$B^0_s \rightarrow J/\psi \phi$ Analysis

The $B^0_s$ mixing phase $\phi_s$ is a sensitive probe for New Physics. In this chapter, the measurement of $\phi_s$ is presented. The decay of $B^0_s$ mesons to the final state $J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ was first observed by the CDF collaboration [55] and the branching ratio has been measured by both the CDF collaboration [56] and the Belle collaboration [57].

The CDF collaboration and the D0 collaboration (both situated at the Tevatron accelerator at Fermilab) have published results on the measurement of $\phi_s$ using datasets of $5.2fb^{-1}$ [58] and $8fb^{-1}$ [59] of integrated luminosity, respectively. The LHCb collaboration has published its first results on $\phi_s$ using $0.37fb^{-1}$ of integrated luminosity [60]. Although the LHCb dataset corresponds to a smaller integrated luminosity, the number of $B^0_s \rightarrow J/\psi \phi$ candidates is larger in LHCb due to the larger production cross section at the LHC, as compared to the Tevatron. The dataset for the analysis presented here is three times larger than the one in [60] and corresponds to $1.0fb^{-1}$ of integrated luminosity. As a result, this analysis leads to the most accurate measurement of $\phi_s$ to date.

6.1 Candidate Selection

The analysis presented in this chapter is based on $B^0_s \rightarrow J/\psi \phi$ candidates recorded during the 2011 LHC running period at a center-of-mass energy of $\sqrt{s}=7$ TeV. The event selection of $B^0_s \rightarrow J/\psi \phi$ candidates is performed in three steps: first the online trigger selection (L0, HLT1 and HLT2), followed by the stripping pre-selection and finally the offline selection.

6.1.1 Trigger Selection

The first stage in the extraction of $B^0_s \rightarrow J/\psi \phi$ candidates is the trigger selection. As discussed in Sec.2.6, the level-0 (L0) hardware trigger selects muon candidates with large transverse momentum $p_T$. Subsequently, the software trigger selects events in two stages: HLT1 and HLT2. In this analysis, only those events that are explicitly triggered on the two muons of the $J/\psi$ candidate are used.
In the HLT1\[1\] a partial event reconstruction is performed and $\mu^+\mu^-$ candidates are required to have two well-identified muons whose trajectories have a distance of closest approach to each other smaller than 0.2 mm, a good vertex fit, i.e. $\chi^2_{\text{vertex}}/n\text{DoF} < 25$, and an invariant mass of at least 2.7 GeV/c\(^2\) \[61\]. Note that there are no explicit requirements that affect the decay-time distribution, hence no decay-time acceptance is introduced in this HLT1 line\[2\].

In the HLT2\[3\] the $J/\psi$ candidates reconstructed from di-muon pairs are required to have a minimum $p_T$ of 1.5 GeV/c, and only candidates with a reconstructed mass within 120 MeV/c\(^2\) from the nominal $J/\psi$ meson mass (3096 MeV/c\(^2\) \[23\]) are accepted: $2976 \text{ MeV}/c^2 < m(\mu^+\mu^-) < 3216 \text{ MeV}/c^2$. In addition, two cuts are applied to reject short-lived candidates where the largest background contribution is expected: for each muon, there is a cut on the $\chi^2$ of the impact parameter (IP) with respect to the PV, i.e. $\chi^2_{\text{IP}}(\mu) > 9$, and in addition there is a di-muon vertex separation cut with respect to the closest PV (decay length significance (DLS) $> 3$ \[61\]). Note that these cuts affect the decay-time distribution of the $B^0_s \to J/\psi\phi$ event candidates, which will introduce a non-trivial efficiency as a function of decay time in the analysis.

### 6.1.2 Stripping and Selection Cuts

After offline reconstruction, a so-called stripping procedure is applied to obtain manageable dataset\[4\]. The final selection is applied to these so-called stripped datasets to obtain the sample of $B^0_s \to J/\psi\phi$ candidates used for the determination of $\phi_s$.

In the final dataset, muon candidates are each required to have $p_T > 0.5 \text{ GeV}/c$. Muons are distinguished from pions by requiring a difference of the particle ID log-likelihood $LL(\mu) - LL(\pi) \equiv DLL(\mu\pi) > 0$ (see Sec. 2.5.1). The corresponding $J/\psi$ candidates are created from pairs of oppositely charged muons that have a common vertex which satisfies $\chi^2_{\text{vertex}}/n\text{DoF}(J/\psi) < 16$. The reconstructed di-muon mass is required to be in the range $3030 \text{ MeV}/c^2 < m(\mu^+\mu^-) < 3150 \text{ MeV}/c^2$.

The $\phi$ meson candidates are selected by requiring two oppositely charged kaon candidates with $DLL(K\pi) > 0$, originating from a common vertex with $\chi^2_{\text{vertex}}/n\text{DoF}(\phi) < 16$. In addition, the $p_T$ of the $\phi$ meson candidate is required to be larger than 1 GeV/c\(^2\). The reconstructed invariant mass should be within $12 \text{ MeV}/c^2$ from the nominal $\phi$ meson mass \[23\]: $1008 \text{ MeV}/c^2 < m(K^+K^-) < 1032 \text{ MeV}/c^2$.

$B^0_s$ candidates are selected from combinations of $J/\psi$ meson candidates and $\phi$ meson candidates with an invariant mass in the range $5200 \text{ MeV}/c^2 < m(J/\psi K^+K^-) < 5550 \text{ MeV}/c^2$. The invariant mass of the $B^0_s$ candidate is computed with the invariant mass of the $\mu^+\mu^-$ pair constrained to the nominal $J/\psi$ mass. The decay time is obtained from a kinematic fit \[52\]. This algorithm constrains the $B^0_s \to \mu^+\mu^-K^+K^-$ candidate to originate from the PV by imposing a $B^0_s$ vertex cut ($\chi^2_{\text{vertex}}/n\text{DoF}(B^0_s) < 10$), an impact parameter cut ($\chi^2_{\text{IP}}(B^0_s) < 25$) and finally a cut on the $\chi^2$ of the kinematic fit itself.

---

1. The HLT1 trigger line used in this analysis is the decay time unbiased trigger line Hlt1DiMuonHighMass.
2. Assuming that the track reconstruction in the VELO is independent of the decay time, see Sec. 6.3.3.
3. The HLT2 trigger line used in this analysis is the decay time biased trigger line Hlt2DiMuonDetachedJpsi.
4. The $B^0_s \to J/\psi\phi$ candidates used in this analysis are processed using reconstruction version Reco12 and stripping version Stripping17.
<table>
<thead>
<tr>
<th>decay mode</th>
<th>cut parameter</th>
<th>stripping</th>
<th>selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>all tracks</td>
<td>( \chi_{\text{track}}^2 / \text{nDoF} )</td>
<td>( &gt; 5 )</td>
<td>( &lt; 4 )</td>
</tr>
<tr>
<td></td>
<td>clone distance</td>
<td>( &gt; 5000 )</td>
<td></td>
</tr>
<tr>
<td>( J/\psi \to \mu^+ \mu^- )</td>
<td>( DLL(\mu\pi) )</td>
<td>( &gt; 0 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_T(\mu) )</td>
<td>( &gt; 0.5 \text{ GeV/c} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{vertex}}^2 / \text{nDoF}(J/\psi) )</td>
<td>( &lt; 16 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{DOCA}}^2(J/\psi) )</td>
<td>( &lt; 20 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( m(\mu^+ \mu^-) )</td>
<td>( \in [3010, 3170] \text{ MeV/c}^2 )</td>
<td>( \in [3030, 3150] \text{ MeV/c}^2 )</td>
</tr>
<tr>
<td>( \phi \to K^+ K^- )</td>
<td>( DLL(K\pi) )</td>
<td>( &gt; -2 )</td>
<td>( &gt; 0 )</td>
</tr>
<tr>
<td></td>
<td>( p_T(\phi) )</td>
<td>( &gt; 0.5 \text{ GeV/c} )</td>
<td>( &gt; 1 \text{ GeV/c} )</td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{track}}^2 / \text{nDoF}(K) )</td>
<td>( &lt; 4 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{vertex}}^2 / \text{nDoF}(\phi) )</td>
<td>( &lt; 16 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{DOCA}}^2(\phi) )</td>
<td>( &lt; 30 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( m(K^+ K^-) )</td>
<td>( \in [980, 1050] \text{ MeV/c}^2 )</td>
<td>( \in [1008, 1032] \text{ MeV/c}^2 )</td>
</tr>
<tr>
<td>( B^0_s \to J/\psi \phi )</td>
<td>( \chi_{\text{vertex}}^2 / \text{nDoF}(B^0_s) )</td>
<td>( &lt; 10 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{kin.fit(B+PV)}}^2 / \text{nDoF}(B^0_s) )</td>
<td>-</td>
<td>( &lt; 5 )</td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{IP}}^2(B^0_s) )</td>
<td>-</td>
<td>( &lt; 25 )</td>
</tr>
<tr>
<td></td>
<td>( \chi_{\text{IP, next}}^2(B^0_s) )</td>
<td>-</td>
<td>( &gt; 50 )</td>
</tr>
<tr>
<td></td>
<td>( m(J/\psi \phi) )</td>
<td>( \in [5200, 5550] \text{ MeV/c}^2 )</td>
<td>( \in [5200, 5550] \text{ MeV/c}^2 )</td>
</tr>
<tr>
<td></td>
<td>( t(\ast) )</td>
<td>( &gt; 0.2 \text{ ps} )</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: Stripping and selection criteria for \( B^0_s \to J/\psi \phi \) candidates. (*): the stripping decay time and offline decay time are calculated from different algorithms.

\( (\chi_{\text{kin.fit(B+PV)}}^2 / \text{nDoF} < 5) \). In case of multiple candidates per event\(^5\), only the candidate with the smallest \( \chi_{\text{kin.fit(B+PV)}}^2 / \text{nDoF} \) is kept. Finally, in order to prevent misassociation of the \( B^0_s \) candidate with the next-best PV, which would clearly result in an incorrect decay time, the \( B^0_s \) candidate must have an impact parameter \( \chi_{\text{IP}}^2 \) to the next-best PV in the event greater than 50. This enforces the \( B^0_s \) candidate to be inconsistent with coming from this next-best PV.

Only \( B^0_s \to J/\psi \phi \) candidates with a decay time within \( 0.3 < t < 14 \text{ ps} \) are considered in the analysis. The lower decay-time cut suppresses a large fraction of the prompt combinatorial background, while losing little in sensitivity to \( \phi_s \). The upper decay-time cut avoids the use of pathological events (misreconstructed or associated to a wrong PV) while losing only very few signal events. A summary of the stripping and selection cuts is given in Table 6.1

To illustrate the mass resolution and small background levels, the mass distributions of the reconstructed \( J/\psi \) and \( \phi \) meson candidates, after applying all selection criteria, are given in Fig. 6.1. The small background in the di-muon invariant-mass distribution suggests that the main background contribution originates from events containing genuine \( J/\psi \) mesons.

