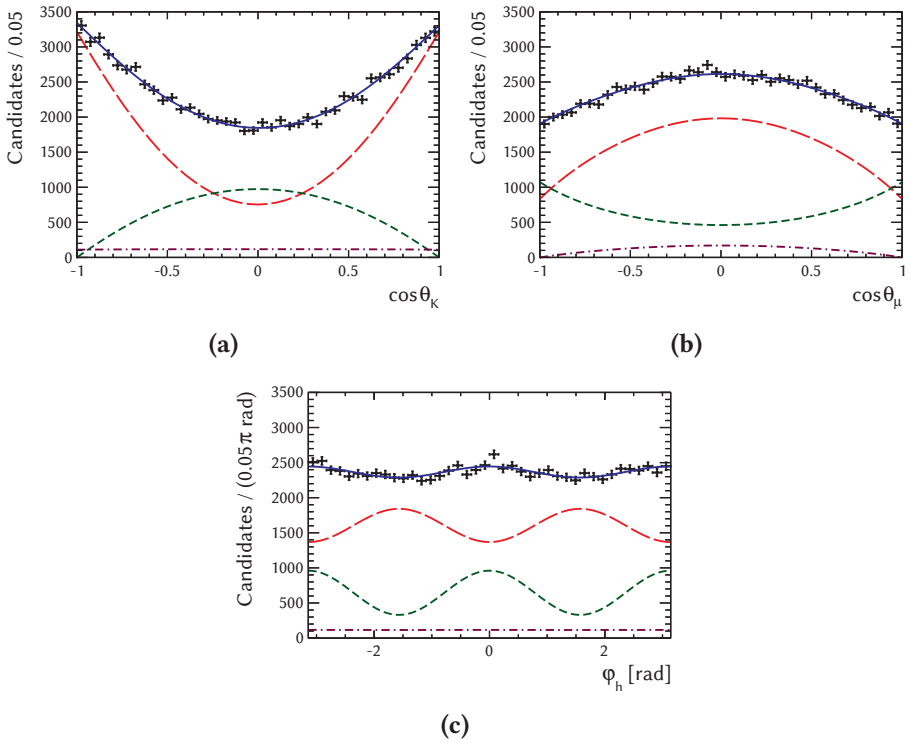


Figure 4.2: *Background-subtracted distribution of decays in decay angles (data points) and the corresponding one-dimensional projections of the PDF (solid, blue line). The distributions of* $\cos\theta_K$, $\cos\theta_\mu$, *and* φ_h *are shown in Figures (a), (b), and (c), respectively. Additional PDF projections are shown for the CP-even (long-dashed, red line) and CP-odd (short-dashed, green line) components of* $B_s^0 \to J/\psi\phi$ *and for the S-wave (dashed-dotted, magenta line).*

respectively. Contributions from interference terms are not shown. The discontinuities in the decay-time PDFs arise from statistical uncertainties in the interval coefficients of the acceptance function.

The shapes of the CP-even and CP-odd projections are clearly different, which enables a statistical separation. In $\cos\theta_K$ and $\cos\theta_\mu$ the difference between the even and odd $B_s^0 \to J/\psi\,\phi$ components is made by the $|A_0|^2$ and $|A_\perp|^2$ angular functions (see Table 2.2). Where the $|A_0|^2$ term has a $\cos^2\theta$ behaviour, the $|A_\perp|^2$ term has a $\sin^2\theta$ behaviour and vice versa. In φ_h the $|A_0|^2$ function is uniform, but the oscillatory behaviour for $|A_\parallel|^2$ is opposite to that of the $|A_\perp|^2$ function. The K^+K^- S-wave term is uniform in $\cos\theta_K$ and φ_h, but proportional to $\sin^2\theta_\mu$.

The largest contribution to the $B_s^0 \to J/\psi\,\phi$ decay rate comes from the CP-even $|A_0|^2$ and $|A_\parallel|^2$ components, which account for approximately 75%. Although this cannot be seen in these projection plots, $|A_\parallel|^2$ and $|A_\perp|^2$ are roughly equal. The S-wave accounts for approximately 5% of the total.

From the decay-time plot with logarithmic vertical scale it can be seen that the CP-even intermediate states decay faster than the CP-odd intermediate states. In case CP violation is small, the even states roughly correspond to the light CP eigenstate of the B_s^0–$\overline{B}_s^0$ system and the odd states to the heavy CP eigenstate (Equation 2.64). Under this assumption it can be inferred from the lifetimes of the states that the decay width of the light state is larger and hence that $\Delta\Gamma_s \equiv \Gamma_L - \Gamma_H$ is positive.

The plots in Figure 4.1 show the sum of the B_s^0 and $\overline{B}_s^0$ decay-time distributions, for which the decay-time oscillation vanishes. Since most of the sensitivity for the CP-violation parameters comes from the amplitudes of the oscillation terms, these projection plots contain little information on the CP-violation measurement. More information is contained in Figure 4.3, which shows the asymmetry between B_s^0 tags and $\overline{B}_s^0$ tags.

The B_s^0–$\overline{B}_s^0$-tag asymmetry is defined as

$$A_{\text{tag}} \equiv \frac{\#B_s^0 \text{ tags} - \#\overline{B}_s^0 \text{ tags}}{\#B_s^0 \text{ tags} + \#\overline{B}_s^0 \text{ tags}}$$

and is shown in intervals of decay-time for the background-subtracted data. The corresponding asymmetry in the PDFs for B_s^0 tags and $\overline{B}_s^0$ tags at each point in decay time is shown for the three parameterizations by the lines in the figure. The data asymmetry in a interval is predicted by the mean

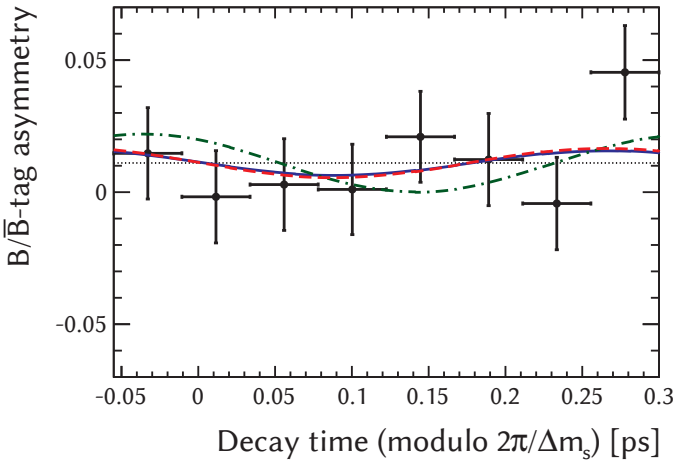


Figure 4.3: *Background-subtracted asymmetry in the numbers of decays with B_s^0 and with $\overline{B}_s^0$ flavour tags as a function of decay time (data points) and the corresponding asymmetries in the PDFs for the ϕ_s^i/C_s^i (solid, blue line), $\phi_s/|\lambda_s|$ (dashed-dotted, green line), and $\phi_s/|\lambda_s| = 1$ (dashed, red line) parameterizations. Data from the full decay-time range are mapped onto one period of the expected oscillation in the asymmetry. Decay candidates are weighted by the product of the corresponding estimated dilution factors from decay-time resolution and wrong tags to optimize the significance of the displayed asymmetry.*

of the PDF asymmetry in that interval. An oscillation can be observed in the asymmetry as a function of decay time provided that CP symmetry is violated (Equation 2.72) and that the dilution factors from flavour tagging and decay-time resolution are non-zero.

To increase the statistical significance of the displayed asymmetry in the data all oscillation periods are projected onto the one period of approximately 0.36 ps that is shown in Figure 4.3. This period starts at about −0.06 ps, onto which the lower edge of the decay-time range in the analysis of 0.3 ps is mapped.

The significance of the displayed oscillation is further enhanced by the use of flavour-tagging and decay-time resolution information. The amplitude of the oscillation in the plot is maximized with respect to its uncertainty by weighting the contribution of each decay candidate by the product of the corresponding dilution factors from tagging and resolution.

With the assumption $|\lambda_s| = 1$ the cosine component of the PDF vanishes. This can be seen in the figure, which shows a sine function for the $\phi_s/|\lambda_s| = 1$ parameterization (dashed, red line). The function has a negative slope at zero decay-time, which indicates a negative value of ϕ_s, assuming the CP-even components dominate the decay (see Equation 2.72). The function for the ϕ_s^i/C_s^i parameterization (solid, blue line) almost overlaps with the $\phi_s/|\lambda_s| = 1$ function. This indicates that the C_s^i parameter values are small or that they effectively cancel, resulting in a vanishing cosine component.

The oscillation for the $\phi_s/|\lambda_s|$ parameterization can be described as a negative sine function with a phase offset of about $0.055\,\text{ps} \cdot \Delta m_s \approx 0.31\,\pi$ rad (dashed-dotted, green line in Figure 4.3), implying that both the sine and the cosine coefficients in Equation 2.72 have non-zero values for this parameterization. The contribution of the cosine also makes the amplitude of the oscillation larger than the amplitudes for the other two parameterizations. With this phase offset the magnitude of the cosine coefficient is roughly a factor 1.5 larger than the magnitude of the sine coefficient. The latter coefficient is again negative, which implies a negative value for ϕ_s^{av}.

Despite the fact that the asymmetry plot shows an oscillation in the B_s^0 and $\overline{B}_s^0$ PDFs, the statistical uncertainties on the data points show that this oscillation is not significant. This indicates that that the estimates of the ϕ_s and C_s from the fit also do not differ from zero significantly.

Also the overall mean B_s^0–$\overline{B}_s^0$-tag asymmetry in the data, indicated by the dotted, black line, is not statistically significant. A non-zero mean could arise from production and tagging asymmetries. The PDFs for B_s^0 tags and $\overline{B}_s^0$ tags are weighted by the number of decays in the corresponding category, which makes the mean asymmetry in the PDF equal to the value in the data by construction.

4.1 Parameter Estimates

Parameter estimates from the three CP-violation parameterizations are listed in Tables 4.1, 4.2, and 4.3. Correlations between the main parameters of the ϕ_s^i/C_s^i parameterization are shown in Table 4.4.

The techniques discussed in Section 3.1 are applied to obtain the estimates of the parameter values, corresponding statistical uncertainties, and correlation coefficients. The uncertainty estimates are obtained from the one-sigma confidence interval of the profile likelihood, which yields an asymmetric un-

Table 4.1: *Parameter estimates with the ϕ_s^i/C_s^i model. See text for details.*

parameter	value	uncertainty	pull mean	pull width
ϕ_s^{av} [rad]	−0.047	0.051	−0.013 ± 0.010	0.970 ± 0.007
$\Delta\phi_s^{\parallel}$ [rad]	−0.019	0.043	−0.013 ± 0.009	0.915 ± 0.006
$\Delta\phi_s^{\perp\prime}$ [rad]	−0.003	0.029	+0.008 ± 0.009	0.875 ± 0.006
$\Delta\phi_s^{\text{S}}$ [rad]	+0.014	0.062	–	–
C_s^{av}	−0.006	0.039	+0.048 ± 0.010	1.004 ± 0.007
$\Delta C_s^{\parallel}$	−0.025	0.122	−0.011 ± 0.011	1.044 ± 0.007
$\Delta C_s^{\perp}$	+0.043	0.162	+0.017 ± 0.010	1.024 ± 0.008
C_s^{avS}	+0.060	0.032	–	–
Γ_s [ps^{-1}]	0.6591	0.0033	−0.015 ± 0.010	0.982 ± 0.007
$\Delta\Gamma_s$ [ps^{-1}]	+0.0784	0.0092	+0.051 ± 0.010	0.989 ± 0.007
Δm_s [ps^{-1}]	17.697	0.062	−0.005 ± 0.010	1.032 ± 0.008
$\lvert A_0^{\text{CP}}\rvert^2$	0.5236	0.0034	+0.016 ± 0.010	1.012 ± 0.007
$\lvert A_\perp^{\text{CP}}\rvert^2$	0.2513	0.0049	−0.135 ± 0.010	1.018 ± 0.008
$F_{\text{S1}}^{\text{CP}}$	0.424	0.054	–	–
$F_{\text{S2}}^{\text{CP}}$	0.057	0.018	–	–
$F_{\text{S3}}^{\text{CP}}$	0.009	+0.007 −0.005	–	–
$F_{\text{S4}}^{\text{CP}}$	0.009	+0.006 −0.005	–	–
$F_{\text{S5}}^{\text{CP}}$	0.048	0.015	–	–
$F_{\text{S6}}^{\text{CP}}$	0.191	0.026	–	–
$\delta_\parallel - \delta_0$ [rad]	+3.247	+0.104 −0.201	–	–
$\delta_\perp - \delta_0$ [rad]	+3.037	+0.160 −0.177	−0.021 ± 0.011	1.059 ± 0.007
$\delta_{\text{S1}} - \delta_{\perp 1}$ [rad]	$[+0.3, +2.6]$		–	–
$\delta_{\text{S2}} - \delta_{\perp 2}$ [rad]	$[+0.6, +2.7]$		–	–
$\delta_{\text{S3}} - \delta_{\perp 3}$ [rad]	$[+0.1, +2.7]$		–	–
$\delta_{\text{S4}} - \delta_{\perp 4}$ [rad]	$[-2.2, +0.1]$		–	–
$\delta_{\text{S5}} - \delta_{\perp 5}$ [rad]	$[-2.7, -0.2]$		–	–
$\delta_{\text{S6}} - \delta_{\perp 6}$ [rad]	$[-1.6, -0.5]$		–	–

Table 4.2: *Parameter estimates with the $\phi_s/|\lambda_s|$ model. See text for details.*

parameter	value	uncertainty	pull mean	pull width
ϕ_s [rad]	−0.057	0.050	+0.008 ± 0.010	0.980 ± 0.007
$\lvert\lambda_s\rvert$	0.9627	0.0188	−0.096 ± 0.011	1.052 ± 0.007
Γ_s [ps^{-1}]	0.6592	0.0033	+0.026 ± 0.010	0.990 ± 0.007
$\Delta\Gamma_s$ [ps^{-1}]	+0.0785	0.0092	+0.014 ± 0.010	0.991 ± 0.007
Δm_s [ps^{-1}]	17.723	0.057	+0.009 ± 0.010	1.020 ± 0.007
$\lvert A_0^{CP}\rvert^2$	0.5237	0.0034	−0.002 ± 0.010	1.012 ± 0.007
$\lvert A_\perp^{CP}\rvert^2$	0.2512	0.0049	−0.112 ± 0.010	1.015 ± 0.007
F_{S1}^{CP}	0.426	0.054	–	–
F_{S2}^{CP}	0.059	0.018	–	–
F_{S3}^{CP}	0.010	+0.007 −0.006	–	–
F_{S4}^{CP}	0.009	+0.006 −0.005	–	–
F_{S5}^{CP}	0.048	0.015	–	–
F_{S6}^{CP}	0.192	0.025	–	–
$\delta_\parallel - \delta_0$ [rad]	+3.257	+0.100 −0.172	–	–
$\delta_\perp - \delta_0$ [rad]	+3.099	+0.141 −0.151	+0.001 ± 0.011	1.075 ± 0.008
$\delta_{S1} - \delta_{\perp 1}$ [rad]	[+0.3, +2.6]		–	–
$\delta_{S2} - \delta_{\perp 2}$ [rad]	[+0.6, +2.7]		–	–
$\delta_{S3} - \delta_{\perp 3}$ [rad]	[+0.1, +2.7]		–	–
$\delta_{S4} - \delta_{\perp 4}$ [rad]	[−2.3, +0.1]		–	–
$\delta_{S5} - \delta_{\perp 5}$ [rad]	[−2.7, −0.2]		–	–
$\delta_{S6} - \delta_{\perp 6}$ [rad]	[−1.6, −0.6]		–	–

Table 4.3: *Parameter estimates with the $\phi_s/|\lambda_s| = 1$ model. See text for details.*

parameter	value	uncertainty	pull mean	pull width
ϕ_s [rad]	−0.056	0.049	−0.018 ± 0.010	0.973 ± 0.007
Γ_s [ps^{-1}]	0.6591	0.0033	+0.018 ± 0.010	0.995 ± 0.007
$\Delta\Gamma_s$ [ps^{-1}]	+0.0785	0.0091	+0.006 ± 0.010	0.991 ± 0.007
Δm_s [ps^{-1}]	17.697	0.060	−0.011 ± 0.010	1.025 ± 0.008
$\lvert A_0^{CP}\rvert^2$	0.5236	0.0034	+0.005 ± 0.010	1.012 ± 0.007
$\lvert A_\perp^{CP}\rvert^2$	0.2512	0.0049	−0.102 ± 0.010	0.997 ± 0.007
F_{S1}^{CP}	0.426	0.054	–	–
F_{S2}^{CP}	0.059	0.018	–	–
F_{S3}^{CP}	0.010	+0.007 −0.006	–	–
F_{S4}^{CP}	0.008	+0.006 −0.005	–	–
F_{S5}^{CP}	0.045	0.016	–	–
F_{S6}^{CP}	0.192	0.025	–	–
$\delta_\parallel - \delta_0$ [rad]	+3.264	+0.099 −0.180	–	–
$\delta_\perp - \delta_0$ [rad]	+3.043	+0.158 −0.166	−0.026 ± 0.011	1.058 ± 0.008
$\delta_{S1} - \delta_{\perp 1}$ [rad]	$[+0.3, +2.6]$		–	–
$\delta_{S2} - \delta_{\perp 2}$ [rad]	$[+0.5, +2.7]$		–	–
$\delta_{S3} - \delta_{\perp 3}$ [rad]	$[+0.1, +2.8]$		–	–
$\delta_{S4} - \delta_{\perp 4}$ [rad]	$[-2.6, +0.1]$		–	–
$\delta_{S5} - \delta_{\perp 5}$ [rad]	$[-2.8, -0.2]$		–	–
$\delta_{S6} - \delta_{\perp 6}$ [rad]	$[-1.6, -0.6]$		–	–

Table 4.4: *Parameter correlation matrix with the ϕ_s^i/C_s^i model. Correlations with an absolute value above 30% are highlighted.*

	ϕ_s^{av}	$\Delta\phi_s^{\parallel}$	$\Delta\phi_s^{\perp\prime}$	$\Delta\phi_s^{S}$	C_s^{av}	$\Delta C_s^{\parallel}$	$\Delta C_s^{\perp}$	C_s^{avS}	Γ_s	$\Delta\Gamma_s$	Δm_s	$\lvert A_0^{CP}\rvert^2$	$\lvert A_\perp^{CP}\rvert^2$	$\delta_\parallel - \delta_0$	$\delta_\perp - \delta_0$
ϕ_s^{av}	1	+0.09	–	+0.16	−0.15	−0.08	−0.09	–	–	−0.06	−0.11	–	–	–	−0.07
$\Delta\phi_s^{\parallel}$		1	−0.09	+0.12	−0.20	−0.27	−0.07	–	–	–	−0.14	–	–	+0.13	−0.06
$\Delta\phi_s^{\perp\prime}$			1	–	**+0.41**	–	**+0.36**	–	–	−0.06	–	–	–	–	+0.06
$\Delta\phi_s^{S}$				1	−0.05	−0.12	+0.05	−0.50	–	–	−0.24	–	–	−0.09	−0.29
C_s^{av}					1	–	–	−0.09	–	–	+0.23	–	–	+0.10	+0.19
$\Delta C_s^{\parallel}$						1	−0.24	–	–	–	+0.12	–	–	−0.09	–
$\Delta C_s^{\perp}$							1	–	–	–	–	–	−0.05	+0.09	–
C_s^{avS}								1	–	–	–	–	–	–	+0.06
Γ_s									1	**−0.38**	–	−0.26	**+0.33**	−0.06	–
$\Delta\Gamma_s$										1	–	**+0.65**	**−0.69**	–	–
Δm_s											1	–	–	+0.08	**+0.72**
$\lvert A_0^{CP}\rvert^2$												1	**−0.59**	–	–
$\lvert A_\perp^{CP}\rvert^2$													1	−0.30	−0.12
$\delta_\parallel - \delta_0$														1	**+0.41**
$\delta_\perp - \delta_0$															1

certainty for some parameters. For the S-wave phase differences the three-sigma confidence interval is shown instead of a point estimate.

