



Universiteit
Leiden
The Netherlands

Measurement of the CP-violating parameter ϕ_s in $B_s \rightarrow J/\psi K_s$ decays

Westerveld, Jeroen

Citation

Westerveld, J. (2024). *Measurement of the CP-violating parameter ϕ_s in $B_s \rightarrow J/\psi K_s$ decays*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/3775363>

Note: To cite this publication please use the final published version (if applicable).



Measurement of the CP-violating parameter ϕ_S in $B_S \rightarrow J/\psi K_S$ decays

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

PHYSICS

Author :	J. Westerveld
Student ID :	s3037975
Supervisor :	Dr. W. Hulsbergen
Second corrector :	Dr. D. Samtleben

Leiden, The Netherlands, February 12, 2024

Measurement of the CP-violating parameter ϕ_s in $B_s \rightarrow J/\psi K_s$ decays

J. Westerveld

Instituut-Lorentz, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

February 12, 2024

Abstract

The goal of this project was to obtain the CP-violating parameter ϕ_s from the run 2 LHCb data and analyse the size of the bias and statistical uncertainty due to the fitting method used. In order to obtain the bias toy data sets were generated with varied parameters and results of three different fitting methods were compared. When compared, the performance of the unbinned likelihood fit(ULF) method did not differ from that of the binned asymmetry fit(BAF) and the RooFit(ROF) method for most cases. From this analysis it became clear that the bias magnitudes of all methods depended on the value of $\Delta\Gamma_s$, the value of S and also on σ_t . Fitting the LHCb run 2 data set with the ULF method resulted in $S = -0.11 \pm 0.29(\text{stat}) \pm 0.03(\text{syst})$ and $C = 0.08 \pm 0.33(\text{stat}) \pm 0.01(\text{syst})$. These values were similar to that of an independent analysis of LHCb run 1 data done by R. Aaij et al. 2013 [1], taking into account that they had access to both the OS and SS flavor tagging. When compared to the theoretical prediction by K. De Bruyn and R. Fleischer 2015 [2] no contradictions were found.

Contents

1	Introduction	1
2	Theory	3
2.1	Quarks and flavor changing	3
2.2	The mixing process	5
2.3	The decay process	9
3	Data Analysis	13
3.1	Flavor tagging	13
3.2	Background subtraction	16
3.3	Time resolution	16
3.4	Validation of the fit method	18
3.4.1	CP violation variable from toy data	18
3.4.2	Bias and uncertainty of the fit	21
3.4.3	Toy data fitting	25
3.4.4	Bias due to $\Delta\Gamma_s$	25
3.4.5	Influence of σ_t	26
3.4.6	Pull width results	26
3.4.7	Fit error in S_{fit} and C_{fit}	27
3.4.8	Method bias comparison	27
4	Results	31
5	Conclusion	35

Introduction

Particle physics has been a subject that has grown in size and has become increasingly more interesting over the last decades. One of the big developments in this field of science is CERN, with their Large Hadron Collider (LHC). This LHC allows for the detection of many different particles and processes that involve them. One of these decay processes involves the Beauty meson and is of particular interest since this decay process results in what is known as CP violation, it was first detected in 1964 by James Cronin and Val Fitch in the neutral kaon system. This was an indirect way of measuring CP violation since then many different ways of directly detecting these CP violating processes were conceived which also led to the development of the LHCb (LHC beauty) experiment, that focuses primarily on B-physics. It was only in 2013 that the LHCb announced the discovery of CP violation in strange Beauty mesons [3]. Many analyses of the LHCb run 1 and run 2 data have been done to compare the observed CP violating parameters to their theoretical predictions. These analyses, of which this thesis will be another example, serve as a check of the completeness of the standard model as these CP violating parameters are sensitive probes of physics beyond the standard model. Any large deviation in these parameter gives insight into the possible existence of physics beyond the standard model.

In this paper we will analyse the LHCb run 2 data but before that we will explore three different methods of fitting data by analysing toy data. First we will build up a theoretical framework in order to fully understand the B meson decay process and understand what parameter we are trying to obtain from the data. Then we will go over some considerations that we made in order to correctly interpret and analyse the data. Followed by a

display of the results, a discussion of their validity and a conclusion about the results of the analysis of the data.

Theory

2.1 Quarks and flavor changing

In the Standard Model (SM) quarks are fermions that serve as the building block for all baryonic matter. In total there are six of these quarks that can be arranged into three families as follows:

$$\begin{pmatrix} \text{up}(u) \\ \text{down}(d) \end{pmatrix} \quad \begin{pmatrix} \text{charm}(c) \\ \text{strange}(s) \end{pmatrix} \quad \begin{pmatrix} \text{top}(t) \\ \text{bottom}(b) \end{pmatrix} \quad (2.1)$$

We have the up type quarks that have an elementary charge of $\frac{2}{3}e$ and the down type quarks that have a charge of $-\frac{1}{3}e$. All of these quarks (q) have a mirror image called the anti-quark (\bar{q}). These anti-quarks have the same quantum numbers as their quark counterpart, except for the flavour and charge numbers, as these are opposite. In quark-flavor physics there exists a charged weak current that allows quarks to change flavor. Examples of the Feynmann diagrams of this charged weak current interaction can be seen in figure 2.1.

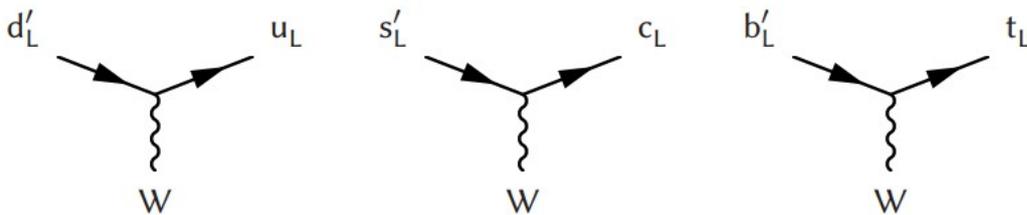


Figure 2.1: Examples of charged-current weak interactions of quarks. [4]

Here the W represents the W -boson, these particles mediate the weak interaction. In addition d'_L , u_L , s'_L , c_L , b'_L and t_L are all quarks of the kind

mentioned in equation 2.1. The subscript "L" is to indicate that only quarks with left-handed chirality partake in this interaction. Note the primes on the d'_L , s'_L and b'_L , these are an indication that these quarks in figure 2.1 that couple to the W-boson are not quark mass eigenstates but are instead given by a rotation of the mass quark mass eigenstates. This rotation can be described using the following equation[5]:

$$\begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} \equiv V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} \quad (2.2)$$

Here V_{CKM} , the Cabibbo-Kobayashi-Maskawa matrix, is composed of the following elements:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.3)$$

The elements of this matrix tell us how strong the coupling, in other words the magnitude, is between different quarks. On the diagonal there is coupling between up-type and down-type within the same generation of quarks, these are all approximately equal to one. The magnitudes for mixing interactions involving quarks of the second and third generation are significantly lower with first and second generation mixing magnitudes being four to five times lower and where mixing between the first and third and between the second and third generation are suppressed even more.

We introduced the mixing for regular quarks, but the same process also happens for anti-quarks. By replacing the left handed quarks in figure 2.1 with right handed anti-quarks one can obtain the charged-current weak interactions of anti-quarks. In the CKM matrix this would result in all elements being replaced with their complex conjugates. Note that these would result in the same elements only if the elements would be real, however this is not the case. One could introduce a complex phase to describe the elements of the matrix, and this complex phase is exactly what causes charge-parity (CP) violation. It is not possible to observe such CP invariance directly in processes that only have a single W-boson amplitude in their description, since squaring such an amplitude would get rid of the phase part. Hence we need interference between two (or more) processes that have different CKM elements in order to observe CP violation, since different CKM elements introduce different weak phases which ultimately influence the observable magnitude of the total amplitude. In section 2.3

we will see how we can use this information to derive the variable of interest ϕ_s .

2.2 The mixing process

The particle of interest in this thesis is the so called Beauty meson B^0 , in particular the version of the particle with a strange quark s in it denoted by B_s^0 . The strange Beauty meson (B_s^0) consists of a combination of an anti-bottom (\bar{b}) and a strange (s) quark and the anti-particle counterpart (\bar{B}_s^0) consists of a bottom (b) and an anti-strange (\bar{s}) quark. Since the B_s^0 particle and the anti-particle \bar{B}_s^0 are both charge-neutral particles it is possible for them to change into each other. This process is called neutral mixing and the leading order diagrams for this transition are the following:

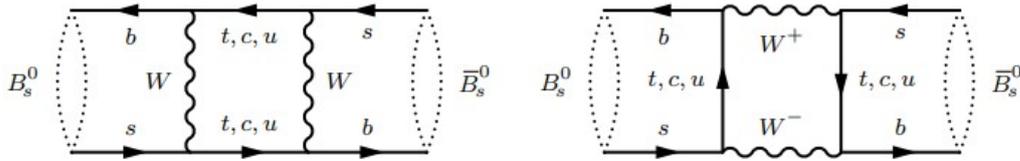


Figure 2.2: Figure showing the leading order diagrams of the B_s^0 meson mixing process. [6]

Note that this diagram can be read both left to right and right to left, respectively for the $B_s^0 \rightarrow \bar{B}_s^0$ and the $\bar{B}_s^0 \rightarrow B_s^0$ transition. The neutral mixing processes shown in figure 2.2 are heavily suppressed [7], they are very sensitive to physics beyond the standard model. Hence this mixing process can be used to check the quality of the SM and can even act as a probe for physics beyond the SM. It should also be said that these different flavour states have different masses and different lifetimes. One important consequence of this mixing is that the decay time distribution is no longer an exponential but contains additional sinusoidal terms, but we will get back to this later.