\(^5\)The average number of candidates per event is 1.10 with an average number of primary vertices of 2.42.
6.2 Angular Acceptance

The detector geometry and the selection criteria introduce an acceptance effect as a function of the decay angles, which must be accounted for in the angular analysis. It is estimated and corrected for by using Monte Carlo simulated data.

6.2.1 MC Dataset

The MC dataset consists of five million simulated $B^0_s \to J/\psi \phi$ events. The parameters of the decay model used to generate these events are given in Table 6.2.

Table 6.2: Decay model parameters for the simulated sample of $B^0_s \to J/\psi \phi$ events. The values for the transversity amplitudes and strong phases are taken from [63].

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_s$</td>
<td>0.02</td>
</tr>
<tr>
<td>$\Gamma_s$</td>
<td>0.6793 ps$^{-1}$</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$</td>
<td>0.060 ps$^{-1}$</td>
</tr>
<tr>
<td>$\Delta m_s$</td>
<td>17.8 ps$^{-1}$</td>
</tr>
<tr>
<td>$</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp(0)</td>
</tr>
<tr>
<td>$</td>
<td>A_\parallel(0)</td>
</tr>
<tr>
<td>$\delta_0$</td>
<td>0</td>
</tr>
<tr>
<td>$\delta_\perp$</td>
<td>-0.17</td>
</tr>
<tr>
<td>$\delta_\parallel$</td>
<td>2.50</td>
</tr>
</tbody>
</table>

Figure 6.1: (a) Invariant mass distribution of the $J/\psi$ meson candidates after applying all selection criteria. (b) Invariant mass distribution of the $\phi$ meson candidates after applying all selection criteria.

Events are processed by Sim05-Reco12-Stripping17 (also known as MC11A). The HLT configuration used is TCK 0x40760037, which is representative for the data recorded during and after summer 2011.
6.2.2 Reparameterization in Terms of Basis Functions

The distribution of \( B_0^s \rightarrow J/\psi \phi \) events in terms of decay angles as presented in Table 5.1 and Table 5.3 can be re-expressed in terms of associated Legendre polynomials \( P_i(\cos \psi) \) and spherical harmonics \( Y_{im}(\cos \theta, \phi) \), where \( \psi, \theta \) and \( \phi \) are the decay angles in the transversity frame. The reason to do so is that the orthogonality properties of these functions will simplify the implementation of the angular acceptance in the fitting procedure. The basis functions \( P_i(\cos \psi) \) and \( Y_{im}(\cos \theta, \phi) \), as well as the resulting expressions for the angular dependence of the differential decay rate are given in Appendix E.

While writing the angular-acceptance function as an expansion in terms of the same basis functions \( P_i(\cos \psi) \) and \( Y_{im}(\cos \theta, \phi) \), the coefficients of this expansion can be calculated from a MC simulated data sample as will be shown in the following section.

6.2.3 Angular-Acceptance Correction

The angular acceptance is implemented in the analysis by multiplying the signal PDF with the angular-acceptance function \( \epsilon(\cos \psi, \cos \theta, \phi) \):

\[
P_{\text{corrected}}(t, \cos \psi, \cos \theta, \phi) = P_{\text{uncorrected}}(t, \cos \psi, \cos \theta, \phi) \cdot \epsilon(\cos \psi, \cos \theta, \phi) \quad . \tag{6.1}
\]

The angular-acceptance function is written as an expansion in terms of associated Legendre polynomials \( P_i(\cos \psi) \) and spherical harmonics \( Y_{im}(\cos \theta, \phi) \):

\[
\epsilon(\vec{\Omega}) = \sum_{i,l,m} c^i_{lm} P_i(\cos \psi) Y_{lm}(\cos \theta, \phi) \quad , \tag{6.2}
\]

where \( \vec{\Omega} = (\cos \psi, \cos \theta, \phi) \). The coefficients \( c^i_{lm} \) of the expansion are calculated from simulated events as follows.

The simulated data sample is generated according to the theoretical \( B_0^s \rightarrow J/\psi \phi \) decay distribution \( g(\vec{\Omega}) \). For an efficiency \( \epsilon(\vec{\Omega}_e) \), and a general function \( h(\vec{\Omega}) \), the following equation holds (following the concept of MC integration):

\[
\int \epsilon(\vec{\Omega}) h(\vec{\Omega}) g(\vec{\Omega}) d\vec{\Omega} \simeq \frac{1}{N_{\text{gen}}} \sum_{e \in \{\text{generated}\}} \epsilon(\vec{\Omega}_e) h(\vec{\Omega}_e) = \frac{1}{N_{\text{gen}}} \sum_{e \in \{\text{accepted}\}} h(\vec{\Omega}_e) \quad . \tag{6.3}
\]

Here, the integral is rewritten as a finite sum over the generated events \( \sum_{e \in \{\text{generated}\}} \). As the event is either accepted or not accepted, the per-event efficiency \( \epsilon(\vec{\Omega}_e) \) is \( \pm 1 \) and the sum over generated events is rewritten as a sum over accepted events only \( \sum_{e \in \{\text{accepted}\}} \).

It is now argued that an appropriate choice of the function \( h(\vec{\Omega}) \) allows the calculation of the coefficients \( c^i_{lm} \) of the expansion of the angular-acceptance function. Substituting the expansion of the acceptance function from Eq.6.2 yields

\[
\int \sum_{n,j,k} c^n_{jk} P_n(\cos \psi) Y_{jk}(\cos \theta, \phi) h(\vec{\Omega}) g(\vec{\Omega}) d\vec{\Omega} = \frac{1}{N_{\text{gen}}} \sum_{e \in \{\text{accepted}\}} h(\vec{\Omega}_e) \quad . \tag{6.4}
\]
When choosing
\[ h(\vec{\Omega}) = \frac{2i + 1}{2} P_i(\cos \psi) Y_{lm}(\cos \theta, \phi) \frac{g(\vec{\Omega})}{g(\vec{\Omega})}, \]
the orthonormality of the basis functions \( P_i(\cos \psi) Y_{lm}(\cos \theta, \phi) \) (see Appendix E.4) can be used to find
\[ \int g(\vec{\Omega}) d\vec{\Omega} c_{jk}^n P_n(\cos \psi) Y_{jk}(\cos \theta, \phi) \cdot \left[ \frac{2i + 1}{2} P_i(\cos \psi) Y_{lm}(\cos \theta, \phi) \right] = c_{lm}^i, \]
and hence, using Eq. 6.4
\[ c_{lm}^i = \frac{1}{N_{\text{gen}}} \sum_{e \in \{\text{accepted}\}} \frac{2i + 1}{2} P_i(\cos \psi_e) Y_{lm}(\cos \theta_e, \phi_e) \frac{g(\vec{\Omega}_e)}{g(\vec{\Omega})}. \]

In other words, the coefficients \( c_{lm}^i \) are calculated as a finite sum over accepted Monte Carlo events of the function \( h(\vec{\Omega}_e) \), as defined in Eq. 6.5. These coefficients fully determine the shape of the acceptance function \( \epsilon(\vec{\Omega}) \). The decay distributions of the signal events are multiplied with this angular-acceptance function to correct for angular acceptance effects.

**Results**

Using the MC dataset mentioned in Sec. 6.2.1, which is assumed to be representative for the real data (see Sec. 6.7.5 for differences between the MC dataset and the real dataset) used in this analysis, the coefficients of the expansion of the angular-acceptance function are calculated and the acceptance function is projected on the three transversity angles in the upper three figures of Fig. 6.2. In this figure, 41 coefficients \( c_{lm}^i \) are used in the expansion. In addition, in the lower three figures of Fig. 6.2, the MC data is shown, together with the theoretical decay distributions of the signal events and the decay distributions corrected for angular acceptance.

**6.2.4 Comparison with Normalization-Weights Method**

Another method to correct for angular acceptance effects is to calculate ten so-called normalization weights corresponding to the ten angular functions \( f_i(\vec{\Omega}) \), using Eq. 6.3 and using these weights to normalize the PDF. The two methods can be compared and checked by reconstructing the ten normalization weights from (a subset of) the coefficients \( c_{lm}^i \). This comparison also demonstrates that only a subset of \( c_{lm}^i \) actually matters for the minimization in the final fit. The comparison of the two methods is shown in detail in Appendix F. The advantage of the expansion method compared to the normalization-weights method is that the shape of the acceptance function is known, as shown in Fig. 6.2. Another advantage of the expansion method is that no assumption is made on the similarity of the angular-acceptance function between signal and background, while in the normalization-weights method this

\(^7\)In the final analysis, only four coefficients are taken into account, since no difference in the final fit result is observed when using these four coefficients instead of 41 coefficients. The only coefficients used in the final analysis are \( c_{00}^0, c_{02}^2, c_{02}^0 \) and \( c_{20}^2 \).
assumption is made by construction. Finally, using the expansion method allows to plot differential distributions, which is impossible if only normalization weights are known.

### 6.3 Decay-Time Acceptance

#### 6.3.1 Acceptance for Small Decay Time

As mentioned in Sec. 6.1, the HLT1 trigger line used in this analysis does not explicitly affect the decay-time distribution of the $B^0_s$ candidates, whereas the HLT2 line does reject short-living $B^0_s$ candidates. To quantify this effect, events selected with a similar, prescaled, HLT2 trigger without decay-time biasing cuts are studied. A non-parametric description, i.e. a histogram, of the HLT2 decay-time acceptance is produced from the overlap between events triggered by the unbiased and biased HLT2 lines and is shown in Fig. 6.3. The structure around $t = 2\,\text{ps}$ is understood as so-called vertex splitting [65], when the secondary vertex is unjustly reconstructed as a primary vertex, with respect to which the $J/\psi$ vertex does not pass the DLS cut in the HLT2 trigger as described in Sec. 6.1.1 therefore losing the event. This effect only occurs for moderate decay times, since both at low decay times and at high decay times, no second PV will be reconstructed from the tracks originating from the secondary vertex. The drop in decay-time acceptance is also observed in MC.
simulated data. The decay-time acceptance is implemented in the analysis by multiplying the theoretical decay-time distribution with the acceptance histogram.

Figure 6.3: (a) The decay-time acceptance of the biased HLT2 trigger line, determined from the overlap between events triggered by the unbiased and biased HLT2 lines. (b) Zoom-in of the low-lifetime region. Notice the adjusted scale on the y-axis. The offline decay-time cut at 0.3 ps is indicated by the dashed line. The structure around \( t = 2 \) ps is understood as vertex splitting (see text).