To indicate the meaning of the parameter uncertainties, the mean and width of the parameter pull distribution in ten thousand pseudo experiments are shown in the last two columns of the parameter tables. The pull is defined as the difference between estimate of the parameter value and the true value, divided by the estimated uncertainty.

The pull parameters are only shown for parameters with a pull distribution that resembles a Gaussian shape. In the limit of a large number of decay candidates, the parameter distribution is expected to be Gaussian and the pull is expected to be distributed according to the standard normal distribution, which has a zero mean and a width equal to one. A non-zero mean indicates a bias in the estimate of the parameter value. If the uncertainty is overestimated or underestimated, the pull width will be smaller or greater than one, respectively.

The pseudo experiments were generated with the nominal parameter estimates, except for the difference between the S-wave phase and the $B_s^0 \to J/\psi\,\phi$ phase in the first K^+K^--mass interval. This parameter was generated at $\delta_{S1} - \delta_{\perp 1} = 2.2$ rad to get the expected ordering of phases differences across the K^+K^--mass intervals (see also Section 3.5 and Figures 4.8c and 4.10).

Even though the pull distributions have means close to zero and widths close to one, there are some small deviations. These deviations may originate from the fact that the parameter distributions only approximately have Gaussian shapes, but also from the inherent bias in parameter estimates from a maximum-likelihood fit or the distortion of the likelihood by event weights. Since these effects are small, they will not be investigated further and estimates of parameter values and uncertainties will not be corrected.

Estimates of all CP-violation parameters are both compatible with the scenario of no CP violation ($\phi_s^i = 0$, $C_s^i = 0$) and with the naive Standard Model prediction ($\phi_s^{\text{av}} = -2\beta_s \approx -0.037$ rad, $\Delta\phi_s^i = 0$, $C_s^i = 0$). Not considering the S-wave and assuming Gaussian distributions for the six parameters that describe CP-violation in the $B_s^0 \to J/\psi\,\phi$ process, the probability to observe values in a given range under these two hypotheses can be calculated. Under the hypothesis of no CP violation the probability to observe the measured values or values further from the hypothesis values is 97%. Under the hypothesis of naive Standard Model CP violation, this probability is even 99.9%.

On the other hand, using the naive Standard Model prediction for the

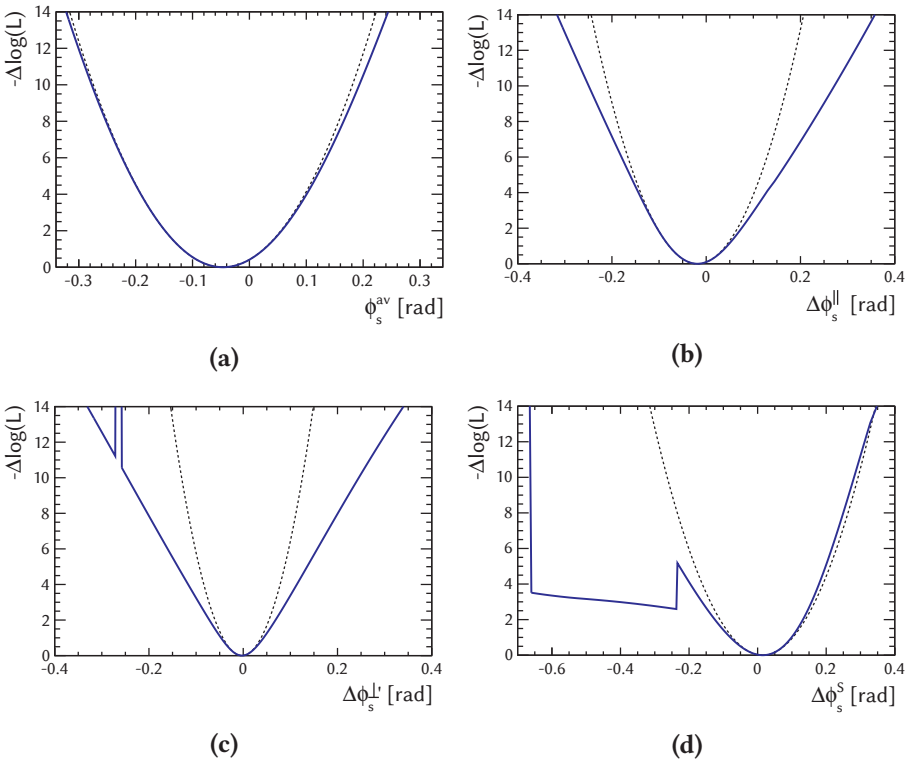


Figure 4.4: *Log-likelihood scans of the CP-violating phases (a) ϕ_s^{av}, (b) $\Delta\phi_s^{\parallel}$, (c) $\Delta\phi_s^{\perp'}$, and (d) $\Delta\phi_s^{S}$ in the ϕ_s^i/C_s^i model. See text for details.*

two S-wave parameters ($\Delta\phi_s^S = 0$ and $C_s^{avS} = 0$), the probability to observe the measured values of $\Delta\phi_s^S$ and C_s^{avS} or values further from zero is only 2%. This low probability is driven by the two-sigma deviation from zero of C_s^{avS} and the fact that the deviations in $\Delta\phi_s^S$ and C_s^{avS} are both positive, while there is a negative correlation between the parameter estimates.

As expected, the value of ϕ_s in the $\phi_s/|\lambda_s|$ and $\phi_s/|\lambda_s| = 1$ parameterizations (-0.06 ± 0.05 rad) are similar to the value of $\phi_s^{av} = -0.05 \pm 0.05$ rad in the ϕ_s^i/C_s^i model. The parameter C_s, however, is a combination of the parameters C_s^{av} and C_s^{avS}. Its estimated value in the $\phi_s/|\lambda_s|$ parameterization is 0.038 ± 0.019, which makes the estimate of the corresponding $|\lambda_s| \approx 1 - C_s$ equal to 0.963 ± 0.019. Since the estimates of ϕ_s and $|\lambda_s|$ are almost uncorrelated, the effect of assuming $|\lambda_s| = 1$ on the ϕ_s estimate is negligible.

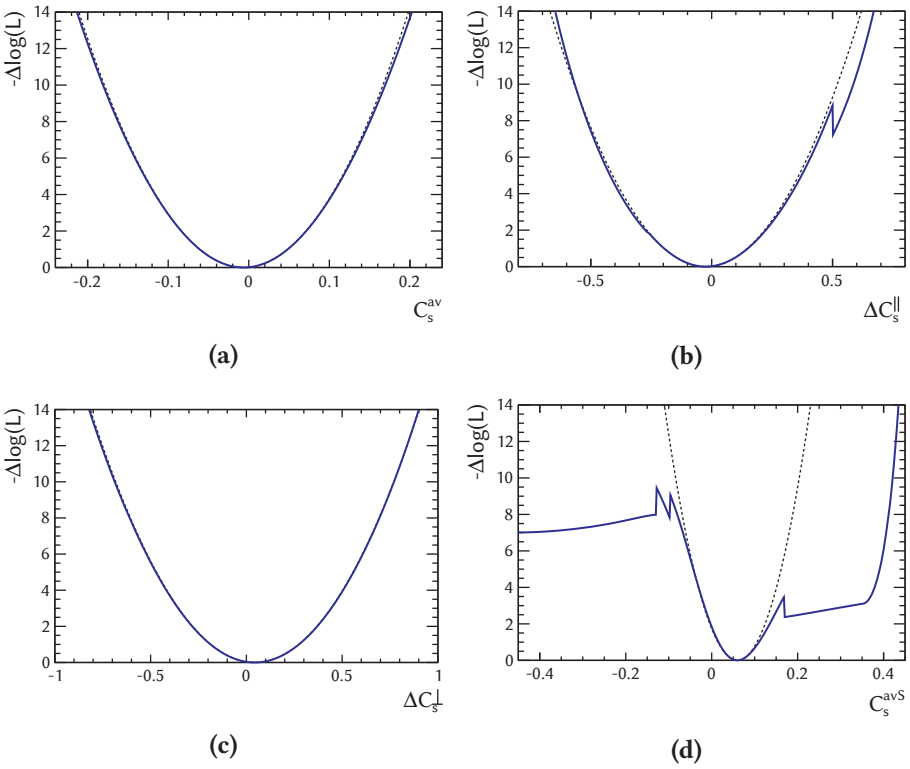


Figure 4.5: *Log-likelihood scans of the asymmetries from CP violation in mixing and in decay (a) C_s^{av}, (b) $\Delta C_s^{\parallel}$, (c) $\Delta C_s^{\perp}$, and (d) C_s^{avS} in the ϕ_s^i/C_s^i model. See text for details.*

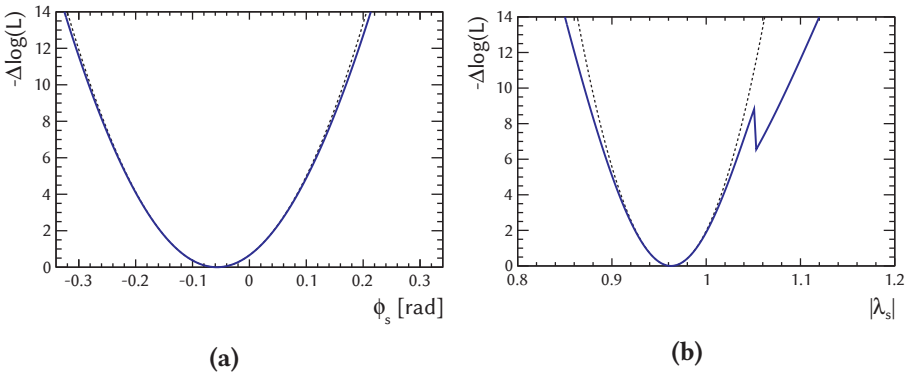


Figure 4.6: *Log-likelihood scans of (a) ϕ_s and (b) $|\lambda_s|$ in the $\phi_s/|\lambda_s|$ model. See text for details.*

The assumptions of a parabolic NLL and a Gaussian parameter distribution are only approximately correct. Figures 4.4, 4.5, and 4.6 show the profiled NLL for the ϕ_s^i, C_s^i, and $\phi_s/|\lambda_s|$ parameters, respectively. The solid, blue line represents the difference in profiled NLL with respect to its minimum value. The dotted, black line represents the parabolic approximation based on the inverse of the second derivative in the minimum. In the Gaussian case the two lines would coincide.

Although none of the NLLs shown exactly has a parabolic shape, for most parameters this is a sufficient assumption up to a difference in NLL of 4.5, which corresponds to three standard deviations in the Gaussian case. Clear exceptions are $\Delta\phi_s^S$ and C_s^{avS}, for which the NLL difference is smaller than 4.5 in a large region. The sudden jumps in NLL for these (and also other) parameters occur at points where a local minimum in one of the $\delta_S - \delta_\perp$ differences becomes the global minimum.

For the parameters with an approximately parabolic NLL also the distribution in pseudo experiments will resemble a Gaussian shape, since the vast majority of experiments will be generated within three standard deviations from the mean. In these cases the pull distribution for the parameter are summarized by the mean and width listed in Tables 4.1 to 4.3.

In the ϕ_s^i parameter estimates there are no significant biases, but widths of the pull distributions indicate that the uncertainties are somewhat overestimated. For the ϕ_s^{av} and ϕ_s parameters the widths are different from one by only a few per cent, but for the $\Delta\phi_s^{||}$ and $\Delta\phi_s^{\perp\prime}$ parameters this is around ten

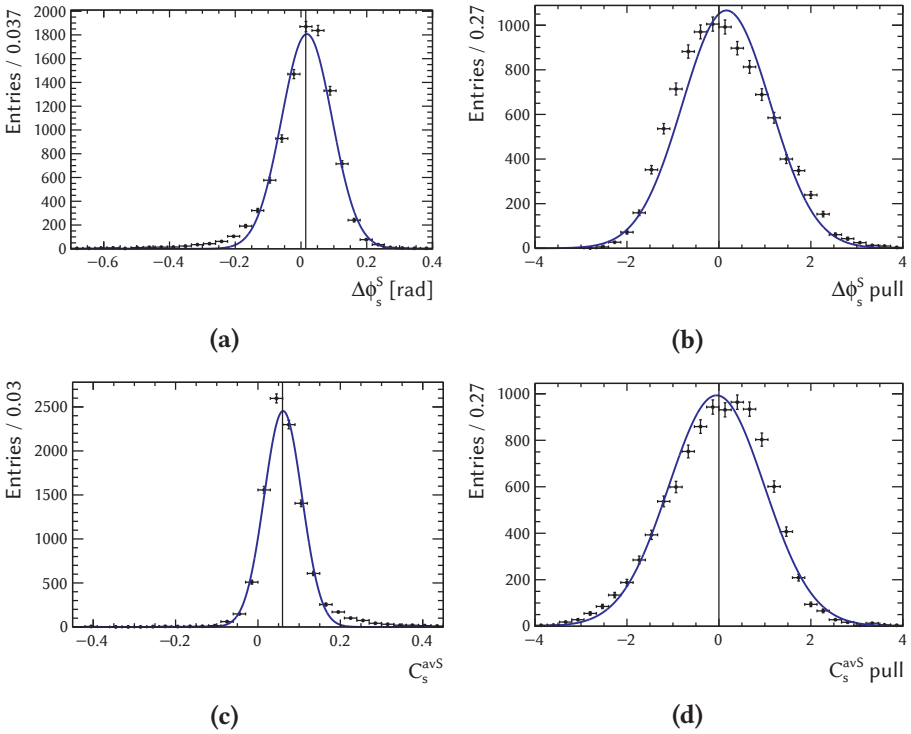


Figure 4.7: *Distributions and Gaussian-fit results of S-wave CP-violation parameters in pseudo experiments: (a) $\Delta\phi_s^S$ ($\mu = 0.018$; $\sigma = 0.077$), (b) $\Delta\phi_s^S$ pull ($\mu = 0.16$; $\sigma = 0.96$), (c) C_s^{avS} ($\mu = 0.061$; $\sigma = 0.046$), and (d) C_s^{avS} pull ($\mu = -0.050$; $\sigma = 1.04$).*

per cent.

The uncertainties of the $\Delta C_s^{\parallel}$ and $\Delta C_s^{\perp}$ parameters are overestimated by a few per cent. While the width of the C_s^{av} pull distribution is compatible with one, the estimate of the value of this parameter is biased by approximately five per cent of the statistical uncertainty. This bias is also present for the $|\lambda_s|$ parameter, for which it is approximately ten per cent.

The parameter and pull distributions of the S-wave CP-violation parameters, which are indeed not Gaussian, are shown in Figure 4.7. The solid, blue lines in this plots represent Gaussian PDFs that were fitted to the distributions. The vertical lines in the parameter plots represent the input value in the pseudo experiments. In the regions where the NLL plots show unexpectedly

low values, the corresponding parameter distributions have tails that would not have been present for a Gaussian shape. The tails do not directly appear in the pull plots, indicating that the NLL shape, and therefore the uncertainty estimate, varies with the position of the NLL minimum.

As discussed in Section 1.3.2, predictions for the decay-width parameters are $\Gamma_s \approx \Gamma_d = 0.6583 \pm 0.0030\,\text{ps}^{-1}$ and $\Delta\Gamma_s = 0.087 \pm 0.021\,\text{ps}^{-1}$. The estimates of these parameters are almost identical for the three different CP-violation parameterizations and statistically compatible with the predictions.

Also the parameter Δm_s can be estimated in the $B_s^0 \to J/\psi K^+K^-$ measurement, because of the interference between the different angular momentum states (see Section 2.6, in particular Table 2.3). The values obtained for Δm_s are compatible with the most precise measurement of this parameter, $\Delta m_s = 17.768 \pm 0.024\,\text{ps}^{-1}$ [58].

The three estimates of the value and uncertainty of Δm_s are different due to correlations with the $C_s^{\text{av}}/|\lambda_s|$ and $\delta_\perp - \delta_0$ parameters. For the ϕ_s^i/C_s^i and $\phi_s/|\lambda_s| = 1$ parameterizations the estimates of the values of these parameters are very similar. Fixing the value of $|\lambda_s|$ does make the uncertainties with the $\phi_s/|\lambda_s| = 1$ parameterization smaller, because of the correlations with this parameter. With the $\phi_s/|\lambda_s|$ model the parameter values change. Because the shape of the NLL depends on the parameter values, the uncertainty estimates also change and become even smaller for Δm_s and $\delta_\perp - \delta_0$.

The means and widths of the lifetime pull distributions for the $\phi_s/|\lambda_s|(=1)$ models are as expected. In the ϕ_s^i/C_s^i model the $\Delta\Gamma_s$ estimate develops a small bias of approximately five per cent of the statistical uncertainty.

Estimates for the transversity amplitudes yield fractions of 52% for the longitudinal $B_s^0 \to J/\psi\,\phi$ polarization ($|A_0^{\text{CP}}|^2$) and 25% for the perpendicular polarization ($|A_\perp^{\text{CP}}|^2$), which leaves 23% for the parallel polarization ($|A_\parallel^{\text{CP}}|^2$). The fraction of K^+K^- S-wave varies from approximately 40% in the first K^+K^--mass interval to approximately 1% in the two central intervals. The total S-wave fraction is approximately 4%.

The NLLs of $|A_0^{\text{CP}}|^2$ and $|A_\perp^{\text{CP}}|^2$ are parabolic, but the estimate of $|A_\perp^{\text{CP}}|^2$ is biased by about 14% of its statistical uncertainty. The NLLs and distributions of the S-wave fractions are asymmetric, because these parameters cannot assume negative values. At one sigma this only results in slightly asymmetric uncertainty estimates in the two central K^+K^--mass intervals, where the fractions are closest to zero.