In our starting situation a meson is produced in either the $|B_s^0\rangle$ or $|\bar{B}_s^0\rangle$ flavour eigenstate. The state of the meson will evolve over time and turn into a mixture of these flavour eigenstates, the wavefunction as a function of time (t) for this process can be written as follows:

$$|\Psi(t)\rangle = a(t)|B_s^0\rangle + b(t)|\bar{B}_s^0\rangle \quad (2.4)$$

Here the coefficients $a(t)$ and $b(t)$ contain the time dependency of the

flavour eigenstates. Under the assumption that the time scale we are looking at is much larger than the time scale of strong interactions, the evolution of these coefficients can be described by the Schrödinger equation, with a constant Hamiltonian \mathbf{H} :

$$i \frac{\partial}{\partial t} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathbf{H} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \quad (2.5)$$

In this equation $\frac{\partial}{\partial t}$ denotes the (partial) time derivative and i is the imaginary number. This equation is true given that one works with natural units where: $\hbar \equiv c \equiv 1$. The Hamiltonian \mathbf{H} can be decomposed into a Hermitian and an anti-Hermitian matrix in the following manner: $\mathbf{H} \equiv \mathbf{M} - \frac{i}{2}\mathbf{\Gamma}$. Where respectively both the mass matrix \mathbf{M} and the decay matrix $\mathbf{\Gamma}$ are Hermitian. Charge, Parity and Time reversal symmetry invariance requires that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. Furthermore since both \mathbf{M} and $\mathbf{\Gamma}$ are Hermitian we also have $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. Using this we can rewrite the Hamiltonian \mathbf{H} from equation 2.5 in the following manner:

$$\mathbf{H} \equiv \begin{pmatrix} H_0 & H_{12} \\ H_{21} & H_0 \end{pmatrix} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} \equiv \begin{pmatrix} M_s & M_{12} \\ M_{12}^* & M_s \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma_s & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_s \end{pmatrix} \quad (2.6)$$

Where $M_s = M_{11}$ is the mass of the B_s^0 state and $\Gamma_s = \Gamma_{11}$ is the decay width of the B_s^0 state. In order for equation 2.5 to be solved and obtain the time evolution of the $|B_s^0\rangle$ and $|\bar{B}_s^0\rangle$ states we need to decouple the system. This is done by transforming the flavor eigenstates in a manner that diagonalizes the Hamiltonian \mathbf{H} . These decoupled states are called the "mass eigenstates" of the system, which have a definite mass and lifetime. Using the transformation matrix \mathbf{P} , we can define the diagonalized Hamiltonian \mathbf{H}' and mass-eigenstate coefficients a' and b' as follows:

$$\mathbf{H}' = \mathbf{P}^{-1}\mathbf{H}\mathbf{P} \quad \text{and} \quad \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathbf{P} \begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} \quad (2.7)$$

Looking back at the Hamiltonian in equation 2.6 it can be shown that it has two eigenvalues, being $H_0 \pm \sqrt{H_{12}H_{21}}$ with corresponding eigenvectors of $(\sqrt{H_{12}}, \pm\sqrt{H_{21}})$. So the eigenvalues will be the new diagonal elements of \mathbf{H}' and the transformation matrix has to be built using the eigenvectors. The new diagonalized Hamiltonian \mathbf{H}' becomes:

$$\mathbf{H}' = \begin{pmatrix} H_0 - \sqrt{H_{12}H_{21}} & 0 \\ 0 & H_0 + \sqrt{H_{12}H_{21}} \end{pmatrix} \equiv \begin{pmatrix} M_L & 0 \\ 0 & M_H \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma_L & 0 \\ 0 & \Gamma_H \end{pmatrix} \quad (2.8)$$

Where we have taken the liberty of redefining the variables using the real matrices M and Γ similar to what we have seen before in equation 2.6. Here the subscript L (light) and H (heavy) are used to indicate the state with smaller and heavier mass respectively. Note that the eigenvalues for the mass eigenstates are: $\omega_{L,H} \equiv M_{L,H} - \frac{i}{2}\Gamma_{L,H}$. Similarly one can obtain an expression for the transformation matrix \mathbf{P} :

$$\mathbf{P} = \begin{pmatrix} \sqrt{H_{12}} & \sqrt{H_{12}} \\ -\sqrt{H_{21}} & +\sqrt{H_{21}} \end{pmatrix} \quad (2.9)$$

The mass and decay parameters of the states $|B_s^0\rangle$ and $|\bar{B}_s^0\rangle$ are related to the masses and decay widths of the new mass states B_L and B_H as defined in equation 2.8. By taking the sum and difference of the diagonal entries of \mathbf{H}' we can obtain the following expressions:

$$H_0 = \frac{1}{2}(M_H + M_L) - \frac{i}{4}(\Gamma_L + \Gamma_H) \quad (2.10a)$$

$$\sqrt{H_{12}H_{21}} = \frac{1}{2}(M_H - M_L) + \frac{i}{4}(\Gamma_L - \Gamma_H) \quad (2.10b)$$

Combining this with equation 2.6 it is also found that:

$$M_S \equiv \Re(H_0) = \frac{1}{2}(M_H + M_L) \quad (2.11a)$$

$$\Gamma_S \equiv -2\Im(H_0) = \frac{1}{2}(\Gamma_L + \Gamma_H) \quad (2.11b)$$

Useful to define for later use are the following quantities:

$$\begin{aligned} \Delta m_S &\equiv M_H - M_L = 2\Re(\sqrt{H_{12}H_{21}}) \\ &= 2\Re\left(\sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}\right) \end{aligned} \quad (2.12a)$$

$$\begin{aligned} \Delta\Gamma_S &\equiv \Gamma_L - \Gamma_H = 4\Im(\sqrt{H_{12}H_{21}}) \\ &= 4\Im\left(\sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}\right) \end{aligned} \quad (2.12b)$$

Where the choice for the definition of $\Delta\Gamma_S$ is made such that it is always expected to be a positive quantity in the standard model. With the quantities that we have just defined it is possible to write out the expression for

$\sqrt{H_{12}H_{21}}$ and we obtain the following:

$$\sqrt{H_{12}H_{21}} = \sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)} = \frac{1}{2}\Delta m_s + \frac{i}{4}\Delta\Gamma_s \quad (2.13)$$

Squaring this expression and comparing the real and imaginary parts of this equation yields the following relations:

$$\Delta m_s^2 - \frac{1}{4}\Delta\Gamma_s^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2 \quad (2.14a)$$

$$\Delta m_s\Delta\Gamma_s = -4\Re(M_{12}\Gamma_{12}^*) = 4|M_{12}||\Gamma_{12}|\cos(\phi_{12}) \quad (2.14b)$$

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

Coming back to the mass eigenstates of the system, we want to know how to express these mass eigenstates in terms of the flavor eigenstates that we started with. A general way of formulating this would be:

$$|B_L\rangle = p|B_s^0\rangle + q|\bar{B}_s^0\rangle \quad (2.15a)$$

$$|B_H\rangle = p|B_s^0\rangle - q|\bar{B}_s^0\rangle \quad (2.15b)$$

In this equation both p and q are complex numbers and are normalised such that $|p|^2 + |q|^2 = 1$. The value of the ratio of these numbers can be derived by using equation 2.6, equation 2.8 and equation 2.15 yielding:

$$\frac{q}{p} = -\sqrt{\frac{H_{21}}{H_{12}}} = e^{i\phi_m} \sqrt{\frac{|M_{12}| + \frac{i}{2}|\Gamma_{12}|e^{i\phi_{12}}}{|M_{12}| + \frac{i}{2}|\Gamma_{12}|e^{-i\phi_{12}}}} \quad (2.16)$$