6.3.2 Acceptance for Large Decay Time

From MC studies, a linear drop in acceptance is observed for large decay time, which is attributed to a reduced track-finding efficiency for tracks originating from vertices that are radially displaced from the beam axis. The acceptance is parameterized as \( \epsilon(t) = 1 + \beta t \), with \( \beta = -0.0112 \pm 0.0013 \) ps. By performing two fits to the data, one including and the other not including this decay-time acceptance, it is found that only the average lifetime \( \Gamma_s \) is affected. Therefore, in the final analysis, the acceptance for large decay time is not explicitly taken into account, but the parameter \( \Gamma_s \) is corrected analytically a posteriori. \(^8\)

6.4 Flavour Tagging

The CP-violation phase \( \phi_s \) can be measured thanks to the interference between a direct decay amplitude and a decay amplitude which includes oscillation, which results in a differ-

\(^8\) The corrected value \( \Gamma_s \) is obtained from the fitted value \( \Gamma_{s,\text{fit}} = \frac{1}{\tau_{s,\text{fit}}} \), where \( \tau_{s,\text{fit}} \) is assumed to be the average of the decay-time distribution \( \langle t \rangle \), in the following way:

\[
\frac{1}{\Gamma_{s,\text{fit}}} = \langle t \rangle = \frac{\int_0^\infty t(1 + \beta t)e^{-\Gamma_s t} \, dt}{\int_0^\infty (1 + \beta t)e^{-\Gamma_s t} \, dt} = \frac{2\beta + \Gamma_s}{\Gamma_s(\beta + \Gamma_s)} ,
\]

which yields

\[
\Gamma_s = \frac{\Gamma_{s,\text{fit}}}{2} \left( 1 - \beta/\Gamma_{s,\text{fit}} + \sqrt{1 + 6(\beta/\Gamma_{s,\text{fit}}) + (\beta/\Gamma_{s,\text{fit}})^2} \right) .
\]
ence in decay rates between $B_s^0$ and $\bar{B}_s^0$ initial states. Therefore, the identification of the initial flavour (at production) of the reconstructed $B_s^0$ candidate is a crucial element in the $\phi_s$ analysis. The flavour tag $q_T$, which takes the values -1 for $\bar{B}_s^0$, +1 for $B_s^0$ and 0 for untagged events, is assigned by combining different tagging algorithms [66]. The combination procedure assigns an estimate of the per-event mistag probability $\eta$. In this thesis, only so-called opposite-side (OS) tagging algorithms are used that exploit the following features in the decay of the other, accompanying (non-signal) $B$ hadron:

- the charge of a muon or electron with large transverse momentum produced by semileptonic $B$ decays,
- the charge of a kaon from a subsequent charmed hadron decay,
- the momentum-weighted charge of all tracks included in the inclusively reconstructed decay vertex.

These three different OS taggers are illustrated by a schematic picture of a $B_s^0 \rightarrow J/\psi \phi$ decay in Fig. 6.4. For this analysis, the so-called same-side (SS) kaon tagger that uses the charged kaons originating from the hadronization of the signal $B_s^0$ meson, is not used.

The tagging algorithms are calibrated by comparing the estimated mistag probability to the measured mistag probability in so-called self-tagging decays (e.g. $B^+ \rightarrow J/\psi K^+$ decays). A linear dependence is assumed between the estimated mistag probability $\eta$ and the actual mistag probability $w$:

$$w = p_0 + p_1 (\eta - \langle \eta \rangle),$$

(6.10)

Figure 6.4: Illustration of the different taggers used in LHCb. For this analysis, only the opposite-side (OS) taggers are used.
with \( p_0 \) and \( p_1 \) calibration parameters and \( \langle \eta \rangle \) the average estimated mistag probability in the calibration sample. If \( p_1 \) is found to be close to 1 and \( p_0 \) close to \( \langle \eta \rangle \), then this indicates that the estimated mistag probability based on information from simulated data is already close to the actual value (\( w \simeq \eta \)). A correct estimation of \( w \) is important, since it directly affects the determination of \( \phi_s \).

The calibration parameters \( p_0 \) and \( p_1 \) are obtained from a background-subtracted \( B^+ \to J/\psi K^+ \) sample as shown in Fig. 6.5 and the results are summarized in Table 6.3. The found values of \( p_0 \simeq \langle \eta \rangle \) and \( p_1 \simeq 1 \) indicate indeed that the estimated mistag probability \( \eta \) is close to the real mistag probability \( w \).

The uncertainties in the tagging calibration parameters \( p_0 \) and \( p_1 \) are implemented as Gaussian constraints in the final analysis, allowing the tagging calibration parameters \( p_0 \) and \( p_1 \) to vary within their uncertainties in the maximum likelihood fit for \( \phi_s \).

### Table 6.3: Fitted calibration parameters for the combined OS taggers, obtained from a \( B^+ \to J/\psi K^+ \) sample.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_0 )</td>
<td>0.392 ± 0.009</td>
</tr>
<tr>
<td>( p_1 )</td>
<td>1.035 ± 0.024</td>
</tr>
<tr>
<td>( &lt; \eta &gt; )</td>
<td>0.391</td>
</tr>
</tbody>
</table>

Figure 6.5: Actual mistag probability \( w \) versus estimated mistag probability \( \eta \) for a background-subtracted \( B^+ \to J/\psi K^+ \) sample. The fit parameters are given in Table 6.3.

The fact that the mistag probability \( w \) is non-zero, i.e. that the tag decision \( q_T \) is not perfect, leads to so-called dilution factors in the signal PDF. When considering terms with a tag decision \( q_T = 1 \), there is a contribution \( P \) from events that are tagged as \( B_s^0 \) with
probability $1 - w$, but also a contribution $P$ from events that are true $B_s^0$ but mis-tagged as $B_s^0$ with probability $w$. The resulting expression for terms in the signal PDF with $q_T = 1$ becomes

$$ P(q_T = 1) = (1 + w)P + wP = (1 - w)(U + T) + w(U - T) = U + (1 - 2w)T, \quad (6.11) $$

where $U$ indicates the untagged terms that do not involve $q_T$ and $T$ denotes the tagged terms that do involve $q_T$. For example, looking back at Table 5.2, the columns labelled by $a_i$ and $c_i$ represent the untagged terms, while the columns labelled by $b_i$ and $d_i$ represent the tagged terms that indeed flip sign in the $B_s^0$ PDF due to the appearance of $q_T$. From Eq. 6.11 it follows that all the terms in the PDF involving $q_T$ are replaced by $Dq_T$, where $D \equiv (1 - 2w)$ is the so-called dilution factor. Untagged events are accounted for by setting their mistag probability to $w = 0$.

The raw CP asymmetry $A_{CP}$ is multiplied by the effective dilution $D_{eff}$ to find the observed CP asymmetry:

$$ A_{observed} = D_{eff}A_{CP}. \quad (6.12) $$

The sensitivity to the observed CP asymmetry (and thus to $\phi_s$) is written as

$$ \sigma(A_{CP}) = \frac{1}{D_{eff}}\sigma(A_{observed}) \propto \frac{1}{D_{eff}}\frac{1}{\epsilon_T N^2} = \frac{1}{\epsilon_T D_{eff}^2 N^2} \equiv \frac{1}{\sqrt{QN}}, \quad (6.13) $$

where $\epsilon_T$ is the tagging efficiency. In other words, the effective statistical power of the data sample scales with $QN$. The effective dilution $D_{eff}$ of the full data sample is defined as

$$ D_{eff}^2 = \frac{1}{N} \sum_i D_i^2 = \frac{1}{N} \sum_{i \in \text{tagged}} (1 - 2w_i)^2 \equiv (1 - 2w_{\text{eff}})^2. \quad (6.14) $$

From the calibration parameters in Table 6.3 and the distribution of $\eta$ in the $B_s^0 \to J/\psi \phi$ dataset, the effective mistag probability is calculated as $w_{\text{eff}} = (36.8 \pm 0.2 \text{(stat.)} \pm 0.7 \text{(syst.)})\%$. The fraction of tagged events in the sample is $\epsilon_T = (32.99 \pm 0.33)\%$. This leads to an effective tagging power $Q = \epsilon_T D_{eff}^2 = (2.29 \pm 0.07 \text{(stat.)} \pm 0.26 \text{(syst.)})\%$.

### 6.5 Unbinned Likelihood Fits

The fit model that will be described in the following sections will finally result in a PDF that depends on several parameters (including the physics parameters that were introduced in the signal PDF, see Sec.[D]) and aims at describing the data on a per-event level. Using the PDF and the observed data, the so-called likelihood function is constructed:

$$ L(\vec{\Pi}; \vec{O}) \equiv \prod_{i=1}^n \text{PDF}(\vec{O}_i, \vec{\Pi}) \quad , \quad (6.15) $$

where $\vec{O}$ is the set of observables, $\vec{\Pi}$ is the set of parameters and $\text{PDF}(\vec{O}_i, \vec{\Pi})$ is the value of the PDF for event $i$, given the set of parameters $\vec{\Pi}$. The parameters $\vec{\Pi}$ are found by
maximizing the likelihood function (or equivalently, by minimizing the negative log-likelihood function), choosing the parameter set such that the observed data is as likely as possible for this PDF. All fit results in this thesis are obtained from unbinned minimum log-likelihood fits, using RooFit [67].

6.6 Fit Model

In Chap. 5, the differential \( \frac{(-)}{B^0_s \to J/\psi \phi} \) decay rate as a function of the decay time and decay angles was derived. In Sec. 6.4 two more observables were introduced, i.e. the tagging observables \( q_T \) and \( \eta \). In the PDF, both \( q_T \) and \( \eta \) are treated as conditional parameters. Choosing \( q_T \) as a conditional parameter effectively implies that the PDF for \( q_T = 1 \) and the one for \( q_T = -1 \) are normalized separately. The advantage of this choice is that it reduces the impact of a possible \( B^0_s - \overline{B^0_s} \) production asymmetry when measuring the time-dependent CP asymmetry \( \phi_s \).