As explained in Section 2.6.1, there are approximate symmetries in the

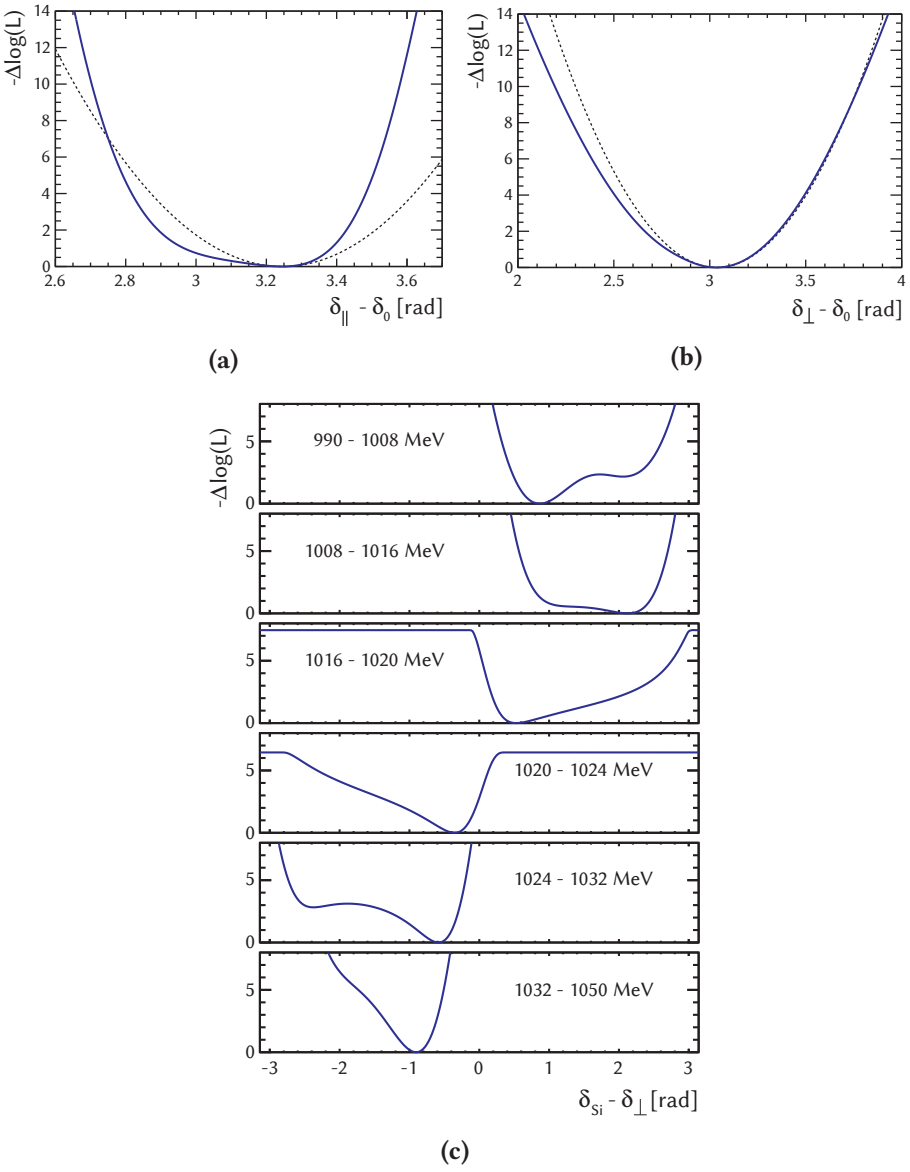
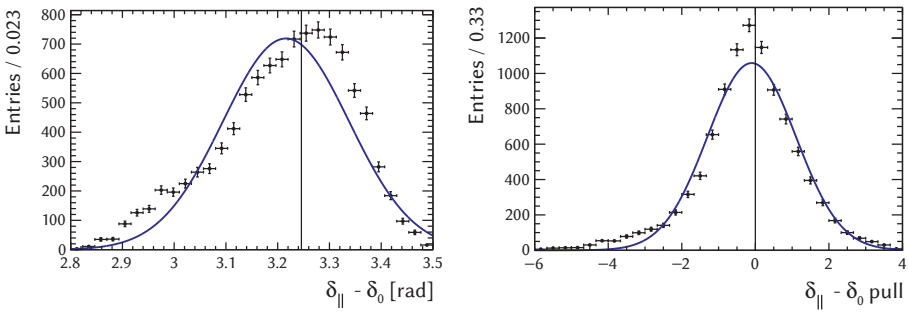


Figure 4.8: *Log-likelihood scans of the phases of the transversity amplitudes (likelihood with polarization-dependent CP violation): (a) $\delta_{\parallel} - \delta_0$, (b) $\delta_{\perp} - \delta_0$, (c) $\delta_{Si} - \delta_{\perp i}$, where the K^+K^--mass interval is indicated with the corresponding K^+K^--mass range. See text for details.*



(a) (b)

Figure 4.9: *Distributions and Gaussian-fit results of the phase of the parallel transversity amplitude in pseudo experiments (maximum of the likelihood with polarization-dependent CP violation): (a) $\delta_{||} - \delta_0$ ($\mu = 3.22$; $\sigma = 0.12$), (b) $\delta_{||} - \delta_0$ pull ($\mu = -0.10$; $\sigma = 1.20$).*

transversity-amplitude phase parameters, which lead to NLLs with a double-minimum structure. The NLL curves for the phase parameters are shown in Figure 4.8. Tables 4.1 to 4.3 list one-sigma asymmetric uncertainties for the $\delta_{||} - \delta_0$ and $\delta_\perp - \delta_0$ parameters. Three-sigma intervals are shown for the $\delta_S - \delta_\perp$ parameters in the tables.

The minimum of the $\delta_{||} - \delta_0$ NLL is located at approximately $\pi + 0.1$. At a value of $\pi - 0.1$ the cosine of $\delta_{||} - \delta_0$, which is determined more precisely than the sine, has the same value. As a result, the NLL is also pulled down at this point, yielding a broad shape around $\delta_{||} - \delta_0 = \pi$. The $\delta_S - \delta_\perp$ NLLs show similar behaviour with a minimum at $(\delta_S - \delta_\perp)_{\text{min}}$ and a second "minimum" at $\pi - (\delta_S - \delta_\perp)_{\text{min}}$ (or, equivalently, $-\pi - (\delta_S - \delta_\perp)_{\text{min}}$), where the sine of $\delta_S - \delta_\perp$ is equal. For two of the K^+K^--mass intervals this results in an actual second minimum.

Despite the broad NLL shapes, a trend of decreasing $\delta_S - \delta_\perp$ with increasing K^+K^- mass is clearly visible in Figure 4.8c. This confirms the assumption of $\phi_s \approx 0$ and $\Delta\Gamma_s > 0$ in favour of $\phi_s \approx \pi$ and $\Delta\Gamma_s < 0$, as discussed in Section 2.6.2.

The S-wave contribution is smallest in the two central K^+K^--mass intervals, which results in the least significant measurement of the parameters related to the S-wave. Above an NLL difference of approximately seven the interval for the $\delta_S - \delta_\perp$ phase differences in these intervals spans the full range

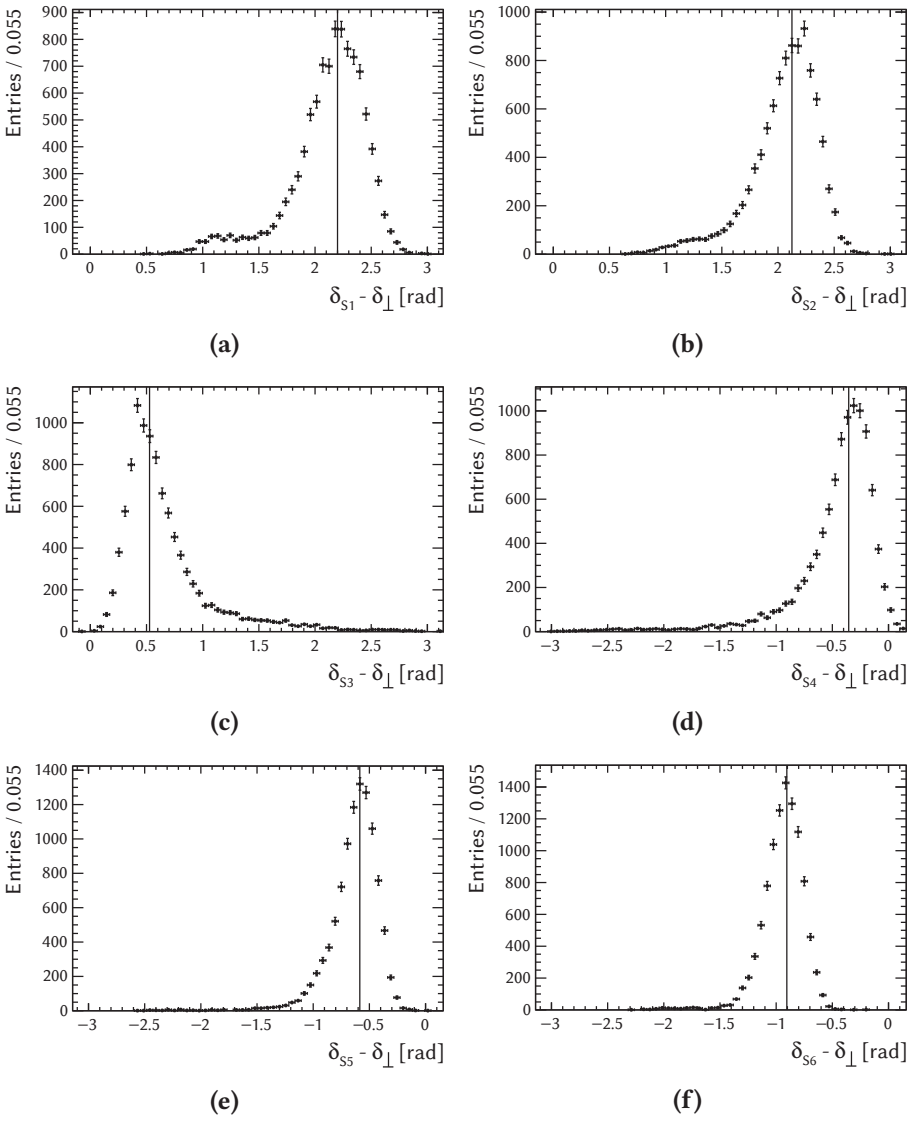


Figure 4.10: *Distributions of the phases of the S-wave amplitudes in pseudo experiments in the six intervals in K^+K^--mass (maximum of the likelihood with polarization-dependent CP violation): (a) 990–1008 MeV/c^2, (b) 1008–1016 MeV/c^2, (c) 1016–1020 MeV/c^2, (d) 1020–1024 MeV/c^2, (c) 1024–1032 MeV/c^2, and (f) 1032–1050 MeV/c^2. Notice the change in horizontal scale between (a), (b), (c) and (d), (e), (f).*

of $[-\pi, +\pi]$. At the plateaus in the NLL plots the S-wave fractions are equal to zero and changes in the phase differences have no effect. The plots in the other four intervals show similar behaviour at much larger NLL difference.

Figures 4.9 and 4.10 show the distributions of the $\delta_{\parallel} - \delta_0$ and $\delta_S - \delta_{\perp}$ values, respectively. Also the pull distribution for $\delta_{\parallel} - \delta_0$ is shown, which is more symmetric than the distribution of parameter values. Also in these distributions of the phase differences the effect of the double-minimum structure in the NLL is visible.

4.2 Alternative Parameterizations

To justify the modelling choices made in Chapter 3, the time and angular fit is repeated with alternative parameterizations. The effects of an external constraint on the estimate of Δm_s, a narrower K^+K^--mass window, and flavour-tagging categories are studied. It is expected that the fits with alternative models yield results that are comparable to or less precise than the results from the nominal fit.

The parameters of the decay-time trigger-acceptance function are fixed to their nominal vales in all alternative fits, since this considerably reduces the time required to perform the fits. As a result, the estimated statistical uncertainty on Γ_s decreases from $0.0033\,\text{ps}^{-1}$ to $0.0031\,\text{ps}^{-1}$.

4.2.1 Constrained Mass-Difference Parameter

Currently, the most precise estimate of Δm_s is given by $17.768 \pm 0.024\,\text{ps}^{-1}$ and comes from an LHCb analysis of $B_s^0 \to D_s^- \pi^+$ decays [58]. This estimate can be combined with the estimate from the $B_s^0 \to J/\psi K^+K^-$ data, as was done for previous measurements [3, 70]. The external measurement is represented by an additional parabolic term in the NLL, using the same procedure as for the external measurements of the reconstruction-acceptance parameters (Section 3.3.2) and the tagging calibration parameters (Section 3.6.2).

The value of the external Δm_s constraint is about one standard deviation larger than the nominal-fit estimate ($\Delta m_s = 17.697 \pm 0.062\,\text{ps}^{-1}$) and its uncertainty is a factor 2.6 smaller. As a result, the estimates of the parameters that are correlated to Δm_s are expected to deviate from the nominal result. Especially the estimate of $\delta_{\perp} - \delta_0$, for which the correlation with Δm_s is 70%, will be significantly affected.

Table 4.5 lists the results of the fit with Δm_s constrained. Estimates of the parameter values are given in the second column. The third column contains the differences of these estimates with respect to the values from the nominal fit. The uncertainty estimates are given in the fourth column and the relative change in the uncertainties in the fifth column.

Table 4.5: *Results of a fit with the estimate of Δm_s constrained to an external measurement (Δm_s = 17.768 ± 0.024 ps^{-1} [58]).*

parameter	estimate	(difference)	uncertainty	(rel. difference)
ϕ_s^{av} [rad]	−0.051	(−0.004)	0.050	(−0.02)
$\Delta\phi_s^{\parallel}$ [rad]	−0.023	(−0.004)	0.041	(−0.04)
$\Delta\phi_s^{\perp\prime}$ [rad]	−0.003	(−0.001)	0.024	(−0.15)
$\Delta\phi_s^{S}$ [rad]	+0.000	(−0.014)	0.064	(+0.03)
C_s^{av}	−0.004	(+0.010)	0.038	(−0.01)
$\Delta C_s^{\parallel}$	−0.010	(+0.015)	0.122	(–)
$\Delta C_s^{\perp}$	+0.035	(−0.009)	0.164	(+0.01)
C_s^{avS}	+0.060	(–)	0.030	(−0.05)
Γ_s [ps^{-1}]	0.6591	(–)	0.0031	(–)
$\Delta\Gamma_s$ [ps^{-1}]	+0.0783	(−0.0001)	0.0092	(–)
Δm_s [ps^{-1}]	17.758	(+0.062)	0.022	(−0.64)
$\lvert A_0^{CP}\rvert^2$	0.5236	(−0.0001)	0.0034	(–)
$\lvert A_\perp^{CP}\rvert^2$	0.2512	(−0.0001)	0.0049	(–)
$\delta_\parallel - \delta_0$ [rad]	+3.255	(+0.009)	+0.100 −0.177	(−0.04 −0.12)
$\delta_\perp - \delta_0$ [rad]	+3.154	(+0.118)	+0.120 −0.125	(−0.25 −0.29)

The estimates of the value and uncertainty of Δm_s agree with what is expected from a combination of the internal and external estimates. Also the value and uncertainty of $\delta_\perp - \delta_0$ change as expected. The decrease of the uncertainty of the $\Delta\phi_s^{\perp\prime}$ estimate might be somewhat surprising, since there is no correlation between $\Delta\phi_s^{\perp\prime}$ and Δm_s given in Table 4.4. However, the likelihood shape in $\Delta\phi_s^{\perp\prime}$ (Figure 4.4c) is not Gaussian, so the linear covariance matrix is not sufficient to predict the changes for this parameter. The similarity of these results to the nominal results justifies omission of the Δm_s constraint in the fit, although this choice may have to be re-evaluated for future analyses, depending on the Δm_s correlations with a larger dataset.

4.2.2 Narrow K^+K^--Mass Window

Since the main interest of this measurement is CP violation in the $B_s^0 \to J/\psi\,\phi$ decay, one could argue to limit the range of K^+K^- masses to the region where the $\phi(1020)$ significantly contributes. This is a region of approximately 24 MeV/c^2 around the centre of the ϕ peak at 1020 MeV/c^2 (see also Figure 3.9f).

Results of a fit in a K^+K^--mass window of [1008, 1032] MeV/c^2 are shown in Table 4.6. The window was split up into the four intervals of the nominal analysis. The structure of the table is the same as for Table 4.5.

The uncertainties of the parameter estimates are larger than the uncertainties obtained with the nominal fit, because the information on the K^+K^- S-wave in the [990, 1008] MeV/c^2 and [1032, 1050] MeV/c^2 intervals is lost. Parameters which are (partially) determined with the S-wave (interference) terms are affected. From these results it can be seen that the two outer K^+K^--mass intervals do add valuable information for the CP-violation measurement.

Table 4.6: *Results of a fit in the K^+K^--mass window of [1008, 1032] MeV/c^2 with four intervals.*

parameter	estimate	(difference)	uncertainty	(rel. difference)
ϕ_s^{av} [rad]	−0.039	(+0.007)	0.053	(+0.05)
$\Delta\phi_s^{\parallel}$ [rad]	−0.044	(−0.025)	+0.053 −0.058	(+0.24 +0.36)
$\Delta\phi_s^{\perp\prime}$ [rad]	−0.024	(−0.021)	+0.040 −0.045	(+0.40 +0.56)
$\Delta\phi_s^{\mathrm{S}}$ [rad]	+0.033	(+0.018)	+0.098 −0.117	(+0.57 +0.89)
C_s^{av}	−0.018	(−0.011)	0.039	(–)
$\Delta C_s^{\parallel}$	−0.104	(+0.129)	+0.150 −0.140	(+0.23 +0.15)
$\Delta C_s^{\perp}$	−0.013	(−0.056)	0.166	(+0.02)
C_s^{avS}	+0.094	(+0.034)	+0.806 −0.063	(+24 +0.95)
Γ_s [ps^{-1}]	0.6572	(−0.0019)	0.0032	(+0.04)
$\Delta\Gamma_s$ [ps^{-1}]	+0.0821	(+0.0036)	0.0094	(+0.02)
Δm_s [ps^{-1}]	17.665	(−0.031)	0.073	(+0.17)
$\lvert A_0^{\mathrm{CP}}\rvert^2$	0.5252	(+0.0016)	0.0035	(+0.02)
$\lvert A_\perp^{\mathrm{CP}}\rvert^2$	0.2496	(−0.0016)	0.0051	(+0.04)
$\delta_\parallel - \delta_0$ [rad]	+2.980	(−0.266)	+0.333 −0.108	(+2.2 −0.46)
$\delta_\perp - \delta_0$ [rad]	+2.794	(−0.242)	+0.282 −0.195	(+0.76 +0.10)

4.2.3 Flavour-Tagging Categories

The flavour-tagging scheme with an estimated wrong-tag probability for each decay candidate is expected to give more precise results than a scheme with an average wrong-tag probability. To test this assumption the data are fitted using three alternative flavour-tagging parameterizations.

Table 4.7 shows the results of a fit in which no flavour-tagging information is used. Neglecting $\mathrm{B_s^0}$–$\overline{\mathrm{B}}{}_\mathrm{s}^0$ nuisance asymmetries, the $\cos(\Delta m_\mathrm{s}\, t)$ and $\sin(\Delta m_\mathrm{s}\, t)$ terms vanish from the PDF used in this untagged fit and only the $\cosh\left(\frac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right)$ terms remain (see Section 3.6). As a result, there is no sensitivity for the parameter Δm_s in this fit, but also the uncertainties in the estimates of other parameters increase significantly.

Table 2.3 shows that with small CP violation $\delta_\parallel - \delta_0$ appears in the coefficients of the $\cosh\left(\frac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right)$ terms, but $\delta_\perp - \delta_0$ does not. The latter is fixed to π in the untagged fit, because its uncertainty becomes too large. In the nominal fit the main sensitivity for $\phi_\mathrm{s}^\mathrm{av}$ and C_s^av comes from the $|A_i|^2$ oscillation terms. In the untagged fit the information has to come from interference terms and, at second order, from the $|A_i|^2\, \cosh\left(\frac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right)$ and $\sinh\left(\frac{1}{2}\Delta\Gamma_\mathrm{s}\, t\right)$ terms. As a result, $\Delta C_\mathrm{s}^\perp$ also becomes too uncertain and is fixed to zero.

Tables 4.9 and 4.10 list the results of fits in which the data sample is split into tagging categories and separate tagging-calibration parameters are used for each category. The split is based on the estimated wrong-tag probabilities for OS and SS tags. The categories are given in Table 4.8.

For the results in Table 4.9 the tagging-dilution factors for the categories are varied independently. Asymmetries in the tagging calibration are neglected, because the information in the data is insufficient to determine the corresponding parameters. The resulting dilution factors are also shown in Table 4.8.

The third column in Table 4.8 shows the mean of the estimated wrong-tag probability in each category. These numbers can be used to calculate the tagging-calibration parameters with the linear model that is used in the nominal fit for each event. The parameter estimates in Table 4.10 are the result of a fit with this model.