Here $|M_{12}|, |\Gamma_{12}|$ are the magnitudes of M_{12} and Γ_{12} respectively and $\phi_m \equiv \arg(M_{12})$. Note that the ratio $|\frac{q}{p}|$ is only different from one if $\phi_{12} \neq 0, \pi$, which is known as CP violation in mixing. Using that $|\Gamma_{12}| \ll |M_{12}|$, which is expected to be true for $B_{(s)}^0$ mesons in the standard model[6], one can obtain the following expression:

$$1 - \left|\frac{q}{p}\right|^2 \approx \left|\frac{\Gamma_{12}}{M_{12}}\right| \sin(\phi_{12}) \quad (2.17)$$

Using the result obtained in equation 2.16 we can then determine the time evolution of the mass eigenstates by solving the Schrödinger equation

with the diagonalised Hamiltonian as given in equation 2.8. The resulting evolution relation for a meson produced in either the B_s^0 or \bar{B}_s^0 flavor state at $t = 0$ can now be written as:

$$|B_s^0(t)\rangle = g_+(t)|B_s^0\rangle + \frac{q}{p}g_-(t)|\bar{B}_s^0\rangle \quad (2.18a)$$

$$|\bar{B}_s^0(t)\rangle = g_+(t)|\bar{B}_s^0\rangle + \frac{p}{q}g_-(t)|B_s^0\rangle \quad (2.18b)$$

The functions $g_{\pm}(t)$ are defined as follows:

$$g_{\pm}(t) = \frac{1}{2}(e^{-i\omega_L} \pm e^{-i\omega_H}) \quad (2.19)$$

where the $\omega_{L,H}$ respectively, are the eigenvalues of the light and heavy mass eigenstates of the B-meson as mentioned earlier. And with this we have gathered the final building blocks for the decay process.

2.3 The decay process

After successfully describing the formalism of the mixing of the flavor/mass eigenstates of the system we can move on to describing the decay process of the Beauty mesons. As stated earlier, it is known that the B-meson is produced in either the $|B_s^0\rangle$ or $|\bar{B}_s^0\rangle$ state and the state evolves as dictated by equation 2.18. We can then label the amplitude of the state $|B_s^0\rangle(|\bar{B}_s^0\rangle)$ that decays into the final state $|f\rangle$ as $\mathcal{A}_f(\bar{\mathcal{A}}_f)$. For this paper we will focus on a specific decay, $B_s^0 \rightarrow J/\psi K_s^0$. In figure 2.3 a representation of the leading contributions to this decay are shown. Since there are two options of decay we can define the two decay amplitudes as:

$$\mathcal{A}_f \equiv \langle f|\mathcal{H}|B_s^0\rangle \quad \text{and} \quad \bar{\mathcal{A}}_f \equiv \langle f|\mathcal{H}|\bar{B}_s^0\rangle \quad (2.20)$$

Here \mathcal{H} is the Hamiltonian that is responsible for the decay of the B-meson. Thus for a meson produced in a state $|B_s^0\rangle$ there are two contributions to the decay width. One where the initial state and the state at time of decay are the same, and one where the $|B_s^0\rangle$ has oscillated into a $|\bar{B}_s^0\rangle$ before decaying. A graphical representation of this decay process is shown in figure 2.4 [6]. Similar to the process we have just taken a look at, there is a process of decay into the CP conjugate of the final state, $|\bar{f}\rangle$. This process also has two terms contributing to the decay width of the process. So overall

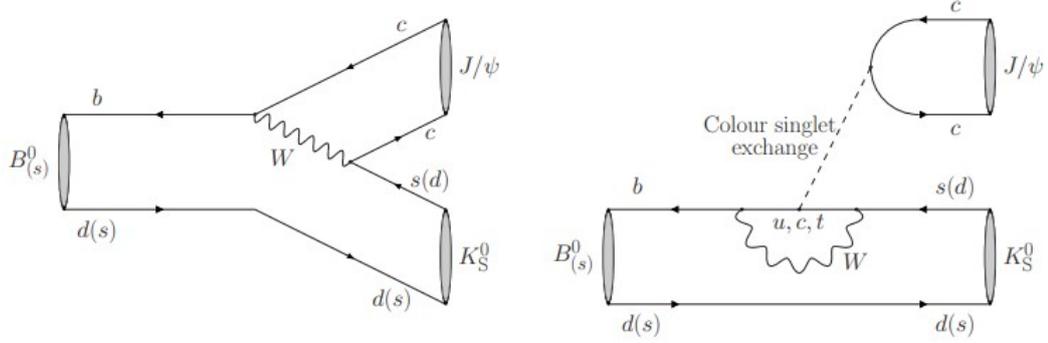


Figure 2.3: Diagrams that contribute to the $B_s^0 \rightarrow J/\psi K_s^0$ decay[1]

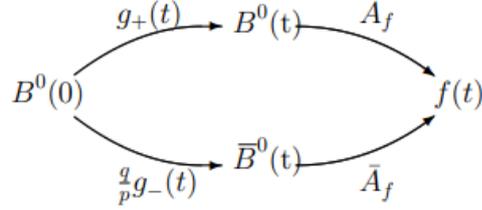


Figure 2.4: Graphical representation of the contributing processes to the decay width of a B_s^0 meson

we are left with four possible combinations of initial and final states and they are given by:

$$\begin{aligned} \mathcal{A}(B_s^0 \rightarrow f) &\propto g_+ \mathcal{A}_f + \frac{q}{p} g_- \bar{\mathcal{A}}_f & ; & \quad \mathcal{A}(B_s^0 \rightarrow \bar{f}) \propto \frac{q}{p} \left(g_- \bar{\mathcal{A}}_{\bar{f}} + \frac{p}{q} g_+ \mathcal{A}_{\bar{f}} \right) \\ \mathcal{A}(\bar{B}_s^0 \rightarrow f) &\propto \frac{p}{q} \left(g_- \mathcal{A}_f + \frac{q}{p} g_+ \bar{\mathcal{A}}_f \right) & ; & \quad \mathcal{A}(\bar{B}_s^0 \rightarrow \bar{f}) \propto g_+ \bar{\mathcal{A}}_{\bar{f}} + \frac{p}{q} g_- \mathcal{A}_{\bar{f}} \end{aligned} \quad (2.21)$$

Looking closely at these expressions one can see that they are very similar in structure. For example the expressions on the right side can be obtained from the expressions of the left hand side by interchanging p and q , by replacing \mathcal{A}_f with $\bar{\mathcal{A}}_{\bar{f}}$ and by replacing $\bar{\mathcal{A}}_f$ with $\mathcal{A}_{\bar{f}}$.

By squaring these process amplitudes one can then obtain the expressions for the differential decay rates in time. Working this out for the $|B_s^0\rangle$ to $|f\rangle$

amplitude gives:

$$\begin{aligned} \frac{d\Gamma(B_s^0 \rightarrow f)}{dt} &\propto |\mathcal{A}(B_s^0 \rightarrow f)|^2 \\ &\propto |g_+|^2 |\mathcal{A}_f|^2 + \left|\frac{q}{p}\right|^2 |g_-|^2 |\bar{\mathcal{A}}_f|^2 + 2\Re(g_+^* g_-) \Re\left(\frac{q}{p} \mathcal{A}_f^* \bar{\mathcal{A}}_f\right) \\ &\quad - 2\Im(g_+^* g_-) \Im\left(\frac{q}{p} \mathcal{A}_f^* \bar{\mathcal{A}}_f\right) \end{aligned} \quad (2.22)$$

Now the equation above can be simplified by taking a closer look at g_{\pm} , since their products are given by:

$$|g_{\pm}|^2 = \frac{1}{2} e^{-\Gamma_s t} \left[\cosh\left(\frac{1}{2} \Delta\Gamma_s t\right) \pm \cos(\Delta m_s t) \right] \quad (2.23a)$$

$$g_+^* g_- = \frac{1}{2} e^{-\Gamma_s t} \left[-\sinh\left(\frac{1}{2} \Delta\Gamma_s t\right) + i \sin(\Delta m_s t) \right] \quad (2.23b)$$

Combining equations 2.22 and 2.23 yields:

$$\begin{aligned} \frac{d\Gamma(B_s^0 \rightarrow f)}{dt} &\propto \frac{1}{2} |\mathcal{A}_f|^2 (1 + |\lambda_f|^2) e^{-\Gamma_s t} \\ &\quad \times \left[\cosh\left(\frac{1}{2} \Delta\Gamma_s t\right) + C_f \cos(\Delta m_s t) \right. \\ &\quad \left. + D_f \sinh\left(\frac{1}{2} \Delta\Gamma_s t\right) - S_f \sin(\Delta m_s t) \right] \end{aligned} \quad (2.24)$$

Where the following definitions have been made and used:

$$\lambda_f = \frac{q}{p} \frac{\bar{\mathcal{A}}_f}{\mathcal{A}_f} \quad (2.25a)$$

$$C_f \equiv \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2} \quad D_f \equiv -\frac{2\Re(\lambda_f)}{1 + |\lambda_f|^2} \quad S_f \equiv \frac{2\Im(\lambda_f)}{1 + |\lambda_f|^2} \quad (2.25b)$$

Note that the way we have defined these variables here might differ from other authors, with regards to the sign of each of the variables. The expressions for the differential decay rates of the other three processes can be obtained by applying to rules mentioned earlier to equation 2.24. In the case of $\bar{B}_s^0 \rightarrow f$ this would result in a change in the sign of both C_f and S_f and multiplying by a factor of $\left|\frac{p}{q}\right|^2$.