In this section, the full differential decay distribution (in terms of normalized probability density functions (PDF)) are discussed, as a function of \( t, \cos \psi, \cos \theta, \phi \), the tagging observables \( q_T \) and \( \eta \), the reconstructed \( B^0_s \) mass, \( m \), and finally, the per-event decay-time error, \( \sigma_t \). The full PDF consists of a signal component and a background component and is written accordingly as

\[
P(\vec{O}) = \left( \frac{N_{\text{signal}}}{N_{\text{signal}} + N_{\text{bkg}}} P_{\text{signal}}(\vec{O}) + \frac{N_{\text{bkg}}}{N_{\text{signal}} + N_{\text{bkg}}} P_{\text{bkg}}(\vec{O}) \right) P_{\text{Poisson}}(N_{\text{signal}} + N_{\text{bkg}}),
\]

where \( P_{\text{Poisson}}(N_{\text{signal}} + N_{\text{bkg}}) \) is a Poisson term that is included because of the choice to fit for both \( N_{\text{bkg}} \) and \( N_{\text{signal}} \). Here, \( \vec{O} \) is a shorthand notation for all observables, \( P_{\text{signal}} \) and \( P_{\text{bkg}} \) are the signal and background PDF, respectively and \( N_{\text{signal}} \) and \( N_{\text{bkg}} \) are the number of signal and background events, respectively.

6.6.1 Signal Component of the PDF

The signal component of the PDF, \( P_{\text{signal}}(\vec{O}) \), that appears in Eq. 6.16 is written as

\[
P_{\text{signal}}(\vec{O}) = P_{\text{signal}}(m) P_{\text{signal}}(t, \cos \psi, \cos \theta, \phi \mid q_T, \eta) P_{\text{signal}}(q_T) P_{\text{signal}}(\eta),
\]

where the observables right of the vertical bar (\( \mid \)) indicate conditional parameters. In this equation, only the PDFs for the reconstructed \( B^0_s \) mass, \( m \), and the tagging observables \( q_T \) and \( \eta \) have not been addressed yet:

- \( P_{\text{signal}}(m) \) is modelled as the sum of two Gaussians \( f_1 G_1(\mu, \sigma_1) + (1 - f_1) G_2(\mu, \sigma_2) \), with the same mean \( \mu \), centered around the reconstructed \( B^0_s \) mass, but with different

---

\[9\] A priori, the difference between the number of \( B^0_s \) and \( \overline{B^0_s} \) mesons is expected to be small when \( \phi_s \) is close to zero. Any source of \( B^0_s - \overline{B^0_s} \) asymmetry (e.g. a production asymmetry) can influence the fitted value of \( \phi_s \) when normalizing the entire PDF (i.e. not choosing \( q_T \) as a conditional parameter) instead of separately normalizing the PDF for \( B^0_s \) and \( \overline{B^0_s} \).

\[10\] The addition of the Poisson term is needed to correctly estimate the errors on \( N_{\text{bkg}} \) and \( N_{\text{signal}} \) (as opposed to fitting for only the fraction of signal events \( f_{\text{signal}} \) when this term is not needed).
width and $\sigma_2/\sigma_1 = 2.258$ fixed from MC simulated events. In addition, also the fraction $f_1 = 0.803$ is fixed from MC. The only free parameters are $\mu$ and $\sigma_1$. Studies reveal that possible peaking backgrounds from similar final states such as $B^0_d \rightarrow J/\psi K*0$ are negligible [60].

- $P_{\text{signal}}(q_T)$: The PDF for the tag decision $q_T$ (where $q_T$ is equal to either $\pm 1$ or 0) is written as

$$P_{\text{signal}}(q_T) = \begin{cases} 
1 - \epsilon_{T,\text{signal}}, & \text{if } q_T = 0 \\
\frac{1}{2} \epsilon_{T,\text{signal}} (1 + q_T \delta_{T,\text{signal}}), & \text{if } q_T = \pm 1 
\end{cases} \quad (6.18)$$

Here, $\epsilon_{T,\text{signal}}$ is the tagging efficiency for signal events and $\delta_{T,\text{signal}}$ is the so-called signal tagging-efficiency asymmetry between the tags $q_T = 1$ and $q_T = -1$.

- $P_{\text{signal}}(\eta)$ is modelled in a non-parametric way. The PDF is sampled from the distribution of $\eta$ of a weighted dataset that represents the signal component of the data (see Sec. 6.6.3) and is shown in Fig. 6.6.

---

**Figure 6.6:** Distribution of $\eta$ for weighted data representing the signal component of the data (black points). The PDF $P_{\text{signal}}(\eta)$ is indicated by the solid line. The events in the last bin at $\eta = 0.5$ represent the untagged events with $q_T = 0$, and correspond to about 67% of the events. Inset: identical, but with the untagged events at $\eta = 0.5$ removed.
6.6.2 Background Component of the PDF

The background component of the PDF, $P_{\text{bkg}}(\vec{O})$, that appears in Eq. 6.16 is written as

$$P_{\text{bkg}}(\vec{O}) = P_{\text{bkg}}(t) P_{\text{bkg}}(\cos \psi, \cos \theta, \phi) P_{\text{bkg}}(m) P_{\text{bkg}}(q_T) P_{\text{bkg}}(\eta).$$  \hspace{1cm} (6.19)

The various distributions of the observables are modelled as follows:

- $P_{\text{bkg}}(t)$ is modelled as the sum of two exponentials:

$$P_{\text{bkg}}(t) = f_{\text{LL}}P_{\text{LL}}(t) + (1 - f_{\text{LL}})P_{\text{SL}}(t),$$  \hspace{1cm} (6.20)

where LL stands for long-living and SL stands for short-living and $f_{\text{LL}}$ is the fraction between the two components. In this model, the three free parameters are the fraction $f_{\text{LL}}$ and the lifetimes $\tau_{\text{SL}}$ and $\tau_{\text{LL}}$. Since the final analysis is performed in the range $0.3 < t < 14 \text{ ps}$, the largest background component, the so-called prompt peak around $t = 0$, can be safely ignored. The projection on $t$ for the data in the reconstructed $B^0_s$ mass sidebands (defined as $5200 \text{ MeV}/c^2 < m < 5330 \text{ MeV}/c^2$ and $5410 \text{ MeV}/c^2 < m < 5550 \text{ MeV}/c^2$) of the full PDF, is shown in Fig. 6.7.

![Figure 6.7](image)

**Figure 6.7:** Data in the $B^0_s$ mass sidebands as defined in the text (black points) and the projection of the full PDF (blue solid line) on the decay time $t$. The signal component outside the signal mass window is small and indicated by the dashed green line, whereas the double-exponential decay-time PDF $P_{\text{bkg}}(t)$ is shown as the red dashed line (almost fully covered by the full PDF, due to the small signal component).
Chapter 6. \( B_s^0 \rightarrow J/\psi \phi \) Analysis

- \( P_{\text{bkg}}(\cos \psi, \cos \theta, \phi) \) is modelled in a non-parametric way. To account for the difference in shape, the number of histogram bins is optimized for each variable: five, seven and nine for \( \cos \psi, \cos \theta \) and \( \phi \), respectively. The angular distributions are shown in Fig. 6.8 for the data in the \( B_s^0 \) mass sidebands. Shown in blue is a rebinned histogram of the same data, representing the PDF \( P_{\text{bkg}}(\cos \psi, \cos \theta, \phi) \).

![Image](image_url)

**Figure 6.8:** Data in the \( B_s^0 \) mass sidebands as defined in the text (black points) and the projection of the full PDF (blue solid line) on the decay angles \( \cos \psi, \cos \theta \) and \( \phi \). The signal component is indicated by the dashed green line, whereas the background component \( P_{\text{bkg}}(\cos \psi, \cos \theta, \phi) \) is shown as the red dashed line (almost fully covered by the full PDF, since the signal component is negligible in the mass sideband). Notice that the events in the peaking structure in \( \cos \theta \) and \( \phi \) are correlated. These are events where the two final state muons are along the decay axis of the \( B_s^0 \) meson candidate.

- \( P_{\text{bkg}}(m) \) is modelled as a simple exponential, with the exponent as free parameter.
- \( P_{\text{bkg}}(q_T) \): The background PDF for the tag decision \( q_T \) is written in the same way as the signal component, but with different parameters \( \epsilon_{T,\text{bkg}} \) and \( \delta_{T,\text{bkg}} \):

\[
P_{\text{bkg}}(q_T) = \begin{cases} 
1 - \epsilon_{T,\text{bkg}} & \text{if } q_T = 0 \\
\frac{1}{2} \epsilon_{T,\text{bkg}} (1 + q_T \delta_{T,\text{bkg}}) & \text{if } q_T = \pm 1
\end{cases}
\]  

(6.21)

- \( P_{\text{bkg}}(\eta) \) is modelled in a non-parametric way. The PDF is sampled from a distribution of \( \eta \) of the weighted dataset that represents the background component of the data (see Sec. 6.6.3) and is shown in Fig. 6.9.
Figure 6.9: Distribution of $\eta$ for weighted data representing the background component of the data (black points). The PDF $P_{\text{bkg}}(\eta)$ is indicated by the solid line. The events in the last bin at $\eta = 0.5$ represent the so-called untagged events with $q_T = 0$. Inset: identical, but with the untagged events at $\eta = 0.5$ removed.
Table 6.4: Fit results for the five free parameters in the PDF for the reconstructed $B_s^0$ mass, $m$, in the range $5200 < m < 5550$ MeV/c$^2$.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{signal}}$</td>
<td>21217 $\pm$ 153</td>
</tr>
<tr>
<td>$N_{\text{bkg}}$</td>
<td>10439 $\pm$ 112</td>
</tr>
<tr>
<td>$\mu_{\text{signal}}$</td>
<td>5368.23 $\pm$ 0.05 MeV/c$^2$</td>
</tr>
<tr>
<td>$\sigma_{\text{signal}}$</td>
<td>6.28 $\pm$ 0.05 MeV/c$^2$</td>
</tr>
<tr>
<td>$\alpha_{\text{bkg}}$</td>
<td>$(-1.63 \pm 0.10) \cdot 10^{-3}$</td>
</tr>
</tbody>
</table>

6.6.3 Signal and Background Yields and Weighted Datasets

Since $q_T$ and $\eta$ are treated as conditional parameters in the analysis, the distributions of $q_T$ and $\eta$ must be included for both the signal and the background component in order to obtain the full PDF. In the case of $\eta$, the signal and background distributions are obtained using the sPlot technique [68]. This is a method that uses information based on a discriminating observable ($m$ in this case) to infer the behavior of the signal and background component with respect to other observables (such as $\eta$, in this case). The two datasets that represent the signal component and the background component of the data are weighted according to the yields $N_{\text{signal}}$ and $N_{\text{bkg}}$ that are obtained from a fit of the mass-only PDF $P(m)$ to the data. This PDF is given by

$$ P(m) = \left( \frac{N_{\text{signal}}}{N_{\text{signal}} + N_{\text{bkg}}} P_{\text{signal}}(m) + \frac{N_{\text{bkg}}}{N_{\text{signal}} + N_{\text{bkg}}} P_{\text{bkg}}(m) \right) \text{Poisson}(N_{\text{signal}} + N_{\text{bkg}}), $$

(6.22)

where the PDFs $P_{\text{signal}}(m)$ and $P_{\text{bkg}}(m)$ were explained in Sec.6.6.1 and Sec.6.6.2 respectively. The mass PDF $P(m)$ in Eq.6.22 has five free parameters: the number of signal and background events, $N_{\text{signal}}$ and $N_{\text{bkg}}$, the mean $\mu_{\text{signal}}$ and width $\sigma_{\text{signal}}$ of the Gaussian $P_{\text{signal}}(m)$ and the exponential $\alpha_{\text{bkg}}$ from $P_{\text{bkg}}(m)$. The PDF is fitted to the data in the range $5200$ MeV/c$^2 < m < 5550$ MeV/c$^2$ and is shown in Fig.6.10. The fit results are given in Table 6.4.