Uncertainties in the parameter estimates with tagging categories are comparable to the uncertainties in the nominal fit, although they are still larger. The parameter estimates are more precise with the linear calibration model and also with event-by-event tagging, as expected.

Table 4.7: *Results of a fit in which no flavour-tagging information is used.*

parameter	estimate	(difference)	uncertainty	(rel. difference)
ϕ_s^{av} [rad]	−0.336	(−0.290)	0.145	(+1.9)
$\Delta\phi_s^{\parallel}$ [rad]	−0.099	(−0.080)	+0.103 −0.194	(+1.4 +3.6)
$\Delta\phi_s^{\perp\prime}$ [rad]	+0.005	(+0.007)	+0.108 −0.068	(+2.8 +1.4)
$\Delta\phi_s^{S}$ [rad]	−0.038	(−0.053)	+0.069 −0.573	(+0.10 +8.2)
C_s^{av}	−0.282	(−0.275)	0.271	(+6.0)
$\Delta C_s^{\parallel}$	−0.054	(−0.030)	0.128	(+0.05)
$\Delta C_s^{\perp}$	0	(−0.044)	–	(–)
C_s^{avS}	+0.071	(+0.011)	+0.190 −0.033	(+4.9 +0.04)
Γ_s [ps^{-1}]	0.6606	(+0.0015)	+0.0038 −0.0034	(+0.24 +0.08)
$\Delta\Gamma_s$ [ps^{-1}]	+0.0849	(+0.0064)	+0.0120 −0.0105	(+0.31 +0.14)
Δm_s [ps^{-1}]	–	(–)	–	(–)
$\lvert A_0^{CP}\rvert^2$	0.5242	(+0.0006)	0.0035	(+0.01)
$\lvert A_\perp^{CP}\rvert^2$	0.2514	(+0.0001)	0.0049	(–)
$\delta_\parallel - \delta_0$ [rad]	+2.943	(−0.303)	+0.260 −0.171	(+1.5 −0.15)
$\delta_\perp - \delta_0$ [rad]	π	(+0.104)	–	(–)

Table 4.8: *Tagging categories and estimates of the corresponding dilution factors from the fit without external calibration constraints (Table 4.9).*

OS index	η^o range	mean η^o	dilution estimate
0 (untagged)	$\eta^o = 0.5$	0.5	0
1	$0.34 \le \eta^o < 0.50$	0.43	0.06 ± 0.04
2	$0.22 \le \eta^o < 0.34$	0.29	0.41 ± 0.08
3	$\eta^o < 0.22$	0.18	0.74 ± 0.14

SS index	η^s range	mean η^s	dilution estimate
0 (untagged)	$\eta^s = 0.5$	0.5	0
1	$0.40 \le \eta^s < 0.50$	0.47	0.07 ± 0.03
2	$0.30 \le \eta^s < 0.40$	0.36	0.17 ± 0.07
3	$\eta^s < 0.30$	0.26	0.43 ± 0.12

Table 4.9: *Results of a fit with four tagging categories for OS tags and four tagging categories for SS tags. Tagging-dilution factors are varied independently for the categories, without external calibration constraints.*

parameter	estimate (difference)		uncertainty (rel. difference)	
ϕ_s^{av} [rad]	−0.061	(−0.014)	0.059	(+0.16)
$\Delta\phi_s^{\parallel}$ [rad]	−0.025	(−0.007)	0.043	(+0.01)
$\Delta\phi_s^{\perp\prime}$ [rad]	+0.003	(+0.005)	+0.036 −0.031	(+0.27 +0.08)
$\Delta\phi_s^{\mathrm{S}}$ [rad]	+0.013	(−0.001)	+0.060 −0.067	(−0.04 +0.07)
C_s^{av}	+0.004	(+0.011)	0.049	(+0.26)
$\Delta C_s^{\parallel}$	−0.011	(+0.013)	0.131	(+0.08)
$\Delta C_s^{\perp}$	+0.066	(+0.021)	0.195	(+0.20)
C_s^{avS}	+0.059	(−0.001)	+0.036 −0.031	(+0.11 −0.04)
Γ_s [ps^{-1}]	0.6592	(+0.0001)	0.0031	(–)
$\Delta\Gamma_s$ [ps^{-1}]	+0.0786	(+0.0001)	0.0092	(–)
Δm_s [ps^{-1}]	17.709	(+0.013)	0.071	(+0.15)
$\lvert A_0^{\mathrm{CP}}\rvert^2$	0.5242	(+0.0001)	0.0035	(–)
$\lvert A_\perp^{\mathrm{CP}}\rvert^2$	0.2514	(+0.0002)	0.0049	(–)
$\delta_\parallel - \delta_0$ [rad]	+2.943	(−0.040)	+0.124 −0.236	(+0.19 +0.17)
$\delta_\perp - \delta_0$ [rad]	+3.034	(−0.003)	+0.177 −0.211	(+0.11 +0.19)

Table 4.10: *Results of a fit with four tagging categories for OS tags and four tagging categories for SS tags. Tagging-dilution factors follow from the linear model that is also applied in the nominal fit.*

parameter	estimate (difference)		uncertainty (rel. difference)	
ϕ_s^{av} [rad]	−0.040	(+0.007)	0.054	(+0.06)
$\Delta\phi_s^{\parallel}$ [rad]	−0.019	(−0.001)	0.045	(+0.05)
$\Delta\phi_s^{\perp\prime}$ [rad]	+0.003	(+0.005)	0.032	(+0.12)
$\Delta\phi_s^{\text{S}}$ [rad]	+0.017	(+0.003)	+0.060 −0.067	(−0.03 +0.08)
C_s^{av}	−0.017	(−0.011)	0.041	(+0.07)
$\Delta C_s^{\parallel}$	−0.040	(−0.016)	0.128	(+0.05)
$\Delta C_s^{\perp}$	+0.119	(+0.075)	0.172	(+0.06)
C_s^{avS}	+0.066	(+0.006)	0.035	(+0.08)
Γ_s [ps^{-1}]	0.6591	(–)	0.0031	(–)
$\Delta\Gamma_s$ [ps^{-1}]	+0.0783	(−0.0001)	0.0092	(–)
Δm_s [ps^{-1}]	17.679	(−0.017)	0.065	(+0.05)
$\lvert A_0^{\text{CP}}\rvert^2$	0.5236	(–)	0.0034	(–)
$\lvert A_\perp^{\text{CP}}\rvert^2$	0.2515	(+0.0003)	0.0049	(–)
$\delta_\parallel - \delta_0$ [rad]	+3.224	(−0.023)	+0.118 −0.253	(+0.14 +0.26)
$\delta_\perp - \delta_0$ [rad]	+2.996	(−0.042)	+0.170 −0.212	(+0.06 +0.20)

4.3 Systematic Uncertainties

Systematic uncertainties in the parameter estimates arise from uncertainties in the models used to describe the distributions of invariant masses, decay time, and decay angles. They are evaluated by repeating the fit of decay time and decay angles with varied values of relevant nuisance parameters or with manipulated data.

The value of some of the nuisance parameters can be determined with the $B_s^0 \to J/\psi K^+K^-$ data that are used in the time and angular fit. In these cases the value of the parameter is varied in the fit, often constrained by external information. Uncertainties originating from these parameters are therefore absorbed in the statistical uncertainties quoted in Section 4.1.

As described in Section 3.3.2, the parameters of the trigger decay-time acceptance function are fully determined from the $B_s^0 \to J/\psi K^+K^-$ data and constrained in the fit by counting decay candidates in the different trigger categories. The externally determined exponential factors for the reconstruction decay-time acceptance in 2011 and 2012 are combined with the information on their difference from the $B_s^0 \to J/\psi K^+K^-$ data by representing the external measurements as additional parabolic contributions to the NLL.

The resulting contributions of the decay-time acceptance uncertainties are estimated by repeating the fit with fixed acceptance parameters and subtracting the statistical uncertainties with varying and with fixed parameters in quadrature. Only the uncertainty of Γ_s is significantly affected. The contribution from the trigger acceptance is 0.0011 ps^{-1} and the contribution from the reconstruction acceptance 0.0015 ps^{-1}.

Other nuisance parameters constrained by external measurements are the parameters that describe the flavour-tagging calibration, as discussed in Section 3.6.2. Fixing the tagging parameters does not significantly affect any of the uncertainty estimates for other parameters.

Systematic effects that are not propagated to the statistical uncertainties are listed in Tables 4.11, 4.12, 4.13, and 4.14 for the principal parameters. Also a total uncertainty is shown for each parameter, determined by adding the statistical and systematic uncertainties in quadrature. The effects for the three CP-violation parameterizations are found to be similar and for most parameters only the results with the ϕ_s^i/C_s^i model are shown. Systematic uncertainties for ϕ_s are assumed to be the same as those for $\phi_s^{\rm av}$ and uncertainties for $|\lambda_s|$ are shown in Table 4.12.

Table 4.11: *Systematic uncertainties for the CP-violating phases.*

	ϕ_s^{av} [rad]	$\Delta\phi_s^{\parallel}$ [rad]	$\Delta\phi_s^{\perp\prime}$ [rad]	$\Delta\phi_s^{S}$ [rad]
$J/\psi K^+K^-$ mass				
resonant	0.003	0.001	0.001	0.009
factorization	0.003	0.005	0.001	0.016
K^+K^- mass				
integrals	–	–	–	0.006
decay time				
resolution	0.003	0.002	0.001	0.001
decay angles				
resolution	0.004	0.014	0.009	0.007
acc. statistical	0.002	0.007	0.004	0.004
acc. simulation	–	0.002	0.001	0.007
total systematic	0.007	0.017	0.010	0.022
statistical	0.051	0.043	0.029	0.062
total	0.05	0.05	0.03	0.07

Table 4.12: *Systematic uncertainties for the asymmetries from CP-violation in mixing and in decay. Uncertainties arising from K^+K^--mass integrals and decay-time resolution are negligible for these parameters.*

	$\lvert\lambda_s\rvert$	C_s^{av}	$\Delta C_s^{\parallel}$	$\Delta C_s^{\perp}$	C_s^{avS}
$J/\psi K^+K^-$ mass					
resonant	0.0028	0.001	0.008	0.004	0.004
factorization	0.0008	0.001	0.042	0.011	0.014
decay angles					
resolution	0.0014	0.004	0.030	0.011	0.010
acc. statistical	0.0024	0.001	0.015	0.006	0.005
acc. simulation	0.0048	0.001	0.003	0.001	0.011
total systematic	0.0063	0.005	0.055	0.018	0.021
statistical	0.0188	0.039	0.122	0.162	0.032
total	0.020	0.04	0.13	0.16	0.04

Table 4.13: *Systematic uncertainties for the B_s^0-lifetime parameters.*

	Γ_s [ps^{-1}]	$\Delta\Gamma_s$ [ps^{-1}]	Δm_s [ps^{-1}]
$J/\psi K^+K^-$ mass			
resonant	–	0.0002	0.001
factorization	–	0.0006	0.002
decay time			
resolution	–	–	0.006
$B_c^+ \to B_s^0 X$	0.0005	–	–
reco. acc. model	0.0013	0.0023	0.012
decay angles			
resolution	–	0.0008	0.006
acc. statistical	–	0.0001	0.001
acc. simulation	–	0.0001	0.001
total systematic	0.0014	0.0025	0.015
statistical	0.0033	0.0092	0.060
total	0.004	0.010	0.06

Table 4.14: *Systematic uncertainties for the transversity amplitudes.*

	$\lvert A_0^{CP}\rvert^2$	$\lvert A_\perp^{CP}\rvert^2$	$\delta_\parallel - \delta_0$ [rad]	$\delta_\perp - \delta_0$ [rad]
$J/\psi K^+K^-$ mass				
resonant	0.0002	0.0003	0.024	0.011
factorization	0.0064	0.0030	0.050	0.048
K^+K^- mass				
integrals	–	–	0.003	0.007
decay time				
resolution	–	–	0.004	0.008
decay angles				
resolution	0.0002	0.0009	0.036	0.022
acc. statistical	0.0005	0.0007	0.019	0.009
acc. simulation	0.0020	0.0011	0.006	0.002
total systematic	0.0067	0.0034	0.069	0.056
statistical	0.0034	0.0049	0.132	0.165
total	0.008	0.006	0.15	0.17

$J/\psi K^+K^-$-mass model: resonant backgrounds
Resonant backgrounds from B^0 and Λ_b^0 decays are subtracted by injecting simulated events with negative weights into the $B_s^0 \to J/\psi K^+K^-$ data sample, as described in Section 3.2. Two types of uncertainties arise from this procedure: uncertainties in the estimates of the numbers of background events that affect the $B_s^0 \to J/\psi K^+K^-$ measurement and uncertainties in the distributions of the relevant variables for the reflection backgrounds.

A systematic uncertainty for the numbers of background events is estimated by changing the subtracted background yields by one standard deviation. The (absolute) variations resulting from an upward and a downward fluctuation are averaged. The yields of all background components are varied simultaneously.

The distributions of the decay angles and tagging variables of simulated background events are reweighted to make them match the distributions in B^0 and Λ_b^0 data as closely as possible. Since the results of this procedure are not expected to be perfect, the difference in results with datasets containing reweighted and not reweighted distributions is taken as an additional systematic uncertainty. The resulting numbers are added in quadrature to the average of the yield uncertainty for the total systematics arising from resonant backgrounds. The total uncertainties are dominated by the results of the reweighting procedure.

$J/\psi K^+K^-$-mass model: statistical
Signal weights for $B_s^0 \to J/\psi K^+K^-$ candidates are determined with the model of $J/\psi K^+K^-$ mass after a fit to the mass distribution in data (Section 3.2). Statistical uncertainties in the signal and the combinatorial background mass models are propagated by repeating the time and angular fit five thousand times with different sets of signal weights. The sets of weights are obtained by calculating sWeights with different sets of mass-model parameters, which are generated according to a multivariate-Gaussian distribution using the means and covariances from the mass fit. Effects on the parameter estimates from the time and angular fit are found to be negligible.

$J/\psi K^+K^-$-mass model: factorization
The $J/\psi K^+K^-$-mass model depends on the angle θ_μ, as shown in Section 3.2. Since the background-subtraction procedure relies on the absence of correlations between the mass and other variables, this dependence leads to a systematic uncertainty. To evaluate the uncertainty, the background is

subtracted in intervals of $\cos\theta_\mu$ and parameter estimates with the resulting dataset are compared to the nominal estimates.

The $J/\psi\,K^+K^-$-mass model is found to be symmetric with respect to the point $\cos\theta_\mu = 0$. Signal weights are calculated separately for three $\cos\theta_\mu$ intervals: $|\cos\theta_\mu| < 0.25$, $0.25 \leq |\cos\theta_\mu| < 0.7$, and $0.7 \leq |\cos\theta_\mu|$. The choice of interval boundaries is motivated by the differences in the mass distribution. The differences between the parameter estimates from fits with these signal weights and the nominal weights are taken as a systematic uncertainty.

K^+K^--mass model: interval integrals
The expression for the differential decay rate is integrated over the K^+K^- mass in six intervals. As discussed in Sections 2.5 and 3.5, this leads to the K_{PS} factors in the $J/\psi\,\phi$–K^+K^- S-wave interference terms, which represent the ratios of the interval integrals for the two contributions. The K_{PS} factors depend on the K^+K^--mass models used in their calculation and a systematic uncertainty arises from uncertainties in the $J/\psi\,\phi$ and S-wave line shapes.

Two uncertainties are considered. The first originates from a finite K^+K^--mass resolution, which is estimated with simulated events. An alternative set of K_{PS} factor is calculated with a resolution that is 20% larger than the nominal estimate (fourth column in Table 3.6). To enhance the significance of the effect of resolution, the differences in K_{PS} factors with the nominal set are doubled before repeating the time and angular fit and the resulting deviations in the parameter estimates are divided by two for the corresponding systematic uncertainties.

The second uncertainty comes from the model of the K^+K^- S-wave. Nominally it is assumed that the S-wave consists only of a $J/\psi\,f_0(980)$ contribution, which is modelled with a Flatté function. However, the true composition of the S-wave is not exactly known. This uncertainty is evaluated by repeating the fit with K_{PS} factors calculated with a uniform line shape in K^+K^- mass (fifth column in Table 3.6). Deviations in the parameter estimates are added in quadrature to the resolution uncertainty.

Decay-time model: resolution
Parameters of the resolution model discussed in Section 3.3.1 are determined for prompt background candidates with zero decay time and it is assumed that these parameter values are applicable to the resolution of signal candidates. The systematic uncertainty associated with this assumption is evaluated with

simulated events. See reference [66] for a detailed description of this procedure.

In simulation the differences between the resolution-model parameters for signal candidates and prompt background are determined. With these differences a transformation between the two models is constructed, which is applied to the measured real-data parameters to obtain estimates of the signal-model parameters for real data. The difference in results between fits with the signal and background resolution models for real data is taken as a systematic uncertainty.

Decay-time model: $B_c^+ \to B_s^0 X$ decays
A fraction of B_s^0 mesons in the $B_s^0 \to J/\psi K^+K^-$ sample originates from $B_c^+ \to B_s^0 X$ decays, where X is a charged particle or combination of particles, for instance a pion. The B_c^+ meson has a mean lifetime of approximately 0.5 ps [45], resulting in a decay at a significant distance from the primary vertex where the B_c^+ is produced. Because it is assumed that the B_s^0 is produced in the primary vertex in the reconstruction of its decay time, the measured decay time for the $B_c^+ \to B_s^0 X$ contribution is wrong and leads to a systematic uncertainty.

The contribution of $B_c^+ \to B_s^0 X$ decays was estimated to be 0.8% of the $B_s^0 \to J/\psi K^+K^-$ signal sample [2]. The corresponding systematic uncertainty was estimated by generating pseudo experiments, where simulated $B_c^+ \to B_s^0(\to J/\psi K^+K^-) X$ decays were injected into each generated data sample. The mean deviations in the resulting parameter estimates are taken as systematic uncertainties. Only the estimate of Γ_s is significantly affected.

Decay-time model: reconstruction acceptance
The acceptance arising from reconstruction is modelled as a function with an exponential shape in decay time, as described in Section 3.3.2. This model only describes the acceptance function approximately, which gives systematic uncertainties in the estimates of the lifetime parameters.

The uncertainties are evaluated with the differences between fit results on a data sample that is corrected for the reconstruction acceptance and the nominal data sample. Data are corrected for acceptance by assigning decay-candidate weights that are inversely proportional to the estimated acceptance. The weights are calculated with the shape of the acceptance in the reconstruction variables of a candidate. Although this acceptance shape is still an

approximation, it describes the exact function significantly better than an exponential shape [2].

Decay-time model: $\mathrm{B_s^0}$–$\overline{\mathrm{B}}{}_\mathrm{s}^0$ normalization asymmetries
In the nominal analysis the absence of all time-independent $\mathrm{B_s^0}$–$\overline{\mathrm{B}}{}_\mathrm{s}^0$ asymmetries is assumed. That is, the assumption is made that there is no asymmetry between the $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$ decay-rate equations from the factor $1 - q_\mathrm{f}\, C_\mathrm{mix}$ (Equation 2.47), no production asymmetry between $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$, and no asymmetry between the numbers of $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$ in each tagging category. The assumption is expected to have no effect on the parameter estimates, as explained in Section 3.6. This is verified by generating ten thousand pseudo experiments with non-zero values for the normalization asymmetries and comparing the resulting parameter distributions with the distributions in the nominal pseudo experiments. No significant differences are found.

Since no precise values of the normalization asymmetries are available, values for each pseudo experiment are drawn from Gaussian distributions that represent the current estimates of these asymmetries. The uncertainties in the asymmetry-parameter estimates propagate to the widths of the parameter and pull distributions in the pseudo experiments.