As we have seen earlier, both the B_s^0 and \bar{B}_s^0 meson are able to decay into

the final state. As a result the interference between decay and mixing gives rise to CP-violation. Specifically in the case where all of the contributing decay amplitudes have the same weak phase one can define the decay amplitude ratio as $\mathcal{A}_f/\bar{\mathcal{A}}_f = \eta_f e^{2i\phi_D}$. Where $\eta_f = \pm 1$ is the CP-eigenvalue of the final state and $\phi_D = \arg(A_f)$. In the range where the amount of CP violation in mixing is small, $|q/p| \approx 1$, one can also define the following parameter:

$$\lambda = \eta_f e^{-i\phi_m} e^{2i\phi_D} \quad (2.26)$$

Here we use the ϕ_m parameter that we defined earlier, underneath equation 2.16. Using the equation above and equations 2.21 and 2.22 we are able to define a time-dependent CP-violation parameter:

$$A_{CP} \equiv \frac{\Gamma_{B_s^0 \rightarrow f} - \Gamma_{\bar{B}_s^0 \rightarrow f}}{\Gamma_{B_s^0 \rightarrow f} + \Gamma_{\bar{B}_s^0 \rightarrow f}} = \frac{\eta_f \sin(\phi_f) \sin(\Delta m_s t)}{\cosh(\frac{1}{2}\Delta\Gamma_s t) - \eta_f \cos(\phi_f) \sinh(\frac{1}{2}\Delta\Gamma_s t)} \quad (2.27)$$

In this equation we have introduced the following quantity ϕ_f , known as the CP violating phase:

$$\phi_f \equiv -\arg(\lambda_f) = \phi_m - 2\phi_D \quad (2.28)$$

If we then measure the amplitude of the sinusoid part of the time-dependent asymmetry we can constrain the CP violating phase ϕ_f . We can use the fact that this parameter is then related to the phase of M_{12} , to use the CP-asymmetry as a kind of probe for the new contributions to M_{12} . What this phase is exactly depends on the type of decay we look at. For decays with a single contributing amplitude the phase ϕ_f can be described using simply the elements of V_{CKM} . In particular, there exist so called "golden modes", that occur when there is a $b \rightarrow c\bar{c}s$ transition:

$$B_s^0 \rightarrow J/\psi K_s \quad : \quad \phi_f = \phi_s^{c\bar{c}s} = -2\beta_s \quad (2.29)$$

Here β_s is defined as[8]:

$$\beta_s \equiv \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \quad (2.30)$$

where V_{ts} , V_{tb}^* , V_{cs} and V_{cb}^* are elements of the CKM matrix as given in equation 2.3. For this decay we can then obtain the value of ϕ_s from the value of S by using that $S = \sin(\phi_s)$. Having arrived at the definition of ϕ_s we have completed the discussion of the theoretical framework and can move on to the data analysis section.

Data Analysis

3.1 Flavor tagging

Since we know that both the B meson (B) and the anti-B meson (\bar{B}) can decay into the observed final state, see figure 2.4, we are left with the task of determining the initial state. As determining the initial state is essential for performing a time-dependent CP asymmetry measurement. The process we use for this is called 'flavor tagging' of which there are two different methods, Opposite Side (OS) tagging and Same Side (SS) tagging, see figure 3.1.

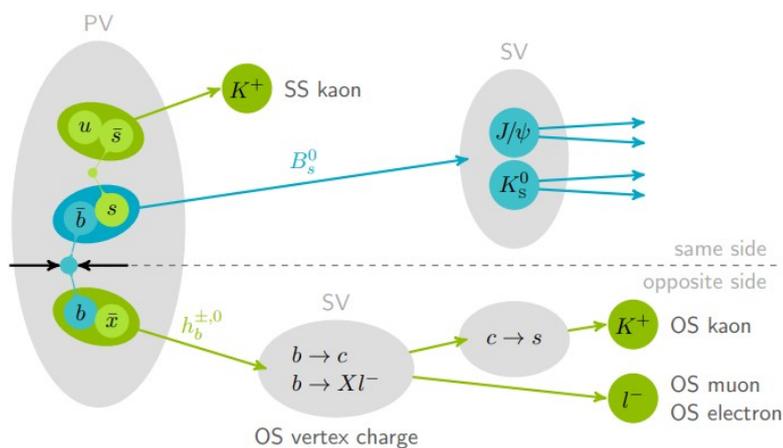


Figure 3.1: Schematic overview of the flavor tagging procedure.[9]

It is known that most b quarks are produced in $b\bar{b}$ pairs, which hadronise and decay independently. Using this one can infer the flavor of the initial state of the B before decay. This is used in OS tagging [10]. By tracking

the non-signal b as it hadronises and decays one can infer the flavor of the signal B meson. Algorithms are implemented that keep track of the decay products of the non-signal b decay.

The SS tagger makes use of that fact that the $s(\bar{s})$ that is bound together with the $\bar{b}(b)$ is also most likely produced in $s\bar{s}$ pairs [11]. The other quark of this pair $s(\bar{s})$ also tends to hadronise into a positively(negatively) charged kaon. This kaon charge then reveals the flavor of the B meson. Note that for the regular $B_{(d)}$ meson with a down quark in it, the other quark would hadronise into a pion. Therefore the tagging algorithm is specialised for either of these decay products, not both.

The performance of the flavor tagging is limited, especially in busy hadronic regions such as the LHC. This means that only a selection of N_R signal B mesons is tagged correctly, whereas there are also N_W mistagged events and N_U untagged events. Using these three numbers one can define the tagging efficiency:

$$\varepsilon_{tag} = \frac{N_R + N_W}{N_R + N_W + N_U} \quad (3.1)$$

With the mistag probability being:

$$\omega \equiv \frac{N_W}{N_R + N_W} \quad (3.2)$$

The reason for mentioning these tagging inefficiencies is that they end up diluting the amplitude of the signal by a factor of $\mathcal{D}_{tag} \equiv (1 - 2\omega)$. One can then attach a figure of merit to these tagging algorithms, which one should optimise to get the perfect performance. We call this the effective tagging efficiency and it is defined as:

$$\varepsilon_{eff} = \varepsilon_{tag}(1 - 2\omega)^2 = \varepsilon_{tag}\mathcal{D}_{tag}^2 \quad (3.3)$$

This tells us that a sample with N events that have $\varepsilon_{eff} = x$ have the same statistical power as xN perfectly sampled events. Typical values for this efficiency are around only a few percent at hadron colliders. For the OS tagger we assumed that the b and \bar{b} quarks have an equal chance of hadronising, in reality the chance for b and \bar{b} to hadronise differs slightly. This would lead to a different ε_{eff} for the $|B_s^0\rangle$ and $|\bar{B}_s^0\rangle$ states, albeit only slight difference.

During the analysis the OS and SS tagger each provide two variables: the tag decision q and an estimate of the probability of the tag being incorrect

η . The tag decision has three possible values +1, 0 and -1, meaning respectively that the candidate is tagged as a B_s , untagged, or tagged as a \bar{B}_s . The estimate η is determined by a neural-network per candidate by using kinematic and geometrical properties of the flavor identifying particles. The predicted mistag rate η needs to be calibrated in data to determine the true mistag rate ω using flavor-specific B meson decays. This means that other decay data is needed to calibrate the tagger before using it on the data of interest. The calibration function is parameterised as:

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle) \quad (3.4)$$

Here p_0 and p_1 are the calibration parameters and $\langle\eta\rangle$ is the mean of the η distribution as predicted by the tagging algorithm. If everything were perfect, meaning $\eta = \omega$, then $p_0 = \langle\eta\rangle$ and $p_1 = 1$. This calibration is done separately for the SS and OS taggers.