6.6.4 Decay-Time Resolution Model

The decay-time resolution needs to be taken into account by convolving all time-dependent functions with the decay-time resolution model, as it dilutes the amplitude of the time-dependent asymmetry. This model is assumed to be the sum of a number of Gaussian functions:

$$ R(t, \sigma_t) = \sum_{i=1}^{k} f_i \frac{1}{\sqrt{2\pi}s_i\sigma_t} \exp \left( -\frac{(t - d)^2}{2(s_i\sigma_t)^2} \right), $$

(6.23)

where $f_i$ is the fraction and $s_i$ is the scale factor of Gaussian $i$ (where $\sum_i f_i = 1$) and $d$ is a common mean value. In addition, $\sigma_t$ is the per-event decay-time error, determined from

$^{11}$ Note that a general PDF $P(x|y)$ that is a function of $x$ and that is conditional on $y$, should be multiplied with the PDF for $y$, $P(y)$, to obtain a PDF that depends on both $x$ and $y$: $P(x,y) = P(x|y)P(y)$. 

---

Note that a general PDF $P(x|y)$ that is a function of $x$ and that is conditional on $y$, should be multiplied with the PDF for $y$, $P(y)$, to obtain a PDF that depends on both $x$ and $y$: $P(x,y) = P(x|y)P(y)$.
Figure 6.10: Reconstructed $B_s^0$ mass distribution (black points) and projection of fitted PDF $P(m)$ in blue. The signal component is shown as the green dashed line, whereas the background component is the red dashed line. The number of signal events is 21217 ± 153.

The parameters of this model are obtained from a fit to the decay-time distribution of events in the range $[-1.5, 8.0]$ ps, from a dedicated sample obtained using trigger, stripping and selection requirements that are not biased with respect to the decay time. Given the a-priori large amount of background, these events are prescaled, both in the trigger and in the stripping. Only (fake) prompt $B_s^0$ candidates are considered, formed from a prompt $J/\psi$ and two random tracks. The non-$J/\psi$ background (i.e. a background component outside the $J/\psi$ mass window in $J/\psi$ mass) is subtracted using the sPlot technique. Long-lived (non-prompt) events are modelled by the sum of two exponentials, as was indicated in Sec. 6.6.2. Studies [69] reveal that the data is reasonably well described by two Gaussians ($k = 2$), as shown in Fig. 6.11. When increasing the number of Gaussians to three, the fit parameters change, but the effective dilution remains the same.

Since in the final analysis only the effective dilution due to decay-time resolution matters, the double Gaussian decay-time resolution is transformed into a single Gaussian model with the same effective dilution. This single Gaussian model uses per-event decay-time errors and has parameters $d = 0$ and scale factor $S_{\sigma_t} = 1.45 \pm 0.06$. All time-dependent functions in

\[ D = \exp(-\Delta m^2 \sigma^2 / 2) \]

\[ P = D^2 \]

For a single Gaussian proper time resolution with width $\sigma_t$, the dilution is $P = \exp(-\Delta m^2 \sigma^2 / 2)$ and the effective power $\langle P \rangle = \sum_j f_j \exp(-\Delta m^2 \sigma^2_j / 2))^2$. When using a per-event decay-time error $\sigma_{j,e}$, the average power of the model is $\langle P \rangle = \sum_e P_e / N$, where $P_e$ is the per-event power.
the final analysis are convolved with this resolution model.

Given the fact that the average of the decay-time error distribution is 32 fs, the effective decay-time resolution is $32 \text{ fs} \times 1.45 = 46 \text{ fs}$. The scale factor of 1.45 implies that, on average, the decay-time resolution is about 45% worse than estimated from the uncertainties on the track parameters. The assigned uncertainty of 0.06 on $S_{\sigma_1}$ accounts for possible differences in the decay-time resolution between prompt $J/\psi \rightarrow \mu^+\mu^-$ background events and long-lived $B^0_s \rightarrow J/\psi \phi$ events and corresponds to a variation of the effective dilution by 2.5%. $S_{\sigma_1}$ is constrained to vary within its uncertainty in the final fit.

### 6.6.5 Angular Resolution

Angular resolution has been studied on simulated events [64]. No biases on the physics parameters were found when neglecting angular resolution, and therefore it is ignored in the analysis.

### 6.7 Results

The shape of the distribution for all variables in the full PDF has been discussed. Using this, a fit to the data is performed, from which physics parameters are extracted. Among these parameters are the amplitudes, strong phases, lifetimes and the weak phase $\phi_s$. The following set of physics parameters $\vec{\Pi}_{\text{phys}}$ is distinguished:

$$\vec{\Pi}_{\text{phys}} = (|A_0|^2, |A_\perp|^2, |A_\parallel|^2, |A_S|^2, \delta_0, \delta_\perp, \delta_\parallel, \delta_S, \Gamma_s, \Delta \Gamma_s, \Delta m_s, \phi_s) \quad .$$

(6.24)
The overall scaling of the $|A_i|^2$ drops out of the PDF due to the normalization, hence one degree of freedom must be eliminated. The amplitude $|A_S|^2$ is expected to be dependent on the $K^+K^-$ invariant mass, while the other transversity amplitudes are not. Therefore, the sum of the P-wave contributions is normalized to one:

$$|A_0|^2 + |A_\perp|^2 + |A_\parallel|^2 = 1 \quad .$$

(6.25)

Consequently, the free parameters in the fit are chosen to be $|A_0|^2$ and $|A_\perp|^2$ and it follows that $|A_\parallel|^2 = 1 - |A_0|^2 - |A_\perp|^2$. Subsequently, the amplitude $|A_S|^2$ is replaced by $f_S$, the fraction of the S-wave component:

$$f_S = \frac{|A_S|^2}{|A_0|^2 + |A_\perp|^2 + |A_\parallel|^2 + |A_S|^2} = \frac{|A_S|^2}{1 + |A_S|^2} \quad .$$

(6.26)

Because only strong phase differences between individual amplitudes enter the PDF, without loss of generality, the strong phase related to the transversity amplitude $A_0$ is taken to be constant and zero: $\delta_0 = 0$. The final set of physics parameters that is determined in the fit is:

$$\vec{\Pi}_{\text{phys}} = (|A_0|^2, |A_\perp|^2, f_S, \delta_\perp, \delta_\parallel, \Gamma_s, \Delta\Gamma_s, \Delta m_s, \phi_s) \quad .$$

(6.27)

The parameter $\Delta m_s$ is constrained in the fit to $\Delta m_s = 17.63 \pm 0.11 \text{ ps}^{-1}$, as measured by LHCb in $B^0_s \to D_s^- (3)\pi$ decays [70].

A final remark should be made about the parameter $\lambda$ as introduced in Eq.5.55. By writing $\lambda = e^{-i\phi_s}$, it was assumed that $|q/p| = 1$ (no CP violation in mixing) and $\frac{|A_{J/\psi \phi}|}{A_{J/\psi \phi}} = 1$ (no CP violation in decay), and it follows that $|\lambda| = 1$. This assumption is used in the fit as well.

### 6.7.1 Fit Results

The fit results of the physics parameters are given in Table6.5. In particular, it is found that

$$\phi_s = 0.00 \pm 0.10$$
$$\Gamma_s = 0.658 \pm 0.005 \text{ ps}^{-1}$$
$$\Delta\Gamma_s = 0.115 \pm 0.018 \text{ ps}^{-1}$$

The quoted uncertainties are statistical only. Systematic uncertainties are determined in Sec.6.7.5. The fitted values of the other free parameters in the fit are given in Table6.6. Note in particular that the values of the parameters that occur in the mass-only PDF as presented in Table6.4 change slightly in the full fit.

The projections of the fitted PDF, as well as its two components are shown in Fig.6.12 for the decay time $t$ and in Fig.6.13 for the decay angles. The observed time-dependent CP asymmetry defined as

$$A_{\text{observed}}(t) = \frac{N_{B \to f}(t) - N_{\overline{B} \to f}(t)}{N_{B \to f}(t) + N_{\overline{B} \to f}(t)} = D_{\text{eff}} A_{\text{CP}}(t) \quad ,$$

(6.28)

\[13\] Rounding rules are applied as indicated in [23].
Table 6.5: Fit results for physics parameters. The parameter $\Delta m_s$, indicated by (*) is constrained to external measurements in the fit.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_s$</td>
<td>0.00 ± 0.10</td>
</tr>
<tr>
<td>$\Gamma_s$</td>
<td>0.658 ± 0.005 ps$^{-1}$</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$</td>
<td>0.115 ± 0.018 ps$^{-1}$</td>
</tr>
<tr>
<td>$\Delta m_s$ (*)</td>
<td>17.59 ± 0.09 ps$^{-1}$</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
</tr>
<tr>
<td>$f_S$</td>
<td>0.022 ± 0.011</td>
</tr>
<tr>
<td>$\delta_\perp$</td>
<td>2.89 ± 0.34</td>
</tr>
<tr>
<td>$\delta_\parallel$</td>
<td>3.33 ± 0.21</td>
</tr>
<tr>
<td>$\delta_S$</td>
<td>2.89 ± 0.34</td>
</tr>
</tbody>
</table>