The asymmetry $C_\mathrm{mix} \approx \frac{1}{2}\, a_\mathrm{fs}$ is estimated with the current average of measurements of CP-violation in mixing, $a_\mathrm{fs} = -0.0109 \pm 0.0040$ [32]. The uncertainty in the production asymmetry is estimated with a measurement of the $\mathrm{B^0}$ production asymmetry in LHCb, 0.006 ± 0.009 [71], where the $\mathrm{B_s^0}$ asymmetry is expected to be smaller than the $\mathrm{B^0}$ asymmetry. As an estimate of the uncertainty for the $\mathrm{B_s^0}$ asymmetry the sum of the value and uncertainty is taken and the production asymmetry is generated as 0 ± 0.15. Uncertainties in the tagging-efficiency differences for $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$ are estimated in tagging calibration [2] and lead to generated values of 0 ± 0.0015 for OS tagging and $= 0 \pm 0.0046$ for SS tagging.

Tagging-efficiency asymmetries for OS-untagged and SS-untagged decay candidates are equal but opposite to the asymmetry in the corresponding tagged category, weighted by the ratio of the tagged and untagged efficiencies. As a consequence, the OS and SS tagging efficiencies are required to calculate the untagged asymmetries. Assuming small asymmetries, the average of the $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}{}_\mathrm{s}^0$ efficiencies is estimated by the fraction of candidates in a category, yielding $\varepsilon_\mathrm{OS} = 0.3076 \pm 0.0017$ and $\varepsilon_\mathrm{OS} = 0.5291 \pm 0.0018$.

Defining tagging efficiencies for the OS and SS categories implicitly assumes factorization for these categories. That is, the efficiency for doubly-tagged candidates is given by $\varepsilon_{\text{OS}}\,\varepsilon_{\text{SS}}$, the efficiency for candidates only tagged by the OS algorithms by $\varepsilon_{\text{OS}}\,(1-\varepsilon_{\text{SS}})$, and so on. Calculating the OS efficiency separately for SS-tagged and SS-untagged candidates and vice versa indicates that factorization is not perfect. To account for this, half of the difference between the two efficiency estimates for each category is added in quadrature to the statistical uncertainties, which finally gives $\varepsilon_{\text{OS}} = 0.308 \pm 0.013$ and $\varepsilon_{\text{OS}} = 0.529 \pm 0.016$.

Decay-angles model: acceptance statistical
A non-trivial acceptance shape in the decay angles is included in the decay model with the normalization weights that were introduced in Section 3.4. The values of the normalization weights are estimated with simulated events. Because of the finite size of the simulated sample the estimates are affected by statistical uncertainties, which are propagated to the final parameter estimates. The time and angular fit is repeated for five thousand sets of varied normalization weights, generated according to a multivariate Gaussian distribution. The square roots of the parameter-estimate variances are taken as systematic uncertainties.

Decay-angles model: acceptance simulation
Kinematic distributions of simulated events do not match the distributions in real data perfectly. Although corrections are applied to the simulated distributions before calculating the angular-acceptance weights (see [2]), the discrepancies in the simulation are not fully understood and an associated systematic uncertainty remains. The size of this effect is estimated with the difference between the fit results with weights from a corrected and an uncorrected simulation.

Decay-angles model: resolution
The finite experimental resolution of the decay angles is not accounted for in the decay model. A study of the effects of a finite resolution is described in reference [2]. The resolution was estimated with simulated events and found to be of the order 0.01 rad. Its effect on the parameter estimates was studied by generating pseudo experiments and smearing the values of the decay angles according to the resolution distributions found in the simulation. Effects were quantified by the distribution of the difference in pull values for the fit results with and without angle smearing for each pseudo experiment. The mean and

the width of the pull-difference distribution are multiplied by the statistical uncertainty from the real-data fit for each parameter and added in quadrature to obtain the systematic uncertainty associated to decay-angle resolution.

The study of angular resolution was only performed for the $\phi_s/|\lambda_s|$ parameterization. Therefore, the systematic uncertainties for the CP-violation parameters in the other two models are estimated in a different way. The systematic for ϕ_s^{av} is taken to be equal to the estimate for ϕ_s. The parameter C_s^{av} appears in similar terms of the PDF as ϕ_s^{av} and, therefore, the systematic for C_s^{av} is also estimated to be equal to the systematic of ϕ_s.

The parameters that describe the differences between the ϕ_s^i and C_s^i parameters are measured from the interference terms in the PDF, as are the phase differences $\delta_{||} - \delta_0$ and $\delta_\perp - \delta_0$. To estimate the effect of decay-angle resolution on the CP-violation parameters, the effects of decay-angle resolution and decay-angle acceptance are compared for $\delta_{||} - \delta_0$ and $\delta_\perp - \delta_0$. For these phase differences, the impact of the propagation of statistical uncertainties in the acceptance is roughly half of the impact of resolution. To get an estimate for the resolution uncertainty in the CP-violation parameters it is assumed that the same factor of two applies.

4.4 Summary and Outlook

Estimates of the principal parameters are given in Table 4.15. As pointed out in Section 4.1, all results are statistically compatible with Standard Model predictions.

For some parameters the systematic uncertainty is comparable to or even larger than the statistical uncertainty. With roughly a factor ten increase in number of decay candidates this would be true for all parameters, which creates a need for reduction of the systematics.

A systematic with a significant contribution for most of the CP-violation parameters (Tables 4.11 and 4.12) is the uncertainty from decay-angle resolution. If this effect is going to be included as a systematic uncertainty in further measurements, the study of the resolution and its effect on the parameter estimates will have to be improved. Instead of describing the resolution separately for the three decay angles, a multidimensional resolution function can be constructed to include correlations. Further studies can also evaluate the effect on all parameters of the ϕ_s^i/C_s^i model.

Table 4.15: *Estimates of the principal parameters.*

parameter	value	uncertainty	
		statistical	systematic
ϕ_s^{av} [rad]	−0.05	±0.05	±0.01
$\Delta\phi_s^{\parallel}$ [rad]	−0.02	±0.04	±0.02
$\Delta\phi_s^{\perp\prime}$ [rad]	−0.00	±0.03	±0.01
$\Delta\phi_s^{S}$ [rad]	+0.01	±0.06	±0.02
C_s^{av}	−0.01	±0.04	±0.01
$\Delta C_s^{\parallel}$	−0.02	±0.12	±0.06
$\Delta C_s^{\perp}$	+0.04	±0.16	±0.02
C_s^{avS}	+0.06	±0.03	±0.02
Γ_s [ps^{-1}]	0.659	±0.003	±0.001
$\Delta\Gamma_s$ [ps^{-1}]	+0.078	±0.009	±0.003
Δm_s [ps^{-1}]	17.70	±0.06	±0.02
$\lvert A_0^{CP}\rvert^2$	0.524	±0.003	±0.007
$\lvert A_\perp^{CP}\rvert^2$	0.251	±0.005	±0.003
$\delta_\parallel - \delta_0$ [rad]	+3.25	$^{+0.10}_{-0.20}$	±0.07
$\delta_\perp - \delta_0$ [rad]	+3.04	$^{+0.16}_{-0.18}$	±0.06

When the systematic from angular resolution resolution becomes dominating, the possibility of including this effect in the model of decay time and angles must be considered. It will be challenging to implement this, since not all functions can be analytically convolved with the angular functions in the expression for the differential decay rate. Either functions should be found for which this is possible or numerical convolution methods should be explored.

Also uncertainties in the angular acceptance contribute to the systematics of the CP-violation parameters. The statistical component of these uncertainties can be reduced by generating more simulated decays. However, this will only emphasize the fact that the simulation does not perfectly describe real decays. If the simulation is to be used for future measurements, it has to be improved to provide a more accurate description of the production, decay, and detection of particles in LHCb events.

Another potential method of obtaining the angular acceptance would be to study detection efficiencies in real data particles and decays for which the underlying physics description is known. This is also an option for the reconstruction component of the decay-time acceptance, which dominates the systematic uncertainties of the B_s^0-lifetime parameters (Table 4.13). For the latter the shape of the acceptance function is already partially determined with real data.

An improved description of the reconstruction decay-time acceptance is already available and was used here to evaluate a systematic uncertainty and in reference [1] as the nominal description. Although this method provides a better description of the acceptance shape, it still introduces the dominant systematic uncertainties on the lifetime parameters and further study of this acceptance effect is required for further measurements.

Removing the uncertainty from decay-time acceptance leaves only the effect of $B_c^+ \to B_s^0 X$ decays for the parameter Γ_s. This uncertainty could be reduced by including a component for this contribution in the resolution model for the decay time, since these are real $B_s^0 \to J/\psi K^+K^-$ decays with a modified decay-time measurement. Remaining uncertainties would then originate from the estimates of the number of contributing $B_c^+ \to B_s^0 X$ decays and the shape of the resolution component.

Dominating contributions to systematic uncertainties for some of the CP-violation parameters and all of the transversity amplitudes (Table 4.14) are the uncertainties introduced by the model of $J/\psi K^+K^-$ mass. The part originating from resonant backgrounds can be reduced by improving the accuracy

of the description of the distributions of these backgrounds in the relevant variables. This can be achieved by improving the simulation of background decays, exploiting analyses of these background distributions in real data, or a combination of these methods.

The dominant $J/\psi K^+K^-$-mass systematics, however, originate from the background-subtraction procedure. A possible way of reducing the systematic from the non-factorization of the mass and time/angles models would be background subtraction in intervals of $\cos\theta_\mu$, as was done in this analysis to evaluate the systematic uncertainties. Alternatively, introducing the estimate of the uncertainty in the $J/\psi K^+K^-$-mass measurement in both the models could provide a way to describe the correlations between mass and angles. This could be implemented in combination with a background-subtraction procedure, but an alternative is to fit unweighted $B_s^0 \to J/\psi K^+K^-$ decays with a five-dimensional model of mass, decay time, and decay angles.

In addition to reducing systematic uncertainties, the precision of the CP-violation measurement can potentially be improved by reducing the number of wrong flavour tags. Given a wrong-tag probability of 40%, a 10% improvement of this value gives almost a factor two increase in the effective number of perfectly tagged decays (see also Section 3.6.2). Improvements of the flavour-tagging procedure may include the use of more sophisticated (multivariate) techniques to select and analyse tagging particles, the development of additional tagging algorithms, and an improved understanding of the properties of particles that are created in association with B_s^0 mesons.

Conclusions

Figure C.11 shows the current status of ϕ_s and $\Delta\Gamma_s$ measurements. It is an update of Figure 1.8 in Section 1.3.2, which gives an overview of the status in the Spring of 2014, when only results with the 2011 dataset from LHCb were available. Results of the measurement presented in this thesis, which uses both the 2011 and 2012 LHCb datasets and was published in [1], are shown in combination with the results from measurements in the $B_s^0 \to J/\psi\,\pi^+\pi^-$ [42] and $B_s^0 \to D_s^+\,D_s^-$ [72] decay channels. The other results in the figure are from a new CMS measurement [73], the updated ATLAS measurement [74], and the CDF [40] and D0 [39] measurements in the $B_s^0 \to J/\psi\,\phi$ channel.

The objective of these measurements, as described in Sections 1.1–1.3, is to test the Standard Model by comparing the CP violation that is measured in $b \to c\bar{c}s$ transitions to the prediction obtained by interpreting other measurements within the Standard Model framework. The current combined precision of the ϕ_s result is 0.04 rad, which is equal to the deviation of the Standard Model prediction from zero. This precision is not yet sufficient to measure potential small deviations from the Standard Model, but these measurements do rule out large contributions from non-Standard Model physics. As discussed in Section 1.3.2, both experimental and theoretical improvements are required for a more precise analysis of CP violation in $b \to c\bar{c}s$ transitions.

At lowest order, the decays that are included in the combination of Figure C.11 are governed by a single tree-level $b \to c\bar{c}s$ transition and CP violation is described by common ϕ_s and $|\lambda_s|$ parameters. To compare a more precise measurement of CP violation to its prediction in the Standard Model, higher order (penguin) contributions must be considered. These potentially yield CP violation parameters with different values for each decay and each angular-momentum state contributing to the $B_s^0 \to J/\psi\,K^+K^-$ decay.

A close interplay between theory and experiment will be required to interpret a precise measurement of these CP-violation parameters. Calculations

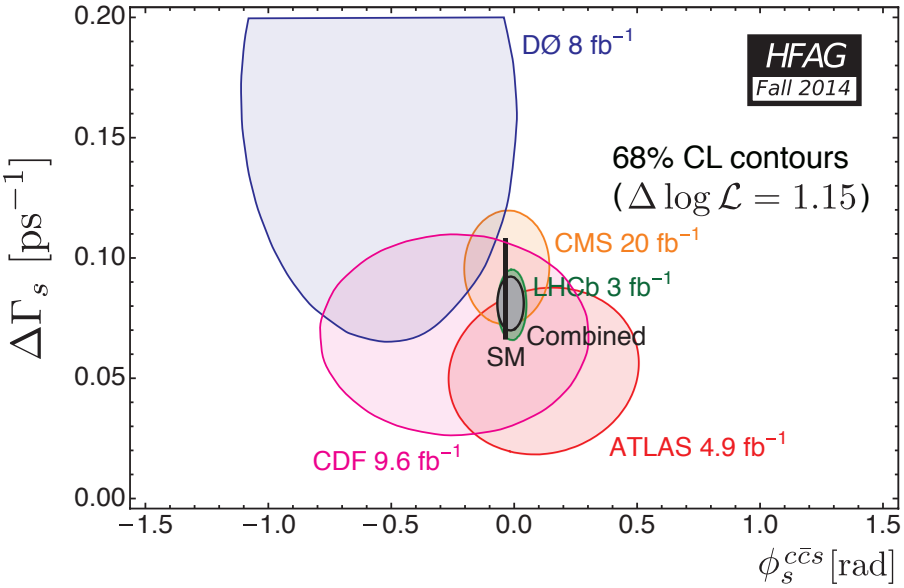


Figure C.11: *Combination of ϕ_s (here represented as $\phi_s^{c\bar{c}s}$) and $\Delta\Gamma_s$ measurements by HFAG [32]. The estimates at 68% confidence level (CL) by the different experiments are shown by the coloured contours. Note that the LHCb contour (green) is a combination of measurements in the $B_s^0 \to J/\psi\phi$, $B_s^0 \to J/\psi\pi^+\pi^-$, and $B_s^0 \to D_s^+ D_s^-$ decays. The combined 68% confidence region is shown by the grey area and the Standard Model prediction by the vertical bar.*

of their values (see e.g. [75]) suffer from uncertainties in the strong interactions within the involved hadrons. However, by exploiting approximate symmetries between different decays of B^0 and B_s^0 mesons, a framework of measurements and calculations can be built to interpret the experimental results [38, 76]. Both precise measurements and precise estimates of how the symmetries between different decays are broken are required for such an analysis.

In reference [76] it is proposed to use the interplay between the decays $B_s^0 \to J/\psi\,\phi$, $B_s^0 \to J/\psi\,K_S^0$, $B_s^0 \to J/\psi\,\overline{K}^{*0}$, $B^0 \to J/\psi\,K_S^0$, and $B^0 \to J/\psi\,\rho^0$ for an analysis of CP violation in $b \to c\bar{c}s$ and $b \to c\bar{c}d$ transitions. For both the $B_s^0 \to J/\psi\,\phi$ and $B^0 \to J/\psi\,\rho^0$ decays measurements of CP violation per intermediate angular-momentum state is required in this framework. The results of the first measurement in this format for the $B_s^0 \to J/\psi\,\phi$ channel

was presented in Chapter 4 (see also reference [1]).

While the current CP-violation results are still compatible with both the Standard Model and no CP violation, it is expected that the measurement with future LHCb data will start to distinguish between the Standard Model and other scenarios. With the final dataset of the LHCb experiment an increase in precision of roughly an order of magnitude is expected [77, 78]. This improvement has to come from an increased number of decays produced by the LHC, but also from improvements in the analysis procedure, as described in Section 4.4.

To implement all the required improvements, it may be necessary to redesign key analysis components. In particular the treatment of background decay candidates and the methods of obtaining and describing the acceptance and resolution shapes must be improved to enable more precise measurements of CP-violation and B_s^0-lifetime parameters. Any enhancements of the flavour-tagging algorithms will contribute to a better statistical CP-violation precision, in addition to the increased set of $B_s^0 \to J/\psi K^+K^-$ data.

As indicated in Section 4.4, there are plenty of openings to achieve the experimental precision needed for the next generation of $B_s^0 \to J/\psi K^+K^-$ measurements. Combined with developments in theory, these measurements have the potential to play a key role in the continuation of Standard Model tests and the establishment of a description of particle interactions beyond the Standard Model.

Appendix

Angular Differential Decay Rate

A.1 Angular Amplitude

The dependence of the $B_s^0 \to J/\psi K^+K^-$ decay on the three decay angles, as described in Section 2.4, can be derived within the helicity formalism [60,61]. In this formalism the decay is described in terms of three two-body decays; first the decay of the B_s^0 into a J/ψ and a second intermediate particle, followed by the decay of the J/ψ into a $\mu^+\mu^-$ pair and the decay of the second particle into a K^+K^- pair. The angular dependence arises from the rotations of the spin vectors of the decaying particles into the momentum directions of the decaying particles.

A generalized form of the decay is shown in Figure A.1. The three coordinate systems in part (a) of the figure are defined in the centre-of-mass frames of particles B, a, and b in the decay $B \to a(\to P_1\, P_2)\, b(\to P_3\, P_4)$. The spherical coordinates θ and φ specify the momentum direction of one of the particles in each two-body decay. The other particle, which is not shown in the figure, has an opposite momentum.

The directions of the decay-product momenta define a *helicity axis* for each of the three decays. The z axis of a particle coordinate system lies along the helicity axis that is associated to the production of the particle. The momentum direction of one of the two particles in a decay is chosen as the positive z direction, as depicted in Figure A.1a for particles a and b.

The sum of the spin projections of the two particles in each decay along

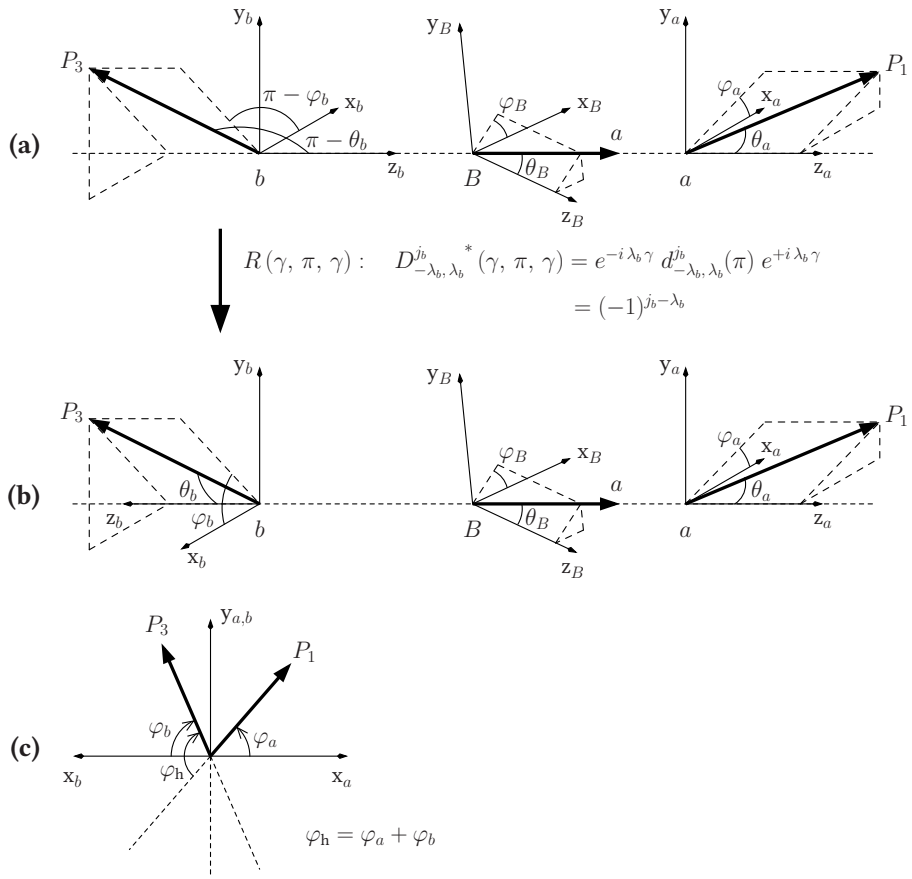


Figure A.1: *Helicity frames for the decay $B \to a(\to P_1\, P_2)\, b(\to P_3\, P_4)$ in (a) a general configuration where the coordinate systems in the a and b rest frames are aligned and (b) with the Jacob–Wick convention, in which the coordinate system in the b rest frame is rotated. A definition of the helicity angle φ_h in the Jacob–Wick convention is shown in (c).*

the helicity axis is given by the difference of the particle helicities. Assuming the spin projection of the decaying particle in the z direction is known, the amplitude for the measuring the required spin projection along the helicity axis is given by the complex conjugate of a *Wigner D-matrix*, which is a function of three Euler angles that specify the rotation of the spin projection from the z axis to the helicity axis.