For this thesis we looked at the $B_s^0 \rightarrow J/\psi K_S$ decay and we ended up using the OS tagger since we did not select a data sample that allowed us to calibrate the other tagging algorithm. Taking into account the dilution caused by the flavor tagging process we can modify the equation for A_{CP} (equation 2.27) into the version of the asymmetry that we observe:

$$A_{CP}^{obs} = \frac{N_{B_s^0}^{obs} - N_{\bar{B}_s^0}^{obs}}{N_{B_s^0}^{obs} + N_{\bar{B}_s^0}^{obs}} = (1 - 2\omega)A_{CP} = \mathcal{D}_{tag}A_{CP} \quad (3.5)$$

where in the last step we used the fact that the observed particle yields can be modified to become:

$$N_{B_s^0}^{obs} = N_{B_s^0}(1 - \omega) + \omega N_{\bar{B}_s^0} \quad (3.6a)$$

$$N_{\bar{B}_s^0}^{obs} = N_{\bar{B}_s^0}(1 - \omega) + \omega N_{B_s^0} \quad (3.6b)$$

For the fit of the parameter ϕ_s decays into three different final states are used, these decays are:

$$\begin{array}{ll} B_s^0 \rightarrow K_S J/\psi & \text{where } J/\psi \rightarrow \mu^+ \mu^- \\ B_s^0 \rightarrow K_S \psi(2S) & \text{where } \psi(2S) \rightarrow \mu^+ \mu^- \\ B_s^0 \rightarrow K_S J/\psi & \text{where } J/\psi \rightarrow e^+ e^- \end{array}$$

In this project the algorithm that we use to fit the LHCb run 2 data also contains a procedure for the tagging of the data, additional decay processes of the Beauty meson were used for the calibration of the OS tagger.

These additional decays are:

$$\begin{aligned} B^+ &\rightarrow J/\psi K^+ \\ B^0 &\rightarrow J/\psi K^{*0} \quad \text{where} \quad K^{*0} \rightarrow K^+ \pi^- \end{aligned}$$

Note that the efficiency ε_{eff} may differ for each of the tagging procedures, which will follow from the result of the data analysis as shown in the next chapter.

3.2 Background subtraction

The raw data also contains a significant fraction of background events. As we are only interested in the relevant decay modes, we need to filter out this background. One method of doing so is by using the mass distribution of the data and using the sPlot method [12][13]. The data contains both information about the mass of the decaying particle and decay time of the particles, otherwise such an analysis would not be possible. An example of how the mass distribution is fitted with the probability density functions can be seen in figure 3.2. We should mention that the data that has been fitted in this figure contains only the $B_s^0 \rightarrow K_s J/\psi$ where $J/\psi \rightarrow \mu^+ \mu^-$ decay, whereas the LHCb run 2 data contains all three decay processes mentioned in the previous section. With this method we can assign weights to each event in the data calculated using a combined probability density function of the B_d^0 (first peak), the B_s^0 (second peak) and an estimate of the background. During the fit of the invariant mass, every event in the data set is given a weight based on the likelihood that is a background contribution. Using these weights from the sPlot method in combination with flavor tagging we can obtain data with the background events statistically removed. The parameters S and C are then obtained by fitting the decay time distribution to the model given in equation 2.24.

Using the sweights calculated from the B meson mass one can calculate directly the asymmetry between $|B_s^0\rangle$ and $|\bar{B}_s^0\rangle$ from their decay time data. This corresponds to one of the methods (BAF) used later on, during the analysis of the toy decay time data.

3.3 Time resolution

In section 3.1 we saw that the tagging procedure causes dilution of the data. This is not the only source that causes a dilution since we also have

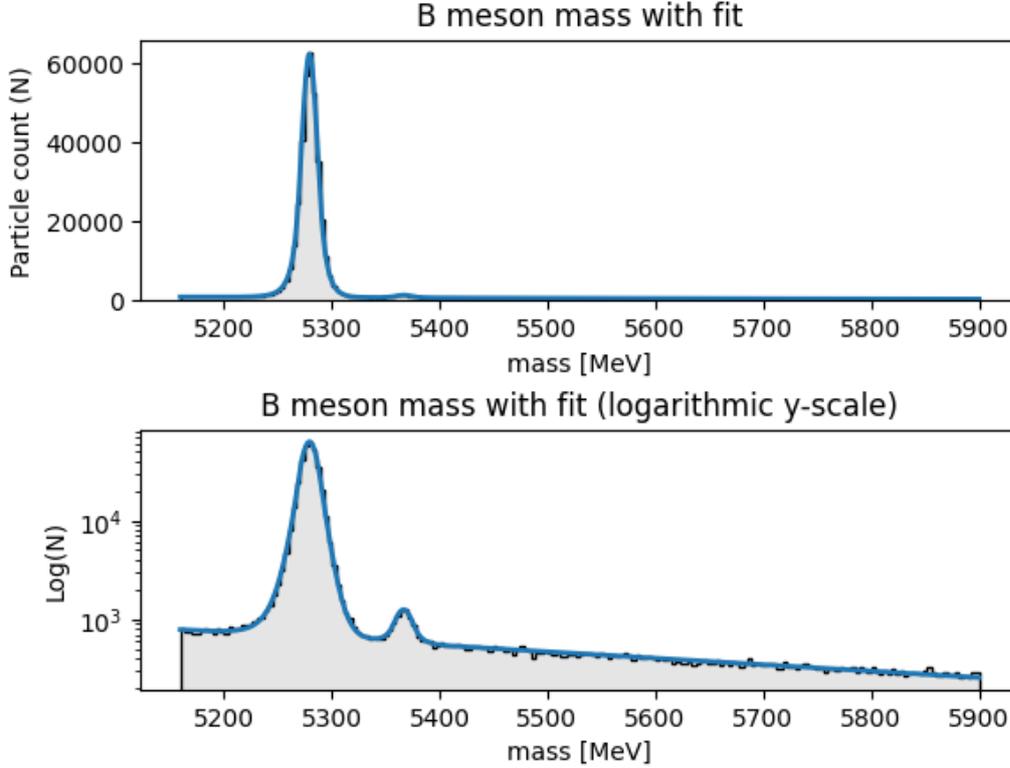


Figure 3.2: Figure showing fit (blue line) of data invariant mass distribution with regular(top) and logarithmic(bottom) y-scale.

the dilution caused by the limited decay time resolution of the detector σ_t . Assuming this resolution σ_t is Gaussian we can compute the dilution. The expression for the dilution is[14]:

$$\mathcal{D}_{reso} = \exp\left(-\frac{1}{2}\sigma_t^2 \Delta m_s^2\right) \quad (3.7)$$

Here Δm_s is the mass difference of the $|B_H\rangle$ and $|B_L\rangle$ mass states as defined in equation 2.12a. How good the resolution is, strongly depends on the final state. It is substantially worse for partially reconstructed decays than for fully reconstructed decays. At the LHCb experiment it is about 0.05 ps [6]. If we then plug in all of the numbers for the B_s^0 oscillation we get a dilution factor \mathcal{D}_{reso} of 0.7. Hence for measurements of the CP violating phase one must properly calibrate for this dilution. We can do this in same way as we did for the tagging dilution in equation 3.5. Obtaining the resolution can be done by running simulations or looking at processes with a so called "zero" decay time distribution, such as 'prompt' J/ψ production.

3.4 Validation of the fit method

3.4.1 CP violation variable from toy data

As stated before, we are interested in the CP-violation variable that characterizes the $B_s^0 \rightarrow J/\psi K_s^0$ decay. This variable can be obtained by fitting the decay time data, the formalism of which we discussed in chapter 2. Crucial to this process is understanding how the decay time data is shaped and why it is shaped that way. Like many other decay time plots we expect there to be an exponential decay, but this is not the only contributing factor that we need to keep in mind. The LHCb run 2 data that we will be fitting later on has undergone some processing already, similar to the effects we have seen in the previous sections of this chapter. An example of what the raw data would look like can be seen in the figure 3.3.

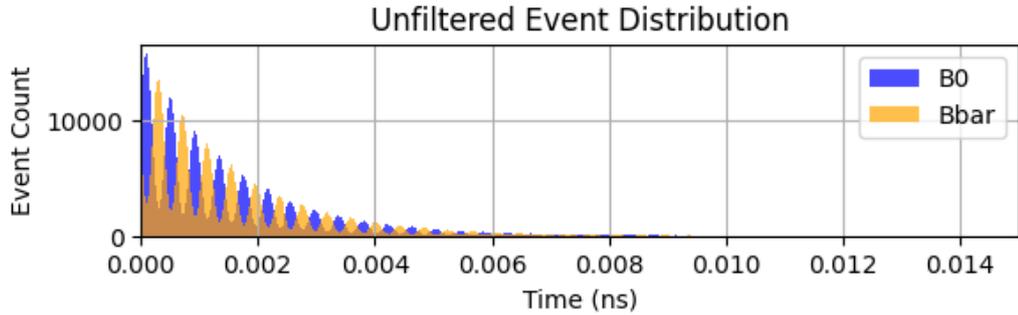


Figure 3.3: Histogram showing the event distribution of the unfiltered decay time toy data of the B_s^0 (\bar{B}_s^0) state, labeled B0 (Bbar).

For figure 3.3 a toy model has been used to generate data with fixed and predetermined values of $C_{val} = 0$, $S_{val} = -0.7$, $D_{val} = 0.51$, $\Delta\Gamma_s = 0ns^{-1}$ and $\Delta m_s = 17.765 * 10^{12} \hbar s^{-1}$. Here we used the subscript 'val' to indicate that these are the predetermined values and not the fit results. The amount of undiluted particles N was set to one million and the resolution of the toy data fitted in this figure is set to be near perfect. In other words the error σ_t in the decay time of the toy data is set to $1fs$.