Table 6.6: Fit results for other parameters, grouped by component (signal/background) or by function in the PDF (tagging/decay-time resolution). The parameters indicated by (*) are constrained to external measurements in the fit.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{signal}}$</td>
<td>21203 ± 151</td>
</tr>
<tr>
<td>$\mu_{\text{signal}}$</td>
<td>5368.20 ± 0.05 MeV/c$^2$</td>
</tr>
<tr>
<td>$\sigma_{\text{signal}}$</td>
<td>6.27 ± 0.04 MeV/c$^2$</td>
</tr>
<tr>
<td>$N_{\text{bkg}}$</td>
<td>10453 ± 110</td>
</tr>
<tr>
<td>$\alpha_{\text{bkg}}$</td>
<td>(-1.63 ± 0.10) × 10$^{-3}$ c$^2$/MeV</td>
</tr>
<tr>
<td>$f_{\text{LL}}$</td>
<td>0.213 ± 0.010</td>
</tr>
<tr>
<td>$\tau_{\text{LL}}$</td>
<td>1.06 ± 0.04 ps</td>
</tr>
<tr>
<td>$\tau_{\text{SL}}$</td>
<td>0.1507 ± 0.0029 ps</td>
</tr>
<tr>
<td>$\delta_{T,\text{bkg}}$</td>
<td>-0.033 ± 0.016</td>
</tr>
<tr>
<td>$\epsilon_{T,\text{bkg}}$</td>
<td>0.397 ± 0.005</td>
</tr>
<tr>
<td>$\delta_{T,\text{signal}}$</td>
<td>0.000 ± 0.012</td>
</tr>
<tr>
<td>$\epsilon_{T,\text{signal}}$</td>
<td>0.3296 ± 0.0033</td>
</tr>
<tr>
<td>$p_0$ (*)</td>
<td>0.392 ± 0.009</td>
</tr>
<tr>
<td>$p_1$ (*)</td>
<td>1.036 ± 0.024</td>
</tr>
<tr>
<td>$S_{\sigma_t}$ (*)</td>
<td>1.45 ± 0.06</td>
</tr>
</tbody>
</table>

following Eq.[6.12] is shown in Fig.[6.14] No oscillation is observed due to the fact that $\phi_s$ is found to be zero in the fit and $\Delta m_s$ is large. In addition, there is extra dilution due to the fact that $|A_\perp|^2 \neq 0$ and $D_{\text{eff}}$ is small.
Figure 6.12: Projection of the fitted PDF on the decay time. The data is indicated by the black points. The PDF signal component is the green dotted line and the PDF background component is the red dotted line. The sum of the two components is the full PDF, indicated by the solid blue line.
Figure 6.13: Projection of the fitted PDF on the decay angles $\cos \psi$, $\cos \theta$ and $\phi$, from left to right. The data is indicated by the black points. The PDF signal component is the green dotted line and the PDF background component is the red dotted line. The sum of the two components is the full PDF, indicated by the solid blue line.
Figure 6.14: Observed time-dependent CP asymmetry and the corresponding projection of the fitted signal PDF. No oscillation is observed due to the fact that $\phi_s$ is found to be zero and $\Delta m_s$ is large. In addition, there is extra dilution due to the fact that $|A_\perp|^2 \neq 0$ and $D_{\text{eff}}$ is small. The data shown in this figure is the weighted data that represents the signal component.
6.7.2 Scans of Profile-Likelihood Ratios

The fit results for the physics parameters were presented in Table 6.5 and Table 6.6. The profile-likelihood ratio is defined as the ratio of the negative log-likelihood function, when one parameter is fixed and the negative log-likelihood function is minimized as a function of the remaining parameters, to the minimized negative log-likelihood function when all parameters are free (i.e. the best fit). A scan of the profile-likelihood ratio is a plot of the profile-likelihood ratio as a function of the parameter of interest. In the limit of a large number of events and no correlations between fit parameters, all the profile-likelihood ratio scans are expected to be parabolic and the statistical uncertainty on the parameter estimate is obtained from the $\Delta \log L = 0.5$ point.

1-Dimensional Scans

The scans of the profile-likelihood ratio for the physics parameters $\phi_s$, $\Gamma_s$ and $\Delta\Gamma_s$ are shown in Fig. 6.15. In addition, the scans of the profile-likelihood ratio for the transversity amplitudes $|A_0|^2$, $|A_\perp|^2$, $f_S$ and the strong phases $\delta_\perp$, $\delta_\parallel$ and $\delta_S$ are shown in Fig. 6.16.

The profile-likelihood ratio scan for $f_S$ is non-parabolic due to the choice of parameterization of $f_S$, an effect that is discussed in Sec. 6.7.5. In addition, the profile-likelihood for $\delta_\parallel$ is non-parabolic due to the fact that the fitted value is close to its degenerate value, as will be explained in Sec. 6.7.3.

![Profile-likelihood scans for $\phi_s$, $\Gamma_s$, and $\Delta\Gamma_s$.](image_url)

Figure 6.15: Profile-likelihood scans for $\phi_s$, $\Gamma_s$, and $\Delta\Gamma_s$. 
Figure 6.16: Profile-likelihood scans for $|A_0|^2$, $|A_\perp|^2$, $f_S$, $\delta_{\perp}$, $\delta_{\parallel}$ and $\delta_S$. 
2-Dimensional Scans

Historically, the $\phi_s$ analysis is presented using a two-dimensional contour plot of the profile-likelihood ratio for $\phi_s$ and $\Delta \Gamma_s$. This contour plot is shown in Fig. 6.17. In addition, the profile-likelihood ratio contours determined here are overlaid with the corresponding data from the CDF and D0 experiments and is presented in Fig. 6.18.

**Figure 6.17**: Two-dimensional contour plot of the profile-likelihood ratio of $\Delta \Gamma_s$ versus $\phi_s$. The black square indicates the SM point ($\phi_s = -0.036 \pm 0.002$, $\Delta \Gamma_s = 0.087 \pm 0.021$ ps$^{-1}$, see [51]) and the black dot indicates the fitted value as presented in Table 6.5. The confidence levels (CL) drawn correspond to 68% (solid red line), 90% (green dashed line) and 95% (blue dotted line). Note that systematic uncertainties are not included. Systematic uncertainties are treated in Sec. 6.7.5 and are not expected to have a big influence on this figure, since they are relatively small for $\phi_s$ and $\Delta \Gamma_s$.

6.7.3 Solving the Ambiguity

Fig. 6.17 and Fig. 6.18 show that two disjoint solutions for $\Delta \Gamma_s$ and $\phi_s$ are found. This can be understood by noting that the signal PDF as presented in Appendix D is invariant under the following transformations:

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

(6.29)
Figure 6.18: Two-dimensional contour plots of the profile-likelihood ratios in the \((\Delta \Gamma_s, \phi_s)\)-plane for the D0 collaboration [59] (red), the CDF collaboration [72] (green) and the values found here (blue). The black square indicates the SM point \((\phi_s = -0.036 \pm 0.002, \Delta \Gamma_s = 0.087 \pm 0.021 \text{ ps}^{-1})\). Only the 68\% confidence levels (solid lines) and 90\% confidence levels (dotted lines) are indicated. Note that systematic uncertainties are not included. Systematic uncertainties are treated in Sec. 6.7.5 and are not expected to have a big influence on this figure, since they are small compared to the statistical uncertainties for \(\phi_s\) and \(\Delta \Gamma_s\).

This explains the appearance of a second minimum in the profile-likelihood ratio scans and it also explains the non-parabolic likelihood scan of the parameter \(\delta_{\parallel}\) in Fig. 6.16, since the fitted value of \(\delta_{\parallel} = 3.33 \pm 0.21\) is close to its degenerate solution \(\delta_{\parallel} \rightarrow -\delta_{\parallel}\), modulo \(2\pi\).

However, this ambiguity can be resolved by measuring the phase difference between the S-wave and P-wave amplitudes as a function of the \(K^+K^-\) invariant mass, \(m_{KK}\). The phase of the P-wave amplitude, which can be described by a spin-1 Breit-Wigner function of \(m_{KK}\), rises rapidly through the \(m_{KK}\) region. The phase of the S-wave amplitude, on the other hand, is expected to vary relatively slowly for both an \(f_0(980)\) meson S-wave contribution and a non-resonant S-wave contribution. As a result, the phase difference \(\delta_S - \delta_\perp\) between the S-wave and P-wave amplitudes falls rapidly with increasing \(m_{KK}\). By measuring this phase difference as a function of \(m_{KK}\), and taking the solution with a decreasing trend when going through the \(\phi\) meson mass pole from low to high masses, the ambiguity is resolved. In particular, the sign of \(\Delta \Gamma_s\), which is a priori unknown, as mentioned in Sec. 1.4, is determined...
by solving the ambiguity.

To solve the ambiguity, a dataset with a wider $m_{KK}$ range is used, as compared to the range that was given in Table 6.1. Subsequently, this $m_{KK}$ region is divided in four bins and the phase difference $\delta_{S\perp} \equiv \delta_S - \delta_\perp$ is fitted in every bin. This is shown in Fig. 6.19 which indicates that the solution with $\Delta \Gamma_s > 0$ is the physical solution, in agreement with the SM prediction [21]. A zoomed-in version of the profile-likelihood ratio contours in the $(\Delta \Gamma_s, \phi_s)$-plane showing the physical solution is presented in Fig. 6.20.

6.7.4 $\Delta m_s$ Measurement

Although $\Delta m_s$ is constrained in the fit, the PDF that is used in the $\phi_s$ analysis presented here, does contain terms that are sensitive to $\Delta m_s$. With the current size of the dataset, when inspecting the separate terms of the PDF, some of the terms proportional to $\sin(\Delta m_s t)$ and $\cos(\Delta m_s t)$ are not multiplied by $C = 0$ or $S = -\sin \phi_s \simeq 0$ and are thus sensitive to $\Delta m_s$ (for example, the terms proportional to $\Delta m_s$ in Eq. D.6 are not multiplied by $C$ or $S$, while the terms proportional to $\Delta m_s$ in Eq. D.1 are).
it is possible to float the parameter $\Delta m_s$. The best fit result is

$$
\Delta m_s = 17.51 \pm 0.15 \text{ps}^{-1},
$$

(6.30)

which is in good agreement with the LHCb measurement performed in [70], $\Delta m_s = 17.63 \pm 0.11 \text{ps}^{-1}$ [70]. In this fit, $\phi_s = 0.00 \pm 0.10$. This value is the same as found when $\Delta m_s$ is constrained, see Table 6.5. The profile-likelihood ratio scan for $\Delta m_s$ is shown in Fig. 6.21.

### 6.7.5 Systematic Uncertainties

Systematic uncertainties are assigned by studying possible sources of systematic effects. These studies are summarized here and finally the systematic uncertainties for all parameters are summarized.

**S-Wave Fraction**

By construction, the fraction of the S-wave contribution, $f_S$, can never be smaller than zero. If $f_S$ is close to zero, this leads to non-parabolic profile-likelihood ratio scans and biases in fit parameters. Since $f_S$ is determined to be small ($f_S = 0.022 \pm 0.011$), the scan of the profile-likelihood ratio of $f_S$ in Fig. 6.16 is not parabolic.
To study the effect of possible biases, toy datasets are generated at \( f_S = 0.0, f_S = 0.02 \) and finally \( f_S = 0.04 \). For all parameter biases, no dependence on the generation value of \( f_S \) is observed, where a parameter bias is defined as the difference between the fitted parameter and the value of the parameter that was used for generation. Therefore, no systematic uncertainty is assigned due to this effect.