The D-matrix for each of the two-body decays is expressed in terms of a real-valued d-matrix and two exponential functions as

$$D^{j}_{m,\,n}(\varphi,\theta,\varphi') = e^{-im\varphi}\, d^{j}_{m,\,n}(\theta)\, e^{-in\varphi'} \,, \tag{A.1}$$

where j is the spin of the decaying particle, m the projection of the spin in the direction of the z axis, n the projection of the spin in the positive direction of the helicity axis. The Euler angles that specify the rotation of the spin projection are φ, θ, and φ'. The angles θ and φ are equal to the spherical coordinates in the mother-particle rest frame, while the angle φ' rotates the coordinate system of a decay product around the helicity axis. In the *Jacob-Wick* convention [60] φ' is chosen to be equal to $-\varphi$, resulting in

$$D^{j}_{m,\,n}(\varphi,\theta,-\varphi) = d^{j}_{m,\,n}(\theta)\, e^{-i(m-n)\varphi} \,. \tag{A.2}$$

A boost along the helicity axis does not affect the spin projections along this axis. Therefore, the projection of the spin of a decay product along the helicity axis in the centre-of-mass frame of its production is equal to the spin projection in the z direction of the coordinate system in its rest frame. As a result, this projection relates the spins of the decays that are described in the two frames.

To define angles that are consistent with the helicity angles in Figure 2.2, the coordinate system of particle b is rotated by 180° around the y_b axis, as shown in Figure A.1b. This operation introduces an additional D-matrix, which represents the amplitude for a rotation with $\theta = \pi$ and arbitrary, but equal remaining Euler angles. The spin projection in the z_b direction goes from $-\lambda_b$ to $+\lambda_b$, where λ_b is the helicity of particle b in the B rest frame. Using the d-matrix property $d^{j}_{m,\,n}(\pi) = (-1)^{j-n}\,\delta_{m,-n}$, this D-matrix can be written as

$$D^{j_b}_{-\lambda_b,\,\lambda_b}{}^{*}(\gamma,\pi,\gamma) = e^{-i\lambda_b\gamma}\, d^{j_b}_{-\lambda_b,\,\lambda_b}(\theta)\, e^{+i\lambda_b\gamma} = (-1)^{j_b-\lambda_b} \,. \tag{A.3}$$

Identifying the particles P_1 and P_3 with K^+ and μ^+, respectively, the angle θ_a corresponds to θ_{K} and the angle θ_b to θ_{μ}. The angle φ_{h} is given by the sum of φ_a and φ_b, as shown in Figure A.1c.

Combining the D-matrices for the three decays, the factor $\mathcal{H}$ in Equation 2.37 is given by

$$\begin{aligned}\mathcal{H}_h = &\sqrt{\tfrac{2j_B+1}{4\pi}\,\tfrac{2j_a^h+1}{4\pi}\,\tfrac{2j_b^h+1}{4\pi}} \\ &\times (-1)^{j_b^h-\lambda_b^h}\, D^{j_B}_{\lambda_B,\,\lambda_a^h-\lambda_b^h}{}^{*}(\Omega_B)\, D^{j_a^h}_{\lambda_a^h,\,\lambda_1-\lambda_2}{}^{*}(\Omega_a)\, D^{j_b^h}_{\lambda_b^h,\,\lambda_3-\lambda_4}{}^{*}(\Omega_b)\,,\end{aligned} \tag{A.4}$$

where Ω_p is a shorthand notation for the set of Euler angles in the decay of particle p and the factors $\sqrt{\frac{2j+1}{4\pi}}$ are normalization factors for the two-particle states in the B, a, and b decays. The index h runs over the intermediate helicity states, for which the particles have definite helicities. The final expressions for helicity states will be combined into expressions for the transversity states of Equation 2.37. Notice that the spins and helicities of particles B and 1–4 do not depend on the intermediate state, since these particles are external and their spin state is thus, in principle, observable.

Concentrating only on the case where particle B is a spinless particle, j_B, λ_B, and $\lambda_a - \lambda_b$ are equal to zero. As a result, the D-matrix corresponding to the B decay reduces to $D^0_{0,\,0}{}^{*} = 1$, which makes the decay independent of the orientation of the helicity axis with respect to the direction of the B momentum. Introducing the notation $\lambda_a = \lambda_b \equiv \lambda$, $\lambda_1 - \lambda_2 \equiv \alpha$, and $\lambda_3 - \lambda_4 \equiv \beta$, Equation A.4 reduces to

$$\mathcal{H}_h = \frac{(-1)^{j_b^h-\lambda^h}}{(4\pi)^{3/2}}\sqrt{(2j_a^h+1)\,(2j_b^h+1)}\; D^{j_a^h}_{\lambda^h,\,\alpha}{}^{*}(\Omega_a)\, D^{j_b^h}_{\lambda^h,\,\beta}{}^{*}(\Omega_b)\,. \tag{A.5}$$

A.2 Squared Angular Amplitude

With Equation A.5 the products $\mathcal{H}_h^*\,\mathcal{H}_{h'}$ that appear in Equation 2.41 can be expressed as

$$\begin{aligned}&\mathcal{H}_h^*\,\mathcal{H}_{h'} \\ &\quad= \frac{(-1)^{j_b^h+j_b^{h'}+M^{hh'}-2\lambda^h}}{(4\pi)^3}\sqrt{(2j_a^h+1)\,(2j_a^{h'}+1)\,(2j_b^h+1)\,(2j_b^{h'}+1)} \\ &\qquad\times D^{j_a^h}_{\lambda^h,\,\alpha}(\Omega_a)\, D^{j_a^{h'}}_{\lambda^{h'},\,\alpha}{}^{*}(\Omega_a)\, D^{j_b^h}_{\lambda^h,\,\beta}(\Omega_b)\, D^{j_b^{h'}}_{\lambda^{h'},\,\beta}{}^{*}(\Omega_b)\,.\end{aligned} \tag{A.6}$$

To evaluate the dependence between the D-matrices in this expression the following relations are applied:

$$D^{j}_{m,\,n}{}^{*} = (-1)^{m-n}\, D^{j}_{-m,\,-n} \tag{A.7a}$$

$$D^{j}_{m,\,n}\, D^{j'}_{m',\,n'} = \sum_{J=|j-j'|}^{j+j'} \langle j\, m,\, j'\, m' | J\, m+m' \rangle \times \langle j\, n,\, j'\, n' | J\, n+n' \rangle\, D^{J}_{m+m',\,n+n'} \tag{A.7b}$$

$$\langle j\, m,\, j'\, m' | J\, M \rangle = (-1)^{j-j'+M} \sqrt{2\,J+1} \begin{pmatrix} j & j' & J \\ m & m' & -M \end{pmatrix} \tag{A.7c}$$

where $\langle j\, m,\, j'\, m' | J\, M \rangle$ is a Clebsch-Gordan coefficient, which is related to the Wigner 3j symbol $\begin{pmatrix} j & j' & J \\ m & m' & -M \end{pmatrix}$. The product of a D-matrix and a complex conjugate D-matrix with equal indices n can now be written as

$$\begin{aligned} D^{j}_{m,\,n}\, D^{j'}_{m',\,n}{}^{*} &= (-1)^{m'-n}\, D^{j}_{m,\,n}\, D^{j'}_{-m',\,-n} \\ &= (-1)^{m'-n} \sum_{J=|j-j'|}^{j+j'} \langle j\, m,\, j'\, -m' | J\, m-m' \rangle \times \langle j\, n,\, j'\, -n | J\, 0 \rangle\, D^{J}_{m-m',\,0} \\ &= (-1)^{m-n} \sum_{J=|j-j'|}^{j+j'} (2J+1) \begin{pmatrix} j & j' & J \\ m & -m' & -m+m' \end{pmatrix} \times \begin{pmatrix} j & j' & J \\ n & -n & 0 \end{pmatrix} D^{J}_{m-m',\,0}\,. \end{aligned} \tag{A.8}$$

With this expression the products of two D-matrices in Equation A.6 that are functions of the same set of angles can be substituted by a sum of single D-matrices. After this substitution and with the definition $M^{hh'} \equiv \lambda^{h} - \lambda^{h'}$, the product $\mathcal{H}^{*}_{h}\, \mathcal{H}_{h'}$ is given by

$$\begin{aligned} &\mathcal{H}^{*}_{h}\, \mathcal{H}_{h'} \\ &\quad = \frac{(-1)^{j_b^h + j_b^{h'} + M^{hh'} - \alpha - \beta}}{(4\pi)^3} \sqrt{(2j_a^h+1)\,(2j_a^{h'}+1)\,(2j_b^h+1)\,(2j_b^{h'}+1)} \\ &\qquad \times \sum_{J_a^{hh'}=|j_a^h - j_a^{h'}|}^{j_a^h + j_a^{h'}} \left(2J_a^{hh'}+1\right) \begin{pmatrix} j_a^h & j_a^{h'} & J_a^{hh'} \\ \lambda^h & -\lambda^{h'} & -M^{hh'} \end{pmatrix} \begin{pmatrix} j_a^h & j_a^{h'} & J_a^{hh'} \\ \alpha & -\alpha & 0 \end{pmatrix} \end{aligned}$$

$$\times \sum_{J_b^{hh'}=|j_b^h-j_b^{h'}|}^{j_b^h+j_b^{h'}} \left(2J_b^{hh'}+1\right) \begin{pmatrix} j_b^h & j_b^{h'} & J_b^{hh'} \\ \lambda^h & -\lambda^{h'} & -M^{hh'} \end{pmatrix} \begin{pmatrix} j_b^h & j_b^{h'} & J_b^{hh'} \\ \beta & -\beta & 0 \end{pmatrix} \times D^{J_a^{hh'}}_{M^{hh'},0}(\Omega_a)\, D^{J_b^{hh'}}_{M^{hh'},0}(\Omega_b)\,. \tag{A.9}$$

Notice that the angular dependence is still written in terms of the sets of Euler angles corresponding to the decays of particles a and b, Ω_a and Ω_b, respectively. The dependence on the helicity angles, $\theta_{\mathrm{K}} \equiv \theta_a$, $\theta_{\mu} \equiv \theta_b$, and $\varphi_{\mathrm{h}} \equiv \varphi_a + \varphi_b$, becomes apparent by expressing the product of D-matrices in Equation A.9 in terms of d-matrices and exponential functions (see Equation A.2):

$$\begin{aligned} & D^{J_a^{hh'}}_{M^{hh'},0}(\Omega_a)\, D^{J_b^{hh'}}_{M^{hh'},0}(\Omega_b) \\ &\quad = d^{J_a^{hh'}}_{M^{hh'},0}(\theta_a)\, e^{-iM^{hh'}\varphi_a}\, d^{J_b^{hh'}}_{M^{hh'},0}(\theta_b)\, e^{-iM^{hh'}\varphi_b} \\ &\quad = d^{J_a^{hh'}}_{M^{hh'},0}(\theta_{\mathrm{K}})\, d^{J_b^{hh'}}_{M^{hh'},0}(\theta_{\mu})\, e^{-iM^{hh'}\varphi_{\mathrm{h}}} \\ &\quad = \sqrt{\frac{(J_a^{hh'}-M^{hh'})!}{(J_a^{hh'}+M^{hh'})!}\,\frac{4\pi}{2J_b^{hh'}+1}}\; P^{M^{hh'}}_{J_a^{hh'}}(\cos\theta_{\mathrm{K}})\, Y^{M^{hh'}\,*}_{J_b^{hh'}}(\theta_{\mu},\varphi_{\mathrm{h}})\,. \end{aligned} \tag{A.10}$$

The functions $P_j^m(\cos\theta)$ and $Y_j^m(\theta,\varphi)$ are associated Legendre polynomials and spherical harmonics, respectively, which are defined by

$$P_j(x) \equiv \frac{1}{2^j\, j!}\frac{\mathrm{d}^j}{\mathrm{d}x^j}\left(x^2-1\right)^j \qquad \text{with } j \geq 0 \tag{A.11a}$$

$$P_j^m(x) \equiv \begin{cases} (-1)^m\, \left(1-x^2\right)^{\frac{1}{2}m}\, \frac{\mathrm{d}^m}{\mathrm{d}x^m} P_j(x) & (m \geq 0) \\ \frac{(j-|m|)!}{(j+|m|)!}\, \left(1-x^2\right)^{\frac{1}{2}|m|}\, \frac{\mathrm{d}^{|m|}}{\mathrm{d}x^{|m|}} P_j(x) & (m < 0) \end{cases} \tag{A.11b}$$

$$Y_j^m(\theta,\phi) \equiv \sqrt{\tfrac{2j+1}{4\pi}\,\tfrac{(j-m)!}{(j+m)!}}\; P_j^m(\cos\theta)\, e^{im\phi} \tag{A.11c}$$

$$d^j_{m,0}(\theta) = \sqrt{\tfrac{(j-m)!}{(j+m)!}}\; P_j^m(\cos\theta)\,. \tag{A.11d}$$

The angular functions that eventually appear in the expression of the differential decay rate are given by the real and imaginary parts of the products $\mathcal{H}_h^*\,\mathcal{H}_{h'}$. After substituting the expression for $D^{J_a}_{M,0}(\Omega_a)\, D^{J_b}_{M,0}(\Omega_b)$ from

Equation A.10, the only complex-valued factor on the right-hand side of Equation A.9 is the spherical harmonic $Y_{J_b}^{M^*}(\theta_\mu, \varphi_h)$. Notice that this function is real valued if its upper index is equal to zero:

$$Y_j^0(\theta, \phi) = \sqrt{\tfrac{2j+1}{4\pi}}\, P_j(\cos\theta)\,. \tag{A.12}$$

In the implementation of the angular functions real-valued spherical harmonics are used, which are defined by

$$Y_{j,m}(\theta,\phi) = \begin{cases} Y_j^0(\theta,\phi) & (m=0) \\ \sqrt{2}\,\Re[Y_j^m(\theta,\phi)] & (m>0) \\ \sqrt{2}\,\Im[Y_j^{|m|}(\theta,\phi)] & (m<0) \end{cases} \tag{A.13}$$

The squared amplitude in the helicity basis can be expressed in a form that is similar to Equation 2.53:

$$\begin{aligned} |\mathcal{A}|^2 &\propto \sum_h |H_h\,\mathcal{H}_h|^2 + \sum_{h\neq h'} \Re(H_h^*\,H_{h'}\,\mathcal{H}_h^*\,\mathcal{H}_{h'}) \\ &\propto \sum_h |H_h|^2\,|\mathcal{H}_h|^2 \\ &\quad + \sum_{h\neq h'} \Re(H_h^*\,H_{h'})\,\Re(\mathcal{H}_h^*\,\mathcal{H}_{h'}) - \Im(H_h^*\,H_{h'})\,\Im(\mathcal{H}_h^*\,\mathcal{H}_{h'})\,. \end{aligned} \tag{A.14}$$

Terms in the second sum of Equation A.14 with h and h' swapped are identical, since $H_h^*\,H_{h'}\,\mathcal{H}_h^*\,\mathcal{H}_{h'}$ and $H_{h'}^*\,H_h\,\mathcal{H}_{h'}^*\,\mathcal{H}_h$ are complex conjugates, which have identical real parts. The sign of $M^{hh'} \equiv \lambda^h - \lambda^{h'}$ is opposite for these terms. Adding the two contributions for each of the $h \leftrightarrow h'$ pairs, the angular functions in the helicity basis are given by

$$\begin{aligned} |\mathcal{H}_h|^2 &\propto P^0_{J_a^{hh'}}(\cos\theta_K)\,{Y^0_{J_b^{hh'}}}^*(\theta_\mu,\varphi_h) \\ &= P^0_{J_a^{hh'}}(\cos\theta_K)\,Y_{J_b^{hh'},0}(\theta_\mu,\varphi_h) && \text{(A.15a)} \\ 2\,\Re(\mathcal{H}_h^*\,\mathcal{H}_{h'}) &\propto 2\,\Re\left[P^{|M^{hh'}|}_{J_a^{hh'}}(\cos\theta_K)\,{Y^{|M^{hh'}|}_{J_b^{hh'}}}^*(\theta_\mu,\varphi_h)\right] \\ &= \sqrt{2}\,P^{|M^{hh'}|}_{J_a^{hh'}}(\cos\theta_K)\,Y_{J_b^{hh'},|M^{hh'}|}(\theta_\mu,\varphi_h) && \text{(A.15b)} \\ -2\,\Im(\mathcal{H}_h^*\,\mathcal{H}_{h'}) &\propto -2\,\Im\left[P^{|M^{hh'}|}_{J_a^{hh'}}(\cos\theta_K)\,{Y^{|M^{hh'}|}_{J_b^{hh'}}}^*(\theta_\mu,\varphi_h)\right] \\ &= \sqrt{2}\,P^{|M^{hh'}|}_{J_a^{hh'}}(\cos\theta_K)\,Y_{J_b^{hh'},-|M^{hh'}|}(\theta_\mu,\varphi_h) && \text{(A.15c)} \end{aligned}$$

A.3 Angular Functions for $B_s^0 \to J/\psi K^+K^-$

Identifying particles B, a, and b with the B_s^0, the K^+K^- pair, and the J/ψ, respectively, the angular dependence of the $B_s^0 \to J/\psi K^+K^-$ decay can be derived from Equations A.9, A.10, A.14, and A.15. As discussed in Section 2.2, the indices h and h' in Equation A.14 run over three $B_s^0 \to J/\psi \phi$ polarization states and one $B_s^0 \to J/\psi K^+K^-$ state where the K^+K^- pair is in an S-wave configuration.

Both the spins of the ϕ and the J/ψ are equal to one ($j_a = j_b = 1$), which results in sums over $J_a, J_b = 0, 1, 2$ for each combination of the $B_s^0 \to J/\psi \phi$ helicity states. The three states are given by the possible ϕ and J/ψ helicities: $\lambda = 0$ ("0"), $\lambda = +1$ ("+"), and $\lambda = -1$ ("−"). Since the K^+K^- system has no orbital angular momentum for the K^+K^- S-wave ("S") and the kaons are spinless, both j_a and λ are equal to zero for this state. The value of J_a can only be zero for $h = h' = S$ and one for $h \neq S$ and $h' = S$. Only one value for $|M^{hh'}| \equiv |\lambda^h - \lambda^{h'}|$ is possible for each combination of states, varying from zero ($\lambda^h = \lambda^{h'} = 0$) to two ($\lambda^h = -\lambda^{h'} = 1$).