We use the 'ROOT' package to generate the toy data, but the way that we generate the toy data does not incorporate the detector efficiency in any way. Making the toy model closer to the actual data will require the raw toy data to be multiplied by an efficiency factor that encompasses the

efficiency of the detector. This efficiency factor is chosen to be:

$$\varepsilon(t) = \begin{cases} 1 - e^{-\beta(t-t_0)} & \text{If } t > t_0 \\ 0 & \text{If } t \leq t_0 \end{cases} \quad (3.8)$$

This means that all data is discarded below a threshold decay time of t_0 . For all other decay times t there is a chance, dependent on the value of β , of it being discarded. This chance is larger for small t and the acceptance is higher for larger values of t . The value of t_0 was chosen to be $0.3ps$ and β was chosen to have a value of $1ps^{-1}$. We can then take a look at the effect this has on the toy data from figure 3.3. The result is shown in figure 3.4.

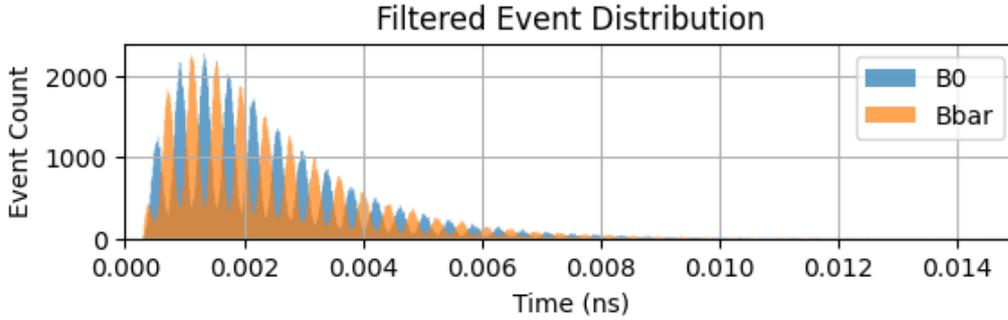


Figure 3.4: Histogram showing the event distribution of the filtered decay time toy data of the B_s^0 (\bar{B}_s^0) state, labeled B0 (Bbar).

One can clearly see the cutoff at t_0 and the reduced acceptance at low values of t . If we look at figures 3.3 and 3.4 we notice an oscillation in both B_s^0 and \bar{B}_s^0 toy decay time data. In equation 2.27 we defined what the CP asymmetry is, one can then define the asymmetry that we get from the data as follows:

$$A_{CP}^{data} = \frac{N_{B_s^0}^{data} - N_{\bar{B}_s^0}^{data}}{N_{B_s^0}^{data} + N_{\bar{B}_s^0}^{data}} \quad (3.9)$$

Where $N_{B_s^0}^{data}$ is the number of B mesons and $N_{\bar{B}_s^0}^{data}$ is the number of anti-B mesons. Using the previous equation we can then calculate the asymmetry of the data shown in figure 3.4, the result is shown in figure 3.5.

Note that the x-axis is different from those in figures 3.3 and 3.4. Only one oscillation period is shown out of convenience since the oscillation frequency is very high and showing multiple oscillations would result in

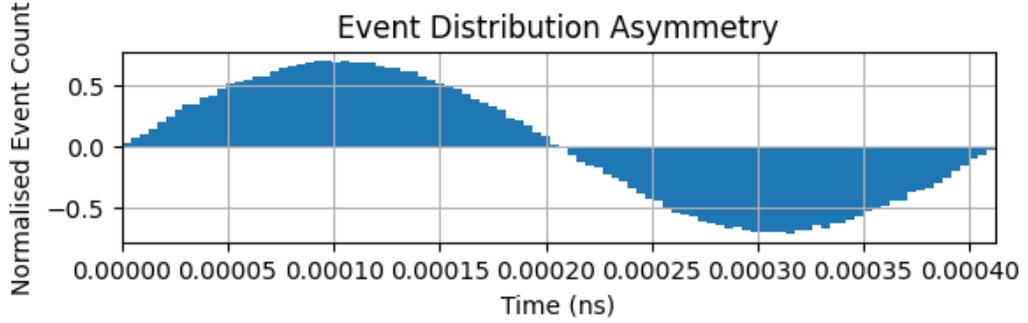


Figure 3.5: Histogram showing the asymmetry in the decay time distribution of the toy data.

an unclear figure. In order to make such a figure we need to fit the ‘folded’ decay time data instead of the decay time data. From theory we know the oscillation period τ of the oscillations in figure 3.4 is equal to $\frac{2\pi}{\Delta m_s}$. If we have an event with decay time t we choose the ‘folded’ decay time of that event to be $t_{fold} = t(\text{mod}\tau)$. By doing this for the entire decay time data set we obtain the ‘folded’ decay time data set. Figure 3.5 shows the asymmetry of the folded decay time data.

The next step is to extract the value of both C and S by fitting this asymmetry with equation 2.24. For the rest of the paper we will call these fit results S_{fit} and C_{fit} .

The toy data we generate is fitted with three different methods of fitting. We will briefly go over these three methods and the strengths and weaknesses they have. The first method is an unbinned likelihood fit (ULF) of the full model which we use later on to fit the LHCb run 2 data as well. This method is able to properly account for the time resolution σ_t but it is not able to include nonzero values of $\Delta\Gamma_s$. The second method is a binned asymmetry fit (BAF), which is essentially a chi squared fit, that uses the ‘zfit’ package. This method is limited since it does not account for the dilutions do to the tagging nor the tagging calibration. For the toy data samples this does not matter as they not contain any tagging information but this would cause a problem when fitting the LHCb run 2 data. The third method, called the RooFit (ROF) method, uses the ‘root’ package to fit the entire model and is able to properly fit data with a nonzero $\Delta\Gamma_s$. After making the toy data we improve the model by adding the effects of the efficiency, but this effect is not in the RooFit model. Therefore, the ROF method can not account for it. If we would have used RooFit to model the

efficiency, then RooFit would work perfectly. However, it is not easy to model the actual efficiency on the data.

We want to know how well the methods perform compared to each other. As we are going to use a method similar to ULF to fit the actual data. By then comparing the results we can check the influence of the toy data parameters on the result and accuracy of the toy data fits and check how well the models, the ULF in particular, perform. In chapter 4 we will show and discuss the result of these decay time toy data fits. The toy data fitting process can be summarised as follows:

1. Generate decay time toy data with 100000 pre-filter events with fixed parameters
2. Apply efficiency to the toy data
3. Fit for the parameters S and C using the three different fit methods
4. Repeat step 1-3 100 times and calculate the average value of the fit results

By repeating this process multiple times and varying the toy data generation parameters $\Delta\Gamma_s$, σ_t and S_{val} we can then infer the biases and influence of each of the parameter values on the fit accuracy. In the next section we will look more into the bias and uncertainty the methods have and we will discuss how to calculate them.

3.4.2 Bias and uncertainty of the fit

We want to estimate with the use of toy data how a mistake in the fitting model affects S and C . In order to achieve this we want to obtain a measure of the bias, which is defined as the estimate of the average error. This bias can be due to a limitation of the model, but it can also be because there are parameters in the model that are incorrect. An example of this would be: Fitting the data with the wrong resolution or tagging calibration. Besides the bias we also have the '(estimated) uncertainty'. This is an estimate of the Root Mean Square(RMS) of the error. Even though we do not know the true value, the variance of the observations still allow us to get an estimate of the RMS of the error. In this project we will use 'pull' distributions in the simulations to test if our fitting procedures estimate the statistical error correctly. So, essentially we use the pull to find a bias in the estimated

uncertainty. A pull value of a data set is calculated as follows:

$$\mathcal{P}(S) = \frac{S_{fit} - S_{val}}{\sigma_S} \quad (3.10)$$

Where the pull of S has been used as an example here, this is also the same for the pull of C as one simply replaces all the S terms with their respective C counterparts. Furthermore, as the names imply S_{fit} is the value of S from the fit, S_{val} is the S value of the generated toy data and σ_S is the estimated uncertainty in the fit of the S value. From these parameters another useful quantity called the 'root mean square' (RMS) can be calculated and it is given by:

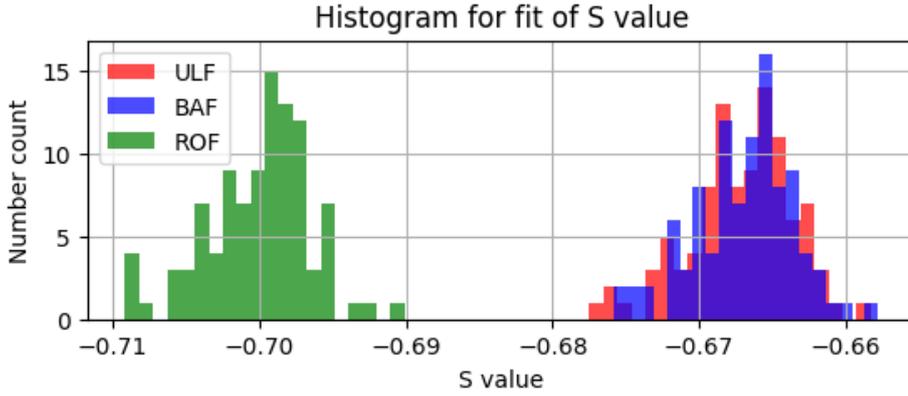
$$RMS(\mathcal{P}(S)) = \sqrt{(\mathcal{P}(S) - \langle \mathcal{P}(S) \rangle)^2} \quad (3.11)$$

Here $\langle \mathcal{P}(S) \rangle$ is the mean value of $\mathcal{P}(S)$, taken over 100 different toy data fits. Once again we have used S in this example but the equation for the RMS is also the same for $\mathcal{P}(C)$. Ideally the fit of the data should give a value of $S_{fit}(C_{fit})$ that is close to the value of $S_{val}(C_{val})$ that we used to generate the toy data with. We can calibrate the estimated uncertainty by using the width of the pull as a scale factor. A pull width of one would correspond to a correct estimation of the statistical error. The statistical error of the bias that we would calculate in that case would be $error/(N)^{0.5}$, where N is the amount of measurements that we average over. For a pull width of 2 this would mean that we have to multiply the estimated uncertainty by a factor of 0.5

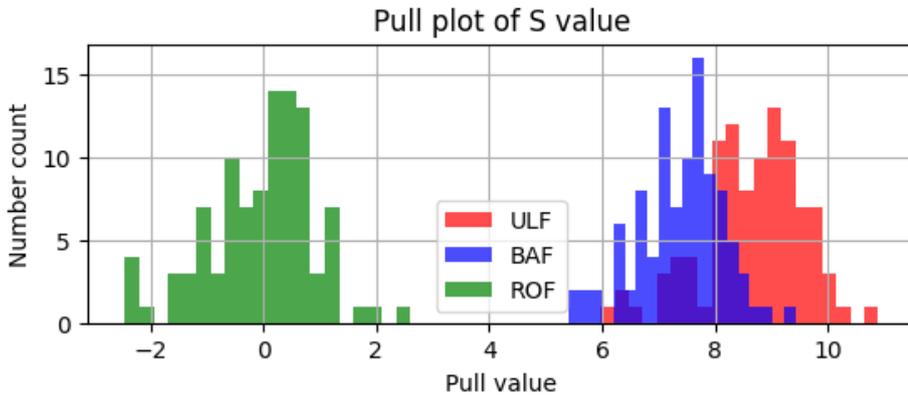
We know from equation 2.27 that the asymmetry also has terms that depend on the value of $\Delta\Gamma_s$. So far only the example where this parameter is equal to zero has been shown, but the value of this parameter in the actual data is nonzero. This means that we must also take a look at toy data where $\Delta\Gamma_s$ is nonzero. To be more specific, we want to look at the case where $\Delta\Gamma_s = 84.5ns^{-1}$, since this is the newest estimate of the actual value [15]. An overview of the fit values of S_{fit} and of C_{fit} and their pull plots are given in figures 3.6 and 3.7 for data generated with $S_{val} = -0.7$, $C_{val} = 0$ and $\Delta\Gamma_s = 84.5ns^{-1}$. For the toy data generation we have chosen to use multiple values, 0 and -0.7 , for S_{val} and to fix C_{val} at a value of zero. This allows to check the influence that the parameter S_{val} has on the bias and estimated uncertainty by comparing the two cases. For σ_t the choice of the three values was also straightforward, an ideal resolution of $1fs$, a realistic value of $50fs$ and a significantly worse than realistic value of $100fs$. For these experiments Δm_s being set equal to the latest estimate

which has a value of $17.765\hbar s^{-1}$.

Three different fit models have been used to determine whether the toy data is generated and fitted properly as seen in figures 3.6 and 3.7. In these figures it becomes clear that the ULF and BAF method are not able to include the non-zero $\Delta\Gamma_s$ properly resulting in a large bias in S . Whereas the ROF method, that does account for nonzero $\Delta\Gamma_s$, does not seem to have a large bias. With this we have enough information to understand the influence that a nonzero value of $\Delta\Gamma_s$ has for the ULF method.

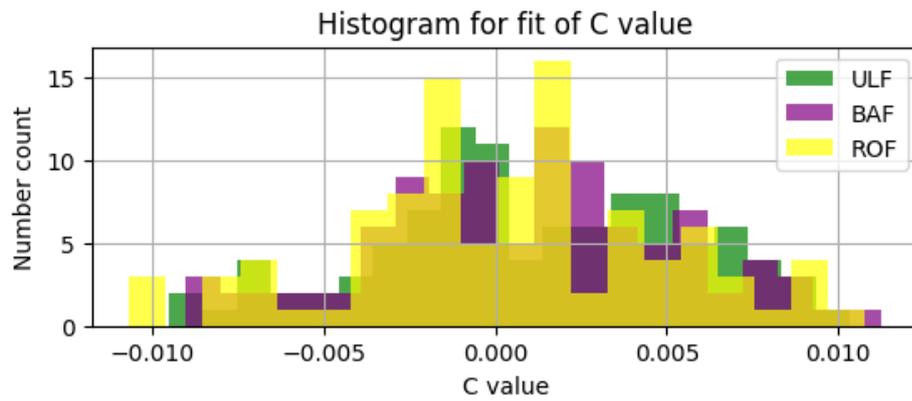


(a) S value plot.

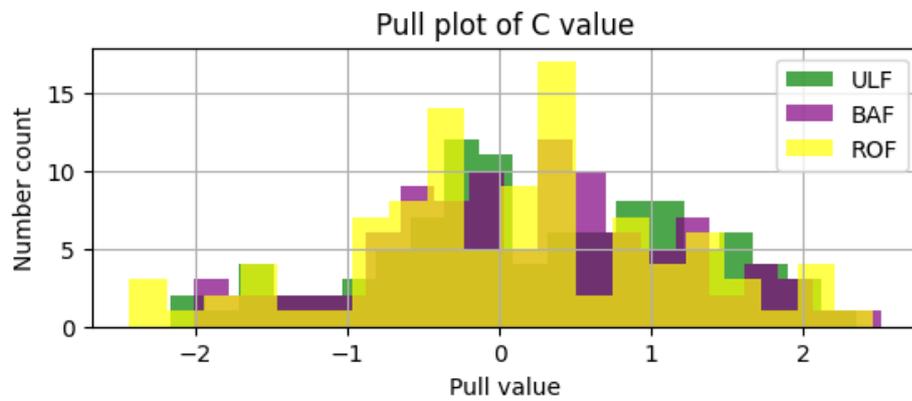


(b) Pull of S plot.

Figure 3.6: Figures showing an example of fit results from toy data fitting of S where $\Delta\Gamma_s = 84.5ns^{-1}$, $S_{val} = -0.7$ and $\sigma_t = 1fs$.



(a) C value plot.



(b) Pull of C plot.

Figure 3.7: Figures showing an example of fit results from toy data fitting of C where $\Delta\Gamma_s = 84.5\text{ns}^{-1}$, $S_{\text{val}} = -0.7$ and $\sigma_t = 1\text{fs}$.

3.4.3 Toy data fitting

We want to take a look at the performance of the three different fitting methods in order to get an idea of the size of the bias, if there is any. To achieve this we used the recipe described in chapter 3.4.1 to generate 100 toy decay time data sets of 100000 pre filter events each. This was done for every combination of the toy data parameters S_{val} , $\Delta\Gamma_s$ and σ_t . The results of this toy decay time data fitting are shown in tables 3.1, 3.2, 3.3, 3.4 and 3.5. In order these contain, the fit errors $\langle\sigma_S\rangle$ and $\langle\sigma_C\rangle$, the $\langle S\rangle$ data, the width of the pull of $\langle S\rangle$, the $\langle C\rangle$ data and the width of the pull of $\langle C\rangle$.

We can briefly go over all of the elements of each table column per column by looking at the items in the top row. Starting from the left we have the decay width $\Delta\Gamma_s$ given in units of ns^{-1} , followed by S_{val} which is the value for S of the toy data set, third is the time resolution σ_t with the units of fs and after that is 'Param' which indicates what result is shown in the next three columns. Here $\langle S\rangle(\langle C\rangle)$ corresponds to the average value of the one hundred fit results $S_{fit}(C_{fit})$ and $\langle RMS(\mathcal{P}(S))\rangle(\langle RMS(\mathcal{P}(C))\rangle)$ corresponds to the averaged width of the pull of $S_{fit}(C_{fit})$. The final three columns contain the relevant results for each of the three fitting methods.