**Fit Bias**

To estimate the parameter biases as a result of the limited data size, toy studies are performed in which 500 toys are generated that are representative for the actual dataset. The events in the toy datasets are generated using the values of the parameters that are extracted from the real data.

For every parameter, the bias is defined as \( p_{\text{fit}} - p_{\text{generated}} \), where \( p_{\text{generated}} \) is the value of the parameter at which the toy dataset was generated and \( p_{\text{fit}} \) is the fitted parameter value. For each parameter, the mean bias and its error is given in Table 6.7, as well as the statistical uncertainty of the original fit. Only for \( \phi_s \) no significant bias is observed (mean bias consistent with zero within 2\( \sigma \)) and therefore no systematic uncertainty is assigned to \( \phi_s \). For all other parameters, the absolute value of the mean bias is assigned as systematic uncertainty. In all cases the assigned systematic error is small compared to the statistical uncertainty of the original measurement.

**Background Description**

The fit model described in Sec. 6.6 (called the ‘baseline fit’ onwards) contains a signal component and a background component. In addition to this fit configuration, a fit of only the signal component PDF on the s-weighted data representing the signal component (as explained in Sec. 6.6.3) is performed. By construction, there is no background component in this dataset and therefore no background PDF is needed. The fit result for this fit...
Table 6.7: Mean bias and mean bias error obtained from toy studies, and the original statistic uncertainty for all parameters. Only $\phi_s$ shows a bias that is consistent with zero within $2\sigma$ and therefore does not have an assigned systematic uncertainty.

<table>
<thead>
<tr>
<th>parameter</th>
<th>mean bias in toys</th>
<th>statistical error from measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_s$</td>
<td>$(-0.2 \pm 0.5) \cdot 10^{-2}$</td>
<td>0.10</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>$(3.44 \pm 0.25) \cdot 10^{-3}$</td>
<td>0.005</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>$(-2.2 \pm 0.8) \cdot 10^{-3}$</td>
<td>0.018</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
<td>^2$</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
<td>^2$</td>
</tr>
<tr>
<td>$f_S$</td>
<td>$(-5.0 \pm 0.5) \cdot 10^{-3}$</td>
<td>0.011</td>
</tr>
<tr>
<td>$\delta_\perp$</td>
<td>$(-3.1 \pm 1.5) \cdot 10^{-2}$</td>
<td>0.34</td>
</tr>
<tr>
<td>$\delta_\parallel$</td>
<td>$(9.8 \pm 0.8) \cdot 10^{-2}$</td>
<td>0.21</td>
</tr>
<tr>
<td>$\delta_S$</td>
<td>$(-4.7 \pm 1.7) \cdot 10^{-2}$</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Table 6.8: Absolute difference in fit results between the baseline fit and the sFit. A hyphen '-' indicates a negligible effect.

<table>
<thead>
<tr>
<th>parameter</th>
<th>baseline fit value</th>
<th>sFit value</th>
<th>absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_s$</td>
<td>$0.00 \pm 0.10$</td>
<td>$-0.01 \pm 0.10$</td>
<td>0.01</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>$0.658 \pm 0.005$</td>
<td>$0.659 \pm 0.005$</td>
<td>0.001</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>$0.115 \pm 0.018$</td>
<td>$0.112 \pm 0.017$</td>
<td>0.003</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
<td>^2$</td>
<td>$0.522 \pm 0.007$</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
<td>^2$</td>
<td>$0.247 \pm 0.010$</td>
</tr>
<tr>
<td>$f_S$</td>
<td>$0.022 \pm 0.011$</td>
<td>$0.024 \pm 0.010$</td>
<td>0.002</td>
</tr>
<tr>
<td>$\delta_\perp$</td>
<td>$2.89 \pm 0.34$</td>
<td>$2.84 \pm 0.29$</td>
<td>0.05</td>
</tr>
<tr>
<td>$\delta_\parallel$</td>
<td>$3.33 \pm 0.21$</td>
<td>$3.36 \pm 0.16$</td>
<td>0.03</td>
</tr>
<tr>
<td>$\delta_S$</td>
<td>$2.89 \pm 0.34$</td>
<td>$2.83 \pm 0.29$</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Angular Acceptance

A source of systematic uncertainty arising from angular acceptance is the observed difference in the kaon momentum spectra between simulated and real data. It turns out that the kaon
momentum and the decay angle $\cos \psi$ are related$^{16}$ and thus that the kaon momentum spectra must be properly simulated. As a systematic check, the angular acceptance is determined from a reweighted MC dataset that by construction matches the observed kaon momentum spectra. The assigned systematic uncertainties are the absolute differences in fitted values between the baseline fit and the fit where the angular acceptance is obtained from the reweighted MC dataset and are summarized in Table 6.9. The systematic uncertainties are rounded to the same number of digits as the baseline fit result for each parameter.

<table>
<thead>
<tr>
<th>parameter</th>
<th>absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_s$</td>
<td>-</td>
</tr>
<tr>
<td>$\Gamma_s [\text{ps}^{-1}]$</td>
<td>0.001</td>
</tr>
<tr>
<td>$\Delta \Gamma_s [\text{ps}^{-1}]$</td>
<td>0.001</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
</tr>
<tr>
<td>$f_\Sigma$</td>
<td>-</td>
</tr>
<tr>
<td>$\delta_\perp$</td>
<td>0.02</td>
</tr>
<tr>
<td>$\delta_\parallel$</td>
<td>0.03</td>
</tr>
<tr>
<td>$\delta_S$</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 6.9: Assigned systematic uncertainties indicating the absolute difference in fit results between the baseline fit and a fit with the angular acceptance as determined from a reweighted MC dataset that matches the observed kaon momentum spectra in data. The numbers are rounded to the same number of digits as the baseline fit result for each parameter. A hyphen ‘-’ indicates a negligible effect.

In addition, the limited size of the MC dataset that was used to determine the angular acceptance is taken into account. This is done by performing additional fits in which the calculated coefficients of the expansion of the angular-acceptance function, $c_{00}^0$, $c_{22}^0$, $c_{20}^0$ and $c_{02}^0$ (see Eq. 6.2), are varied by $\pm 1\sigma$. For each coefficient, the absolute biases from the $\pm 1\sigma$ fits are averaged and finally the averages for each coefficient are summed quadratically. The assigned systematic uncertainties are summarized in Table 6.10 for all parameters.

Acceptance for small decay time

To estimate the systematic uncertainty on the determination of the decay-time acceptance histogram, the baseline fit is repeated without multiplying the signal PDF with this histogram. The differences in the fit results are assigned as systematic uncertainties and are summarized in Table 6.11.

---

$^{16}$This can be understood as follows: when $\cos \psi = 1$ ($\cos \psi = -1$), the $K^+$ is emitted in the direction of (opposite to) the $B$ meson, i.e. with a relatively low (high) $p_T$ (see Fig. 5.1). For the $K^-$ meson, the relationships are opposite. This indicates that (implicit) cuts on the $p_T$ of the kaons induce acceptance effects in $\cos \psi$ $^{64}$. 
### 6.7 Results

<table>
<thead>
<tr>
<th>parameter</th>
<th>absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi_s )</td>
<td></td>
</tr>
<tr>
<td>( \Gamma_s [\text{ps}^{-1}] )</td>
<td>0.001</td>
</tr>
<tr>
<td>( \Delta \Gamma_s [\text{ps}^{-1}] )</td>
<td>0.001</td>
</tr>
<tr>
<td>(</td>
<td>A_0</td>
</tr>
<tr>
<td>(</td>
<td>A_\perp</td>
</tr>
<tr>
<td>( f_S )</td>
<td>0.001</td>
</tr>
<tr>
<td>( \delta_\perp )</td>
<td>0.01</td>
</tr>
<tr>
<td>( \delta_\parallel )</td>
<td>0.02</td>
</tr>
<tr>
<td>( \delta_S )</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Table 6.10:** Assigned systematic uncertainties due to the finite size of the MC dataset. These numbers are obtained by varying the four coefficients of the expansion of the angular-acceptance function within \( \pm 1\sigma \) and summing the fit biases quadratically for the four coefficients. A hyphen '-' indicates a negligible effect.

<table>
<thead>
<tr>
<th>parameter</th>
<th>baseline fit value</th>
<th>fit without decay-time acceptance</th>
<th>absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi_s )</td>
<td>0.00 ± 0.10</td>
<td>0.00 ± 0.10</td>
<td>-</td>
</tr>
<tr>
<td>( \Gamma_s [\text{ps}^{-1}] )</td>
<td>0.658 ± 0.005</td>
<td>0.656 ± 0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>( \Delta \Gamma_s [\text{ps}^{-1}] )</td>
<td>0.115 ± 0.018</td>
<td>0.117 ± 0.018</td>
<td>0.002</td>
</tr>
<tr>
<td>(</td>
<td>A_0</td>
<td>^2 )</td>
<td>0.522 ± 0.007</td>
</tr>
<tr>
<td>(</td>
<td>A_\perp</td>
<td>^2 )</td>
<td>0.247 ± 0.010</td>
</tr>
<tr>
<td>( f_S )</td>
<td>0.022 ± 0.011</td>
<td>0.022 ± 0.011</td>
<td>-</td>
</tr>
<tr>
<td>( \delta_\perp )</td>
<td>2.89 ± 0.34</td>
<td>2.89 ± 0.34</td>
<td>-</td>
</tr>
<tr>
<td>( \delta_\parallel )</td>
<td>3.33 ± 0.21</td>
<td>3.32 ± 0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>( \delta_S )</td>
<td>2.89 ± 0.34</td>
<td>2.89 ± 0.35</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 6.11:** Absolute differences in fit results between the baseline fit including decay-time acceptance and the fit without taking the decay-time acceptance into account. A hyphen '-' indicates a negligible effect.

#### Acceptance for large decay time

In Sec. 6.3.2, the acceptance function for large decay time was found to be \( \epsilon(t) = 1 + \beta t \), with \( \beta = -0.0112 \pm 0.0013 \text{ ps} \). The assigned systematic uncertainty is half the value of \( \beta \) \([69]\), i.e. 0.006. Since this decay-time acceptance influences \( \Gamma_s \) only, this is directly translated as a systematic uncertainty of 0.006 ps\(^{-1}\) on \( \Gamma_s \).