Kaons are spinless, which results in $\alpha = 0$ for the $B_s^0 \to J/\psi K^+K^-$ decay. The helicities of the muons from the J/ψ decay can both be either $+\frac{1}{2}$ or $-\frac{1}{2}$, which gives the combinations $\beta \in \{0, \pm 1\}$. Since the μ^+ and the μ^- are produced with opposite chirality, the case $\beta = 0$ requires the helicity of one of the muons to be opposite to its chirality. These contributions are suppressed by a factor $m_\mu^2/m_{\mu^+\mu^-}^2 \approx 10^{-3}$ [79], where m_μ is the muon mass and $m_{\mu^+\mu^-}$ the dimuon invariant mass, the latter of which is equal to the mass of the J/ψ. For this reason only contributions with opposite helicities are considered, for which $\beta = \pm 1$.

Evaluation of the angular functions for the combinations of the $B_s^0 \to J/\psi \phi$ states leads to Table A.1. The first column of the table shows a combination of coefficients (helicity and transversity amplitudes) and the second column the corresponding angular functions. The angular functions for the K^+K^- S-wave and the interference between $B_s^0 \to J/\psi \phi$ and the K^+K^- S-wave are listed in Table A.2. All functions are multiplied by a factor $8\pi^2$, which makes their integrals over all three angles equal to either one or zero.

The top halves of the tables show the angular functions in the helicity basis, while the bottom halves show the functions in the transversity basis. For both bases helicity angles are used. The transversity functions are obtained by substituting the helicity amplitudes (H_0, H_+, and H_-) by combinations of transversity amplitudes (A_0, $A_\parallel$, and $A_\perp$), after which the helicity functions

Table A.1: *Angular functions for the $B_s^0 \to J/\psi\phi$ decay expressed in terms of associated Legendre polynomials and spherical harmonics in helicity angles. Functions are shown for $\beta = \pm 1$. Top: functions in the helicity basis. Bottom: functions in the transversity basis.*

amplitudes	$f(\Omega) \times 16\sqrt{\pi}$
$\lvert H_0\rvert^2$	$4\,(P_0^0 + 2\,P_2^0)\,(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0})$
$\lvert H_+\rvert^2$	$2\,(P_0^0 - P_2^0)\,(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} \pm \sqrt{3}\,Y_{1,0})$
	$= P_2^2\,(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} \pm \sqrt{3}\,Y_{1,0})$
$\lvert H_-\rvert^2$	$2\,(P_0^0 - P_2^0)\,(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} \mp \sqrt{3}\,Y_{1,0})$
	$= P_2^2\,(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} \mp \sqrt{3}\,Y_{1,0})$
$\Re(H_0^* H_+)$	$+2\sqrt{\frac{3}{5}}\,P_2^1(Y_{2,+1} \pm \sqrt{5}\,Y_{1,+1})$
$\Im(H_0^* H_+)$	$-2\sqrt{\frac{3}{5}}\,P_2^1(Y_{2,-1} \pm \sqrt{5}\,Y_{1,-1})$
$\Re(H_0^* H_-)$	$+2\sqrt{\frac{3}{5}}\,P_2^1(Y_{2,+1} \mp \sqrt{5}\,Y_{1,+1})$
$\Im(H_0^* H_-)$	$+2\sqrt{\frac{3}{5}}\,P_2^1(Y_{2,-1} \mp \sqrt{5}\,Y_{1,-1})$
$\Re(H_+^* H_-)$	$-2\sqrt{\frac{3}{5}}\,P_2^2\,Y_{2,+2}$
$\Im(H_+^* H_-)$	$-2\sqrt{\frac{3}{5}}\,P_2^2\,Y_{2,-2}$
$\lvert A_0\rvert^2$	$4\,(P_0^0 + 2\,P_2^0)\,(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0})$
$\lvert A_\parallel\rvert^2$	$P_2^2\,(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} - \sqrt{\frac{3}{5}}\,Y_{2,+2})$
$\lvert A_\perp\rvert^2$	$P_2^2\,(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} + \sqrt{\frac{3}{5}}\,Y_{2,+2})$
$\Re(A_0^* A_\parallel)$	$+2\sqrt{2}\sqrt{\frac{3}{5}}\,P_2^1\,Y_{2,+1}$
$\Im(A_0^* A_\parallel)$	$\mp 2\sqrt{2}\,\sqrt{3}\,P_2^1\,Y_{1,-1}$
$\Re(A_0^* A_\perp)$	$\pm 2\sqrt{2}\,\sqrt{3}\,P_2^1\,Y_{1,+1}$
$\Im(A_0^* A_\perp)$	$-2\sqrt{2}\sqrt{\frac{3}{5}}\,P_2^1\,Y_{2,-1}$
$\Re(A_\parallel^* A_\perp)$	$\pm 2\sqrt{3}\,P_2^2\,Y_{1,0}$
$\Im(A_\parallel^* A_\perp)$	$+2\sqrt{\frac{3}{5}}\,P_2^2\,Y_{2,-2}$

Table A.2: *Angular functions for the $B_s^0 \to J/\psi K^+K^-$ decay with a K^+K^- S-wave and the $B_s^0 \to J/\psi\phi$ and K^+K^- S-wave interference expressed in terms of associated Legendre polynomials and spherical harmonics in helicity angles. Functions are shown for $\beta = \pm 1$. Top: functions in the helicity basis. Bottom: functions in the transversity basis.*

amplitudes	$f(\Omega) \times 16\sqrt{\pi}$
$\lvert H_S\rvert^2$	$4\,P_0^0\,(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0})$
$\Re(H_0^* H_S)$	$8\sqrt{3}\,P_1^0\,(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0})$
$\Im(H_0^* H_S)$	0
$\Re(H_+^* H_S)$	$+6\,P_1^1\,(\frac{1}{\sqrt{5}}\,Y_{2,+1} \pm Y_{1,+1})$
$\Im(H_+^* H_S)$	$+6\,P_1^1\,(\frac{1}{\sqrt{5}}\,Y_{2,-1} \pm Y_{1,-1})$
$\Re(H_-^* H_S)$	$+6\,P_1^1\,(\frac{1}{\sqrt{5}}\,Y_{2,+1} \mp Y_{1,+1})$
$\Im(H_-^* H_S)$	$-6\,P_1^1\,(\frac{1}{\sqrt{5}}\,Y_{2,-1} \mp Y_{1,-1})$
$\lvert A_S\rvert^2$	$4\,P_0^0\,(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0})$
$\Re(A_0^* A_S)$	$8\sqrt{3}\,P_1^0\,(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0})$
$\Im(A_0^* A_S)$	0
$\Re(A_\parallel^* A_S)$	$+6\sqrt{2}\frac{1}{\sqrt{5}}\,P_1^1\,Y_{2,+1}$
$\Im(A_\parallel^* A_S)$	$\pm 6\sqrt{2}\,P_1^1\,Y_{1,-1}$
$\Re(A_\perp^* A_S)$	$\pm 6\sqrt{2}\,P_1^1\,Y_{1,+1}$
$\Im(A_\perp^* A_S)$	$+6\sqrt{2}\frac{1}{\sqrt{5}}\,P_1^1\,Y_{2,-1}$

can be combined into functions corresponding to the transversity amplitudes. The amplitude for the K^+K^- S-wave is the same in both bases ($H_S = A_S$). The substitution is given by [35]

$$\begin{aligned} H_0 &= A_0 \\ H_\pm &= \tfrac{1}{\sqrt{2}}(A_\parallel \pm A_\perp) \,. \end{aligned} \tag{A.16}$$

The differences between contributions from $\beta = \pm 1$ are distinguished with $\pm$ and $\mp$ signs in the tables. Because the decay of the J/ψ is not a weak process it conserves parity, which results in equal helicity/transversity amplitudes for $\beta = +1$ and $\beta = -1$. Since the muon helicities are not measured in the experiment, the two contributions are added, resulting in cancellation of the terms with opposite sign.

Both Legendre polynomials and real-valued spherical harmonics can be expressed in terms of sines and cosines of the angles. Tables A.3 and A.4 show the angular functions in terms of sines and cosines that correspond to the functions in Tables A.1 and A.2, respectively.

Table A.3: *Angular functions for the $B_s^0 \to J/\psi\phi$ decay expressed in terms of sines and cosines in helicity angles. Functions are shown for $\beta = \pm 1$. Top: functions in the helicity basis. Bottom: functions in the transversity basis.*

amplitudes	$f(\Omega) \times \frac{32\pi}{9}$
$\lvert H_0\rvert^2$	$2\,\cos^2\theta_K\,\sin^2\theta_\mu$
$\lvert H_+\rvert^2$	$\frac{1}{2}\,\sin^2\theta_K\,(1 \pm \cos\theta_\mu)^2$
$\lvert H_-\rvert^2$	$\frac{1}{2}\,\sin^2\theta_K\,(1 \mp \cos\theta_\mu)^2$
$\Re(H_0^* H_+)$	$\pm \sin 2\theta_K\,\sin\theta_\mu\,(1 \pm \cos\theta_\mu)\cos\varphi_h$
$\Im(H_0^* H_+)$	$\mp \sin 2\theta_K\,\sin\theta_\mu\,(1 \pm \cos\theta_\mu)\sin\varphi_h$
$\Re(H_0^* H_-)$	$\mp \sin 2\theta_K\,\sin\theta_\mu\,(1 \mp \cos\theta_\mu)\cos\varphi_h$
$\Im(H_0^* H_-)$	$\mp \sin 2\theta_K\,\sin\theta_\mu\,(1 \mp \cos\theta_\mu)\sin\varphi_h$
$\Re(H_+^* H_-)$	$-\sin^2\theta_K\,\sin^2\theta_\mu\,\cos 2\varphi_h$
$\Im(H_+^* H_-)$	$-\sin^2\theta_K\,\sin^2\theta_\mu\,\sin 2\varphi_h$
$\lvert A_0\rvert^2$	$2\,\cos^2\theta_K\,\sin^2\theta_\mu$
$\lvert A_\parallel\rvert^2$	$\sin^2\theta_K\,(1 - \sin^2\theta_\mu\,\cos^2\varphi_h)$
$\lvert A_\perp\rvert^2$	$\sin^2\theta_K\,(1 - \sin^2\theta_\mu\,\sin^2\varphi_h)$
$\Re(A_0^* A_\parallel)$	$+\frac{1}{\sqrt{2}}\,\sin 2\theta_K\,\sin 2\theta_\mu\,\cos\varphi_h$
$\Im(A_0^* A_\parallel)$	$\mp\sqrt{2}\,\sin 2\theta_K\,\sin\theta_\mu\,\sin\varphi_h$
$\Re(A_0^* A_\perp)$	$\pm\sqrt{2}\,\sin 2\theta_K\,\sin\theta_\mu\,\cos\varphi_h$
$\Im(A_0^* A_\perp)$	$-\frac{1}{\sqrt{2}}\,\sin 2\theta_K\,\sin 2\theta_\mu\,\sin\varphi_h$
$\Re(A_\parallel^* A_\perp)$	$\pm 2\,\sin^2\theta_K\,\cos\theta_\mu$
$\Im(A_\parallel^* A_\perp)$	$+\sin^2\theta_K\,\sin^2\theta_\mu\,\sin 2\varphi_h$

Table A.4: *Angular functions for the $B_s^0 \to J/\psi K^+K^-$ decay with a K^+K^- S-wave and the $B_s^0 \to J/\psi\phi$ and K^+K^- S-wave interference expressed in terms of sines and cosines in helicity angles. Functions are shown for $\beta = \pm 1$. Top: functions in the helicity basis. Bottom: functions in the transversity basis.*

amplitudes	$f(\Omega) \times \frac{32\pi}{9}$
$\lvert H_S\rvert^2$	$\frac{2}{3}\,\sin^2\theta_\mu$
$\Re(H_0^* H_S)$	$\frac{4}{3}\sqrt{3}\,\cos\theta_K\,\sin^2\theta_\mu$
$\Im(H_0^* H_S)$	0
$\Re(H_+^* H_S)$	$\pm\frac{2}{3}\sqrt{3}\,\sin\theta_K\,\sin\theta_\mu\,(1 \pm \cos\theta_\mu)\,\cos\varphi_h$
$\Im(H_+^* H_S)$	$\pm\frac{2}{3}\sqrt{3}\,\sin\theta_K\,\sin\theta_\mu\,(1 \pm \cos\theta_\mu)\,\sin\varphi_h$
$\Re(H_-^* H_S)$	$\mp\frac{2}{3}\sqrt{3}\,\sin\theta_K\,\sin\theta_\mu\,(1 \mp \cos\theta_\mu)\,\cos\varphi_h$
$\Im(H_-^* H_S)$	$\pm\frac{2}{3}\sqrt{3}\,\sin\theta_K\,\sin\theta_\mu\,(1 \mp \cos\theta_\mu)\,\sin\varphi_h$
$\lvert A_S\rvert^2$	$\frac{2}{3}\,\sin^2\theta_\mu$
$\Re(A_0^* A_S)$	$\frac{4}{3}\sqrt{3}\,\cos\theta_K\,\sin^2\theta_\mu$
$\Im(A_0^* A_S)$	0
$\Re(A_\parallel^* A_S)$	$+\frac{1}{3}\sqrt{6}\,\sin\theta_K\,\sin 2\theta_\mu\,\cos\varphi_h$
$\Im(A_\parallel^* A_S)$	$\pm\frac{2}{3}\sqrt{6}\,\sin\theta_K\,\sin\theta_\mu\,\sin\varphi_h$
$\Re(A_\perp^* A_S)$	$\pm\frac{2}{3}\sqrt{6}\,\sin\theta_K\,\sin\theta_\mu\,\cos\varphi_h$
$\Im(A_\perp^* A_S)$	$+\frac{1}{3}\sqrt{6}\,\sin\theta_K\,\sin 2\theta_\mu\,\sin\varphi_h$

References

[1] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. Lett. **114** (2015) 041801, arXiv:1411.3104.

[2] R. Aaij *et al.*, LHCb-ANA-2014-039 (internal LHCb note) (2014).

[3] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600; R. Aaij *et al.*, LHCb-ANA-2012-067 (internal LHCb note) (2013).

[4] S. Glashow, Nucl. Phys. **22** (1961) 579; S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264; A. Salam, Conf. Proc. **C680519** (1968) 367.

[5] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. **B47** (1973) 365.

[6] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321; P. W. Higgs, Phys. Lett. **12** (1964) 132; P. W. Higgs, Phys. Rev. Lett. **13** (1964) 508; G. Guralnik, C. Hagen, and T. Kibble, Phys. Rev. Lett. **13** (1964) 585.

[7] ATLAS Collaboration, G. Aad *et al.*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214; CMS Collaboration, S. Chatrchyan *et al.*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[8] D. Gross and F. Wilczek, Phys. Rev. Lett. **30** (1973) 1343; H. D. Politzer, Phys. Rev. Lett. **30** (1973) 1346.

[9] N. Cabibbo, Phys. Rev. Lett. **10** (1963) 531.

[10] S. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. **D2** (1970) 1285.

[11] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49** (1973) 652.

[12] O. W. Greenberg, Phys. Rev. Lett. **89** (2002) 231602, arXiv:hep-ph/0201258.

[13] S. Weinberg, Phys. Rev. **D13** (1976) 974; L. Susskind, Phys. Rev. **D20** (1979) 2619; G. 't Hooft, NATO Adv. Study Inst. Ser. B Phys. **59** (1980) 101.

[14] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003; SNO Collaboration, Q. Ahmad *et al.*, Phys. Rev. Lett. **89** (2002) 011301, arXiv:nucl-ex/0204008; KamLAND Collaboration, K. Eguchi *et al.*, Phys. Rev. Lett. **90** (2003) 021802, arXiv:hep-ex/0212021; DAYA-BAY Collaboration, F. An *et al.*, Phys. Rev. Lett. **108** (2012) 171803, arXiv:1203.1669.

[15] E. Majorana, Nuovo Cim. **14** (1937) 171.

[16] P. Minkowski, Phys. Lett. **B67** (1977) 421.

[17] A. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[18] M. Gavela, P. Hernandez, J. Orloff, and O. Pene, Mod. Phys. Lett. **A9** (1994) 795, arXiv:hep-ph/9312215; P. Huet and E. Sather, Phys. Rev. **D51** (1995) 379, arXiv:hep-ph/9404302; M. Gavela *et al.*, Nucl. Phys. **B430** (1994) 382, arXiv:hep-ph/9406289.

[19] V. Kuzmin, V. Rubakov, and M. Shaposhnikov, Phys. Lett. **B155** (1985) 36; M. Fukugita and T. Yanagida, Phys. Lett. **B174** (1986) 45.

[20] WMAP, G. Hinshaw *et al.*, Astrophys. J. Suppl. **208** (2013) 19, arXiv:1212.5226.

[21] Y. Golfand and E. Likhtman, JETP Lett. **13** (1971) 323; D. Volkov and V. Akulov, Phys. Lett. **B46** (1973) 109; J. Wess and B. Zumino, Nucl. Phys. **B70** (1974) 39.

[22] J. L. Feng, Ann. Rev. Nucl. Part. Sci. **63** (2013) 351, arXiv:1302.6587.

[23] Y. Nir and D. J. Silverman, Nucl. Phys. **B345** (1990) 301; D. Silverman, Phys. Rev. **D58** (1998) 095006, arXiv:hep-ph/9806489; P. Ball and R. Fleischer, Phys. Lett. **B475** (2000) 111, arXiv:hep-ph/9912319; I. Dunietz, R. Fleischer, and U. Nierste, Phys. Rev. **D63** (2001) 114015, arXiv:hep-ph/0012219.

[24] A. J. Buras, PoS **EPS-HEP2009** (2009) 024, arXiv:0910.1032.

[25] C.-W. Chiang *et al.*, JHEP **1004** (2010) 031, arXiv:0910.2929; A. Datta and S. Khalil, Phys. Rev. **D80** (2009) 075006, arXiv:0905.2105.

[26] B. Pontecorvo, Sov. Phys. JETP **6** (1957) 429; B. Pontecorvo, Sov. Phys. JETP **7** (1958) 172; Z. Maki, M. Nakagawa, and S. Sakata, Prog. Theor. Phys. **28** (1962) 870; B. Pontecorvo, Sov. Phys. JETP **26** (1968) 984.

[27] L. Wolfenstein, Phys. Rev. Lett. **51** (1983) 1945; L.-L. Chau and W.-Y. Keung, Phys. Rev. Lett. **53** (1984) 1802; A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, Phys. Rev. **D50** (1994) 3433, arXiv:hep-ph/9403384.

[28] CKMfitter Group, J. Charles *et al.*, Eur. Phys. J. **C41** (2005) 1, arXiv:hep-ph/0406184, and updates at http://ckmfitter.in2p3.fr/.

[29] UTfit Collaboration, M. Bona *et al.*, JHEP **0507** (2005) 028, arXiv:hep-ph/0501199, and updates at http://www.utfit.org/.

[30] S. Ali, *Visualizations of CKM unitarity triangles and helicity angles* (private communication).

[31] C. Jarlskog, Phys. Rev. Lett. **55** (1985) 1039.

[32] Heavy Flavor Averaging Group, Y. Amhis *et al.*, arXiv:1207.1158, and updates at http://www.slac.stanford.edu/xorg/hfag/.

[33] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. Lett. **110** (2013) 221601, arXiv:1304.6173.

[34] BaBar Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87** (2001) 091801, arXiv:hep-ex/0107013; Belle Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.

[35] A. S. Dighe, I. Dunietz, H. J. Lipkin, and J. L. Rosner, Phys. Lett. **B369** (1996) 144, arXiv:hep-ph/9511363.

[36] A. S. Dighe, I. Dunietz, and R. Fleischer, Eur. Phys. J. **C6** (1999) 647, arXiv:hep-ph/9804253.

[37] S. Stone and L. Zhang, Phys. Rev. **D79** (2009) 074024, arXiv:0812.2832.

[38] S. Faller, R. Fleischer, and T. Mannel, Phys. Rev. **D79** (2009) 014005, arXiv:0810.4248; B. Bhattacharya, A. Datta, and D. London, Int. J. Mod. Phys. **A28** (2013) 1350063, arXiv:1209.1413.

[39] D0 Collaboration, V. M. Abazov *et al.*, Phys. Rev. **D85** (2012) 032006, arXiv:1109.3166.

[40] CDF Collaboration, T. Aaltonen *et al.*, Phys. Rev. Lett. **109** (2012) 171802, arXiv:1208.2967.

[41] ATLAS Collaboration, G. Aad *et al.*, JHEP **1212** (2012) 072, arXiv:1208.0572; ATLAS Collaboration, ATLAS-CONF-2013-039 (2013).

[42] LHCb collaboration, R. Aaij *et al.*, Phys. Lett. **B736** (2014) 186, arXiv:1405.4140.

[43] CMS Collaboration, CMS-PAS-BPH-11-006 (2012).

[44] A. Lenz and U. Nierste, JHEP **0706** (2007) 072, arXiv:hep-ph/0612167; A. Lenz and U. Nierste, arXiv:1102.4274.

[45] Particle Data Group, J. Beringer *et al.*, Phys. Rev. **D86** (2012) 010001.

[46] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. **D87** (2013) 072004, arXiv:1302.1213.

[47] L. Evans and P. Bryant, JINST **3** (2008) S08001.

[48] M. Pojer, EPJ Web Conf. **60** (2013) 01002.

[49] LHCb collaboration, R. Aaij *et al.*, JHEP **08** (2013) 117, arXiv:1306.3663.

[50] LHCb collaboration, R. Aaij *et al.*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.

[51] LHCb collaboration, LHCb-CONF-2012-033 (2012).

[52] D. van Eijk, *Ageing and the Decay of Beauty*, PhD thesis, VU University Amsterdam, 2012, CERN-THESIS-2012-137.

[53] W. D. Hulsbergen, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[54] R. Aaij *et al.*, JINST **8** (2013) P04022, arXiv:1211.3055.

[55] LHCb collaboration, A. A. Alves Jr. *et al.*, JINST **3** (2008) S08005.

[56] LHCb collaboration, R. Aaij *et al.*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[57] V. F. Weisskopf and E. P. Wigner, Z. Phys. **63** (1930) 54; V. F. Weisskopf and E. Wigner, Z. Phys. **65** (1930) 18; T. Lee, R. Oehme, and C.-N. Yang, Phys. Rev. **106** (1957) 340.

[58] LHCb collaboration, R. Aaij *et al.*, New J. Phys. **15** (2013) 053021, arXiv:1304.4741.

[59] L. Zhang and S. Stone, Phys. Lett. **B719** (2013) 383, arXiv:1212.6434.

[60] M. Jacob and G. Wick, Annals Phys. **7** (1959) 404.

[61] S. U. Chung, CERN Yellow Report **71-8** (1971); J. D. Richman, Caltech Preprint **CALT-68-1148** (1984); R. K. Kutschke, *An Angular Distribution Cookbook*, internal CLEO note, 1996.

[62] F. Azfar *et al.*, JHEP **1011** (2010) 158, arXiv:1008.4283.

[63] S. M. Flatté, Phys. Lett. **B63** (1976) 224.

[64] Y. Xie, P. Clarke, G. Cowan, and F. Muheim, JHEP **0909** (2009) 074, arXiv:0908.3627.

[65] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. Lett. **108** (2012) 241801, arXiv:1202.4717.

[66] R. Aaij, *Triggering on CP Violation*, PhD thesis, VU University Amsterdam, 2015, CERN-THESIS-2015-102.

[67] M. Pivk and F. R. Le Diberder, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083; Y. Xie, arXiv:0905.0724.

[68] D. Martínez Santos and F. Dupertuis, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.

[69] T. du Pree, *Search for a Strange Phase in Beautiful Oscillations*, PhD thesis, VU University Amsterdam, 2010, CERN-THESIS-2010-124.

[70] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. Lett. **108** (2012) 101803, arXiv:1112.3183; R. Aaij *et al.*, LHCb-ANA-2011-036 (internal LHCb note) (2011).

[71] LHCb collaboration, R. Aaij *et al.*, JHEP **10** (2013) 183, arXiv:1308.1428.

[72] LHCb collaboration, R. Aaij *et al.*, Phys. Rev. Lett. **113** (2014) 211801, arXiv:1409.4619.

[73] CMS Collaboration, CMS-PAS-BPH-13-012 (2014).

[74] ATLAS, G. Aad *et al.*, Phys. Rev. **D90** (2014), no. 5 052007, arXiv:1407.1796.

[75] X. Liu, W. Wang, and Y. Xie, Phys. Rev. **D89** (2014), no. 9 094010, arXiv:1309.0313.

[76] K. De Bruyn and R. Fleischer, JHEP **03** (2015) 145, arXiv:1412.6834.

[77] LHCb Collaboration, CERN-LHCC-2011-001 (2011).

[78] LHCb collaboration, R. Aaij *et al.*, and A. Bharucha *et al.*, Eur. Phys. J. **C73** (2013) 2373, arXiv:1208.3355.

[79] W. Altmannshofer *et al.*, JHEP **0901** (2009) 019, arXiv:0811.1214.

Summary

Measurement of CP Violation in Mixing and Decay of Strange Beauty Mesons

Despite its precise and accurate description of elementary-particle interactions, the Standard Model of Particle Physics has several shortcomings. To find a more complete description of nature, particle interactions are tested for deviations from Standard Model predictions, which would indicate how to extend the model. The LHCb experiment at CERN's Large Hadron Collider searches for such deviations in the description of particle decays.

In particular, the LHCb experiment studies the decay of the bound states formed by beauty and strange quarks, or "strange beauty mesons". The combination of antibeauty and strange is a B_s^0 meson, beauty and antistrange form the corresponding antiparticle, denoted by $\overline{B}_s^0$.

An important feature of these particles is that they can turn into each other, which creates a mixed system of a particle and its antiparticle. Starting with a B_s^0 meson, the particle evolves and can be either a B_s^0 or a $\overline{B}_s^0$ meson at the time it decays into other particles. Similarly, there is a probability that a particle created as a $\overline{B}_s^0$ meson decays as a B_s^0 meson.

A particularly interesting mode of decay is that into a J/ψ meson and a $\phi(1020)$ meson, which occurs for both B_s^0 and $\overline{B}_s^0$. For this mode there are two possible decay paths for each of the two initial particles. In one path the initial particle first turns into its antiparticle an then decays into the $J/\psi\,\phi$ final state and in the other path the initial particle decays directly into this state.

This mixing and decay process is predicted to be almost identical for B_s^0 and $\overline{B}_s^0$ within the Standard Model framework. Measurements indicate that the difference in the rates for the transitions from B_s^0 to $\overline{B}_s^0$ and vice versa are

very small. Also the rates of the decays into the $J/\psi\,\phi$ state are expected to be nearly equal. This equivalence between matter and antimatter is known as *CP symmetry*.

Contributions from particle interactions that are not described by the Standard Model may increase the amount of violation of CP symmetry in the $B_s^0 \to J/\psi\,\phi$ process. In particular, a difference between the complex phases of the probability amplitudes for the B_s^0–$\overline{B}_s^0$ mixing transitions may be introduced. In general this does not lead to CP violation, since the rate of a process only depends on the magnitude of the corresponding probability amplitude. In this case, however, the amplitude is a sum of the interfering contributions from the two decay paths. The relative phases of these contributions do affect the magnitude of the sum, leading to an observable difference between the $B_s^0 \to J/\psi\,\phi$ and $\overline{B}_s^0 \to J/\psi\,\phi$ processes.

This type of CP violation in the interference between decay paths with and without mixing is measured by examining the distribution of the time between the production and the decay of B_s^0 and/or $\overline{B}_s^0$ mesons. Without CP violation, this distribution is given by the sum of two exponential contributions with slightly different mean lifetimes. CP violation introduces an oscillation on top of this exponential shape with an amplitude of opposite sign for initial B_s^0 and $\overline{B}_s^0$ mesons.

In the LHCb experiment, B_s^0 and $\overline{B}_s^0$ mesons are abundantly produced in roughly equal amounts in the proton–proton collisions of the Large Hadron Collider. Decays into $J/\psi\,\phi$ followed by decays of the J/ψ meson into two muons and the ϕ meson into two kaons are selected by requiring the signature of these muons and kaons in the detector is compatible with this decay chain.

The produced B_s^0 and $\overline{B}_s^0$ mesons have a mean lifetime of about 1.5 ps, which means typical distances of several millimetres are covered before their decay. These distances are measured by determining the positions of the proton–proton collision and the common point of origin of the muons and kaons from the decay. Also including the measurement of the combined momentum of the decay particles, the time between production and decay of the original meson is inferred.

The shape of the decay-time distribution is modelled and the resulting model is fitted to the measured distribution to determine the values of its parameters. Parameters that describe CP violation determine the amplitude of the oscillation in decay time. The frequency of the oscillation and the lifetimes of the two exponential shapes are controlled by parameters that describe the

coupled B_s^0–$\overline{B}_s^0$ system.

Different types of CP violation are included in the decay model, which are measured individually for the three different angular-momentum states of the $J/\psi\,\phi$ system for the first time in this measurement. Small differences between the contributions of these states are expected, which become important in a precision measurement of CP violation. The different contributions are separated by including the measurement of the angles between the momentum directions of the four final-state particles. This results in a four-dimensional distribution of the decay time and three decay angles.

To describe the measured distribution of these variables, experimental effects such as detection and selection efficiencies and finite measurement resolutions are included in the model. Also the fact that the measured distribution is a sum of B_s^0 and $\overline{B}_s^0$ decays is taken into account. Uncertainties in the estimates of these experimental effects lead to systematic uncertainties in the estimated decay-time and CP-violation parameters, in addition to the statistical uncertainties associated with the size of the sample of decays.

The measured distribution of time and angles is constructed from roughly ninety thousand decays, collected in the years 2011 and 2012. Estimates of the parameter values with these data are compatible with Standard Model predictions, given the experimental uncertainties. These results show that potential non-Standard Model contributions to the $B_s^0 \to J/\psi\,\phi$ mixing and decay process must be smaller than the current experimental precision.

An improvement in precision of an order of magnitude is expected with future data from the LHCb experiment, which provides new opportunities for measuring deviations from the Standard Model with $B_s^0 \to J/\psi\,\phi$ decays. The measurement with this larger sample of decays requires some improvements in the experimental procedure, to keep systematic uncertainties smaller than the statistical uncertainties.

Adopting the new strategy of measuring CP violation individually for the different angular-momentum states of the $J/\psi\,\phi$ system would enable interpretation of future precision measurements within a framework of measurements and theoretical calculations of several different meson decays. Such a combined analysis is likely to be required to overcome limitations in the theoretical predictions of CP-violation parameters. Combining these experimental and theoretical tools, the measurement of CP violation in mixing and decay of strange beauty mesons has the potential to continue playing an important role in the search for a more complete description of nature.

Samenvatting

Meting van CP-schending in het vermengen en vervallen van vreemde schoonheidmesonen

Ondanks dat het Standaardmodel van de deeltjesfysica een juiste en nauwkeurige beschrijving geeft van de interacties tussen elementaire deeltjes, heeft het meerdere tekortkomingen. Om een completere beschrijving van de natuur te vinden, worden deeltjesinteracties getest op afwijkingen van voorspellingen door het Standaardmodel, welke een indicatie zouden geven van de manier waarop het model moet worden uitgebreid. Het LHCb experiment bij de "Large Hadron Collider" van CERN zoekt naar dit soort afwijkingen in de beschrijving van het verval van deeltjes.

In het bijzonder wordt in het LHCb experiment het verval bestudeerd van gebonden toestanden van schoonheid- en vreemdquarks, of "vreemde schoonheidmesonen". De combinatie van antischoonheid en vreemd is een B_s^0 meson, terwijl schoonheid en antivreemd het bijbehorende antideeltje vormen, aangeduid met $\overline{B}_s^0$.

Een belangrijke eigenschap van deze deeltjes is dat ze in elkaar kunnen overgaan. Dit geeft een gemengd systeem van een deeltje en zijn antideeltje. Een B_s^0 meson evolueert in de tijd en kan of een B_s^0 of een $\overline{B}_s^0$ meson zijn op het moment dat het vervalt in andere deeltjes. Op dezelfde manier kan een deeltje dat is geproduceerd als $\overline{B}_s^0$ meson vervallen als een B_s^0 meson.

Met name de modus van verval in een J/ψ meson en een $\phi(1020)$ meson is interessant. Deze modus is mogelijk voor zowel B_s^0 als $\overline{B}_s^0$. Er zijn daardoor twee mogelijke vervalpaden voor ieder van de twee aanvankelijke deeltjes. In een van de paden gaat het vervallende deeltje eerst over in zijn antideeltje en vervalt vervolgens naar de $J/\psi\,\phi$ eindtoestand. In het andere pad vervalt het aanvankelijke deeltje direct naar deze toestand.

Het Standaardmodel voorspelt dat dit proces van vermengen en vervallen bijna gelijk verloopt voor B_s^0 en $\overline{B}_s^0$. Metingen laten zien dat het verschil in de snelheden van de overgangen van B_s^0 naar $\overline{B}_s^0$ en vice versa erg klein is. Ook de snelheden van de vervallen naar de $J/\psi\,\phi$ toestand zijn naar verwachting bijna gelijk. Deze overeenkomst tussen materie en antimaterie staat bekend als *CP-symmetrie*.

Bijdragen van deeltjesinteracties die niet worden beschreven door het Standaardmodel zouden de mate waarin CP-symmetrie wordt geschonden in het $B_s^0 \to J/\psi\,\phi$ proces kunnen vergroten. Met name zou er een verschil kunnen ontstaan tussen de complexe fases van de waarschijnlijkheidsamplitudes van de B_s^0–$\overline{B}_s^0$ mengovergangen. In het algemeen leidt dit niet tot CP-schending, omdat de waarschijnlijkheid van een proces alleen afhangt van de absolute waarde van de bijbehorende waarschijnlijkheidsamplitude. In dit geval is de amplitude echter een som van de interfererende bijdragen van de twee vervalpaden. De relatieve fases van deze bijdragen hebben wel invloed op de absolute waarde van de som en dit leidt tot een waarneembaar verschil tussen de $B_s^0 \to J/\psi\,\phi$ en $\overline{B}_s^0 \to J/\psi\,\phi$ processen.

Deze vorm van CP-schending in de interferentie tussen vervalpaden met en zonder vermenging wordt gemeten door de verdeling van de tijd tussen de productie en het verval van B_s^0 en/of $\overline{B}_s^0$ mesonen te bestuderen. Zonder CP-schending is deze verdeling gegeven door de som van twee exponentiële bijdragen met een klein verschil in gemiddelde levensduur. CP-schending introduceert een oscillatie op deze exponentiële vorm met een amplitude die een tegengesteld teken heeft voor aanvankelijke B_s^0 en $\overline{B}_s^0$ mesonen.

In het LHCb experiment worden B_s^0 en $\overline{B}_s^0$ mesonen in overvloed en in gelijke hoeveelheden geproduceerd in de proton–proton botsingen van de "Large Hadron Collider". Vervallen naar $J/\psi\,\phi$ gevolgd door de vervallen van het J/ψ meson naar twee muonen en het ϕ meson naar twee kaonen worden geselecteerd door te eisen dat het patroon van deze muonen en kaonen in de detector overeenkomt met deze vervalketen.

De geproduceerde B_s^0 en $\overline{B}_s^0$ mesonen hebben een gemiddelde levensduur van ongeveer 1,5 ps, waardoor ze typische afstanden van enkele millimeters afleggen voor hun verval. Deze afstanden worden gemeten door de posities te bepalen van de proton–proton botsing en het gezamenlijke punt van oorsprong van de muonen en kaonen. Door ook de meting van de gecombineerde impuls van de vervaldeeltjes mee te nemen, kan de tijd tussen productie en verval van het oorspronkelijke meson worden afgeleid.

De vorm van de verdeling van vervaltijden wordt gemodelleerd en het resulterende model wordt gefit aan de gemeten verdeling om de waarden van parameters in het model te bepalen. Parameters die CP-schending beschrijven, bepalen de amplitude van de oscillatie in de vervaltijd. De frequentie van de oscillatie en de levensduren van de twee exponentiële vervallen worden bepaald door parameters die het gekoppelde B_s^0–$\overline{B}_s^0$ systeem beschrijven.

Het vervalmodel beschrijft verschillende vormen van CP-schending, die in deze meting voor het eerst individueel worden gemeten voor de drie verschillende impulsmomenttoestanden van het $J/\psi\,\phi$ systeem. Naar verwachting bestaan er kleine verschillen tussen de bijdragen van deze toestanden, die belangrijk worden in een precisiemeting van CP-schending. De bijdragen worden gescheiden door de meting van de hoeken tussen de impulsrichtingen van de vier deeltjes in de eindtoestand mee te nemen in het model. Dit geeft een vierdimensionale verdeling van de vervaltijd en drie vervalhoeken.

Om de gemeten verdeling van deze variabelen te kunnen beschrijven, worden experimentele effecten als detectie- en selectie-efficiënties en resoluties van metingen meegenomen in het model. Ook wordt rekening gehouden met het feit dat de waargenomen verdeling een som is van B_s^0 en $\overline{B}_s^0$ vervallen. Onzekerheden in de inschattingen van deze experimentele effecten leiden tot systematische onzekerheden in de parameterschattingen, naast de statistische onzekerheden die worden geassocieerd met de grootte van de verzameling vervallen.

De gemeten verdeling van tijd en hoeken is opgebouwd uit ongeveer negentigduizend vervallen, die verzameld zijn in de jaren 2011 en 2012. Schattingen van de parameterwaarden met deze data komen overeen met de voorspellingen van het Standaardmodel, gegeven de experimentele onzekerheden. Deze resultaten laten zien dat potentiële bijdragen van buiten het Standaardmodel aan het vermengen en vervallen in het $B_s^0 \to J/\psi\,\phi$ proces kleiner zijn dan de huidige experimentele nauwkeurigheid.

Een verbetering in nauwkeurigheid met een orde van grootte wordt verwacht met toekomstige data van het LHCb experiment. Dit biedt nieuwe mogelijkheden voor het meten van afwijkingen van voorspellingen van het Standaardmodel met $B_s^0 \to J/\psi\,\phi$ vervallen. De meting met deze grotere verzameling vervallen vereist enige verbetering in de experimentele procedure om systematische onzekerheden kleiner te houden dan statistische onzekerheden.

Het overnemen van de nieuwe strategie van het individueel meten van

CP-schending voor de verschillende impulsmomenttoestanden van het $J/\psi\,\phi$ systeem, zou het mogelijk maken om toekomstige precisiemetingen te interpreteren in een raamwerk van metingen en theoretische berekeningen van verschillende meson vervallen. Een dergelijke gecombineerde analyse wordt waarschijnlijk nodig om te kunnen omgaan met beperkingen in theoretische voorspellingen van CP-schendingparameters. Door het combineren van deze experimentele en theoretische analyses kan de meting van CP-schending in het vermengen en vervallen van vreemde schoonheidmesonen potentieel een belangrijke rol blijven spelen in de zoektocht naar een completere beschrijving van de natuur.