#### Length Scale and Momentum Scale

The determination of the parameters \( \Delta \Gamma_s \) and \( \Gamma_s \) is a measurement of the lifetimes of the two mass eigenstates \( B^0_{s,H} \) and \( B^0_{s,L} \). The \( B \) meson lifetime \( t \) is determined from the distance between the primary vertex (PV) and the \( B \) meson decay vertex: \( \frac{m_p \cdot d}{p^2} \), with \( d \) the...
decay distance, \( m \) the reconstructed \( B \) mass and \( p \) the \( B \) momentum. From this relation, it follows that 
\[
\left( \frac{\delta t}{t} \right)^2 = \left( \frac{\delta p}{p} \right)^2 + \left( \frac{\delta d}{d} \right)^2.
\]

The length scale \( \left( \frac{\delta d}{d} \right) \) is the accuracy by which the 
\( z \)-coordinate position of each of the VELO modules is known and is estimated to be at
most 0.1%, whereas the momentum scale \( \left( \frac{\delta p}{p} \right) \) is determined from the shift in reconstructed masses of known resonances and is determined to be at most 0.15% [69]. From these, the relative error \( \left( \frac{\delta t}{t} \right) = 0.18\% \) and the assigned systematic uncertainty is 0.001ps\(^{-1}\) for \( \Gamma_s \), while there is a negligible effect for \( \Delta \Gamma_s \). The fit parameters other than \( \Gamma_s \) and \( \Delta \Gamma_s \) are not
affected by the length scale and momentum scale uncertainties.

**Nuisance CP Asymmetries**

A measurement of the asymmetry that results from CP violation in the interference be-
tween \( B^0_s \rightarrow J/\psi \phi \) mixing and decay (parameterized by \( \phi_s \)) is affected by several other (CP) asymmetries or nuisance asymmetries:

- CP violation in decay and/or mixing: in the \( \phi_s \) analysis presented here, it is assumed that there is neither direct CP violation, nor CP violation in mixing. This is reflected in the statement \( |\lambda| = 1 \) in Sec. 5.7. The effect of this assumption is checked by performing toy studies in which events are generated with \( |\lambda|^2 = 0.95 \) and \( |\lambda|^2 = 1.05 \). These datasets are subsequently fitted with the assumption \( |\lambda|^2 = 1 \). The differences in the fit results are assigned as systematic uncertainties to the various parameters. The effect depends on the actual value of \( \phi_s \) and is found to be negligible for both \( \Gamma_s \) and \( \Delta \Gamma_s \), and 0.02 for \( \phi_s \).

- production asymmetry between \( B^0_s \) and \( \bar{B}^0_s \) mesons: in the \( \phi_s \) analysis, it is assumed that there is no production asymmetry between \( B^0_s \) and \( \bar{B}^0_s \) mesons. Toy studies are performed by generating datasets with events that are generated with a generous \( \pm 10\% \) production asymmetry. These datasets are fitted with the assumption of no production asymmetry. Again, the differences in the fit results are assigned as systematic uncertainties to the various parameters. The effect depends on the value of \( \phi_s \) and is found to be (for the fitted value of \( \phi_s \)) negligible for both \( \Gamma_s \) and \( \Delta \Gamma_s \), and 0.01 for \( \phi_s \).

- difference in tagging efficiency for \( B^0_s \) and \( \bar{B}^0_s \) mesons: as indicated in Table 6.6, the signal tagging asymmetry \( \delta_{T,\text{signal}} = 0.000 \pm 0.012 \), therefore, this possible source of nuisance asymmetry is ignored.

- difference in mistag probability between \( B^0_s \) and \( \bar{B}^0_s \) mesons: the effect of this nuisance asymmetry is absorbed in the constraints on the parameters \( p_0 \) and \( p_1 \) as indicated in Sec. 6.4.

**Decay-Time Resolution, Tagging and \( \Delta m_s \)**

The uncertainties related to the decay-time resolution, the tagging calibration and \( \Delta m_s \) are
taken into account by constraining the related parameters in the fit:

- the scale factor of the decay-time resolution model is constrained in the fit to \( S_{\sigma_t} = 1.45 \pm 0.06 \), as mentioned in Sec. 6.6.4.
• the tagging calibration parameters $p_0$ and $p_1$ are constrained in the fit to $p_0 = 0.392 \pm 0.009$ and $p_1 = 1.035 \pm 0.024$, as mentioned in Sec. 6.4.

• the $B^0_s - \bar{B}^0_s$ oscillation frequency $\Delta m_s$ is constrained in the fit to $\Delta m_s = 17.63 \pm 0.11$ ps$^{-1}$, as measured by LHCb, see Sec. 6.7.

As a consequence, for these parameters, the uncertainties are (implicitly) included in the quoted statistical uncertainty.

**Bias from Peaking Backgrounds**

The only identified source of peaking background is from $B^0 \rightarrow J/\psi K^*$ events. The fraction of these events is estimated to be at most 2% and is estimated to be negligible compared to statistical uncertainties [69].

### 6.7.6 Summary of Systematic Uncertainties

All the mentioned systematic uncertainties are summarized for each parameter in Table 6.12. The total systematic uncertainty for each parameter is taken to be the quadratic sum of all the individual systematic uncertainties.

| Source of Uncertainty                           | $\phi_s$ | $\Gamma_s$ [ps$^{-1}$] | $\Delta \Gamma_s$ [ps$^{-1}$] | $|A_0|^2$ | $|A_\perp|^2$ | $f_S$ | $\delta_\perp$ | $\delta_\parallel$ | $\delta_S$ |
|------------------------------------------------|----------|------------------------|-------------------------------|----------|----------------|-------|----------------|-------------------|----------|
| stat. uncert.                                   | 0.10     | 0.005                  | 0.018                         | 0.007    | 0.010          | 0.011 | 0.34           | 0.21              | 0.34     |
| fit bias                                        | -        | 0.003                  | 0.002                         | 0.003    | 0.005          | 0.03  | 0.10           | 0.05              |          |
| bkg. modelling                                  | 0.01     | 0.001                  | 0.003                         | 0.001    | -              | 0.002 | 0.05           | 0.03              | 0.06     |
| ang. acc. rew.                                  | -        | 0.001                  | 0.001                         | 0.032    | 0.018          | -     | 0.02           | 0.03              | 0.02     |
| ang. acc. stat.                                 | -        | 0.001                  | 0.001                         | 0.001    | 0.002          | 0.001 | 0.01           | 0.02              | 0.01     |
| small time acc.                                 | -        | 0.002                  | 0.002                         | -        | 0.001          | -     | -              | 0.01              | -        |
| large time acc.                                 | -        | 0.006                  | -                             | -        | -              | -     | -              | -                 | -        |
| length + mom. scale                             | -        | 0.001                  | -                             | -        | -              | -     | -              | -                 | -        |
| CPV mix + dec.                                  | 0.02     | -                      | -                             | -        | -              | -     | -              | -                 | -        |
| prod. asymm.                                    | 0.01     | -                      | -                             | -        | -              | -     | -              | -                 | -        |
| total syst. uncert.                             | 0.02     | 0.007                  | 0.004                         | 0.032    | 0.018          | 0.005 | 0.06           | 0.11              | 0.08     |

**Table 6.12**: Summary of systematic uncertainties. The total systematic uncertainty for every parameter is the quadratic sum of all the sources of systematic uncertainties. A hyphen ‘-’ indicates no or negligible effect.

### 6.8 Final Results including Systematic Uncertainties

A time-dependent angular analysis is performed on approximately 21200 $B^0_s \rightarrow J/\psi \phi$ candidates, obtained from 1 fb$^{-1}$ of $pp$ collisions collected during the 2011 LHCb runs at $\sqrt{s} = 7$ TeV. With an effective decay-time resolution of 46 fs and an effective tagging efficiency of $\epsilon_T D^2_{\text{eff}} = (2.3 \pm 0.3)\%$, the following results for $\phi_s$, $\Gamma_s$ and $\Delta \Gamma_s$ are found:
\[
\phi_s = 0.00 \pm 0.10 \text{ (stat.)} \pm 0.02 \text{ (syst.)}
\]
\[
\Gamma_s = 0.658 \pm 0.005 \text{ (stat.)} \pm 0.007 \text{ (syst.) ps}^{-1}
\]
\[
\Delta \Gamma_s = 0.115 \pm 0.018 \text{ (stat.)} \pm 0.004 \text{ (syst.) ps}^{-1}
\]

In addition, for the transversity amplitudes the following values are found:
\[
|A_0|^2 = 0.522 \pm 0.007 \text{ (stat.)} \pm 0.032 \text{ (syst.)}
\]
\[
|A_\perp|^2 = 0.247 \pm 0.010 \text{ (stat.)} \pm 0.018 \text{ (syst.)}
\]
\[
f_S = 0.022 \pm 0.011 \text{ (stat.)} \pm 0.005 \text{ (syst.)}
\]

The parameter $|A_\parallel|^2$ is not a fit parameter, since the sum of the P-wave amplitudes is normalized to one, see Eq.6.25. Finally, for the strong phases it is found that
\[
\delta_\perp = 2.89 \pm 0.34 \text{ (stat.)} \pm 0.06 \text{ (syst.)}
\]
\[
\delta_\parallel = 3.33 \pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (syst.)}
\]
\[
\delta_S = 2.89 \pm 0.34 \text{ (stat.)} \pm 0.08 \text{ (syst.)}
\]

### 6.9 Discussion and Outlook

The measurement of $\phi_s = 0.00 \pm 0.10 \text{ (stat.)} \pm 0.02 \text{ (syst.)}$ is the world’s most precise measurement of $\phi_s$. In addition, the measurement of $\Delta \Gamma_s = 0.115 \pm 0.018 \text{ (stat.)} \pm 0.004 \text{ (syst.) ps}^{-1}$ is the first direct observation of a non-zero value of $\Delta \Gamma_s$. These results are all in good agreement with the Standard Model.

For most of the parameters, the uncertainties are still statistics-dominated. This is not the case for $|A_0|^2$ and $|A_\perp|^2$, where the systematic uncertainty is dominated by the angular acceptance correction. In addition, for $\Gamma_s$, the dominating uncertainty is the systematic uncertainty on the large decay-time acceptance.

Several improvements on the measurement of $\phi_s$ are foreseen. First of all, inclusion of the same-side tagger will improve the effective tagging power and thus the sensitivity to $\phi_s$. In addition, a different trigger strategy can increase the event yields.

The weak phase $\phi_s$ is measured in other decays as well, for instance in $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ decays [73], which also allows for the combination of the results from separate independent analyses. Finally, to ensure that the zero value of $\phi_s$ is a genuine SM effect, instead of possible New Physics effects that are cancelled by penguin contributions, the latter should be taken into account as described in Sec.5.9.
6.9 Discussion and Outlook