3.4.4 Bias due to $\Delta\Gamma_s$

With that said and done it is time to take a look at the influence that the parameters have on the bias of S_{fit} and C_{fit} , starting off with the influence of $\Delta\Gamma_s$. From table 3.2 we can conclude that none of the methods have a bias in S when S_{val} and C_{val} are zero, irrespective of the value of $\Delta\Gamma_s$. The previous statement also holds for C upon closer inspection of table 3.4. This is something that we would have expected for the ULF method and the BAF method as they are unable to take into account the influence of $\Delta\Gamma_s$.

In the case where S_{val} is equal to -0.7 both the ULF method and the BAF method have an increase in the bias size of S_{fit} but for the ROF method there is hardly any change. The bias in C_{fit} does not seem to change for any of the three methods. This makes sense for the ROF method as it takes $\Delta\Gamma_s$ into account but there is also little to no effect for the other two methods. It seems that the influence of $\Delta\Gamma_s$ contributes more when the values of both S_{val} and C_{val} are zero. Which could mean that there are other effects at play that also lead to increases in the magnitude of the bias.

3.4.5 Influence of σ_t

This brings us to the parameter σ_t , here we compare the results of $\langle S_{fit} \rangle$ and $\langle C_{fit} \rangle$ with different values of σ_t that have the same $\Delta\Gamma_s$ and S_{val} . Starting with S_f by checking the results shown in table 3.2 it is clear, for every method, that there is an increase in the bias magnitude in the case where $S_{val} = 0$ irrespective of the value of $\Delta\Gamma_s$. This means that even though the ULF and BAF method correct for the value of σ_t it still results in a worse fit which means that these methods do not remove the effect of σ_t entirely. For the case where $S_{val} = -0.7$ we barely see an increase in the bias in S_f for any of the three methods, once again irrespective of the value of $\Delta\Gamma_s$, since the largest increase in the magnitude of the bias is 0.001.

Doing the same analysis of C_f by inspecting table 3.4 we see that in the case where $S_{val} = 0$ there is a significant increase of the bias magnitude as σ_t becomes larger. This increase is bigger for the BAF and ROF method than it is for the ULF method. When $S_{val} = -0.7$ we see that the bias increases as the value σ_t increases, but this increase in bias is much larger for the ULF and the BAF model than it is for the ROF model. When $\Delta\Gamma_s$ is nonzero this difference becomes even more apparent as the ROF model does not seem to be effected by the value of σ_t at all.

3.4.6 Pull width results

So far we have looked at the bias in both S_{fit} and C_{fit} but in order to gauge how well we have estimated the statistical uncertainty in the fit we must also look at tables 3.3 and 3.5. In an ideal scenario the analysis of the data would return a value that is one for the RMS of both $\mathcal{P}(S)$ and $\mathcal{P}(C)$, as this would mean that we have correctly estimated the statistical uncertainty in our fit of the data. When we inspect table 3.3 we see that for all three models most of the cases that the RMS of the pull is between 0.8 and 1.0. Which means that the statistical uncertainty has been slightly overestimated resulting in a narrower pull width. It is only in the case where $\Delta\Gamma_s = 0$, $S_{val} = 0$ and $\sigma_t = 1fs$ that we underestimate the statistical uncertainty slightly for all methods. The BAF method in particular seems to overestimate the statistical uncertainty as the pull width is even smaller than 0.8 in some cases. In table 3.5 we can see that the pull width of C_f is mostly between 0.9 and 1 for all three methods. This means that the statistical uncertainty is estimated better for C_f compared to S_f , although there is still a slight overestimation of the statistical uncertainty. So for most cases the three different methods are able to correctly evaluate the

statistical uncertainty.

3.4.7 Fit error in S_{fit} and C_{fit}

In addition to the pull width and mean value of the fit of S_f and C_f we have also calculated the average statistical uncertainty for all the toy data sets as seen in table 3.1. Using these values we can make an estimate of the error in the fit of S_{fit} and C_{fit} that is different from the bias we looked at earlier. Also of importance is the spread of the average statistical uncertainty as this tells us how much each toy data set varies and how big the difference in accuracy of the methods is. An example of this statistical uncertainty spread for both S_f and C_f is shown in figure 3.8. In these figures we can clearly see that all of the generated data sets differ, albeit only slightly and we see that all of three methods do indeed have an error of a similar magnitude for both S_f and C_f . We want to know the statistical uncertainty of the fit as this is important to know for the LHCb run 2 data analysis, for which we will use the statistical uncertainty value of the ULF method since this method is used for the LHCb run 2 data analysis. Looking at table 3.1 we see that there is a statistical uncertainty in S of $\sigma_S = 6.30e-3$ and a statistical uncertainty in C of $\sigma_C = 6.30e-3$, here $\Delta\Gamma_s = 84.5ns^{-1}$, $S_{val} = -0.7$ and $\sigma_t = 50fs$. We use these values since they are most representative of the expected parameter values of the LHCb run 2 data.

$\Delta\Gamma_s$	S_{val}	Param	ULF	BAF	ROF
0	0	$\langle\sigma_S\rangle$	6.30e-3	5.98e-3	6.29e-3
0	0	$\langle\sigma_C\rangle$	6.30e-3	5.98e-3	6.36e-3
84.5	-0.7	$\langle\sigma_S\rangle$	5.67e-3	5.98e-3	5.54e-3
84.5	-0.7	$\langle\sigma_C\rangle$	6.09e-3	5.98e-3	6.02e-3

Table 3.1: Table containing the toy decay time data fit results of the statistical uncertainty σ_S and σ_C with varying parameters.

3.4.8 Method bias comparison

When we combine the results of the influence that the different parameters have on the bias in both S_f and C_f and the size of pull width of S_f and C_f we can gauge the performance of each model. In particular we want to know how well the ULF method has performed at fitting the toy data samples. For the fitting of S the ULF model performed comparable to the

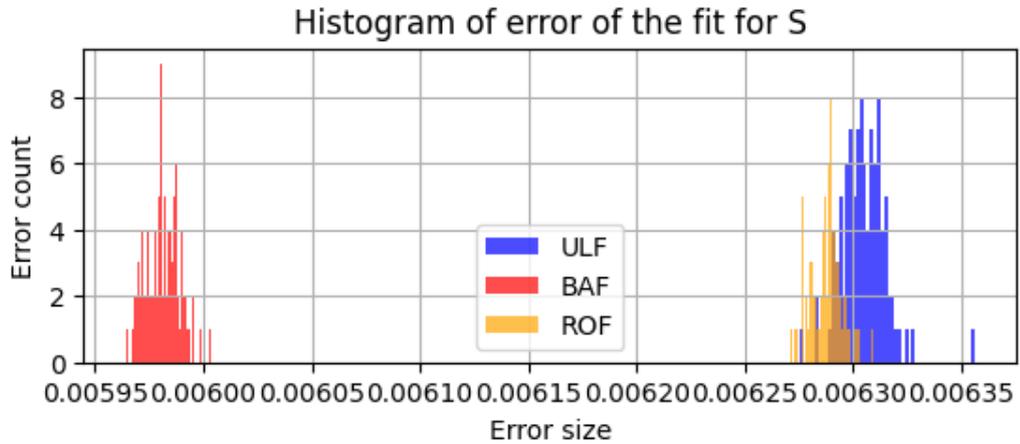
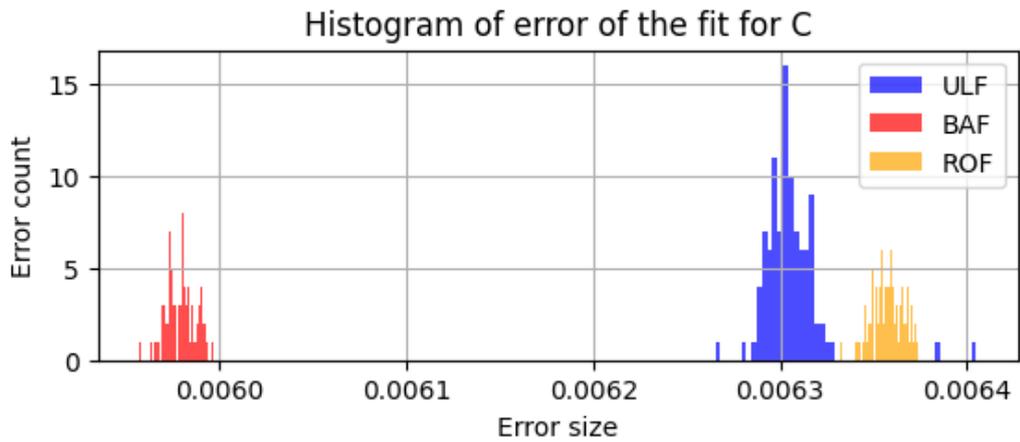
(a) σ_S values for toy data(b) σ_C values for toy data

Figure 3.8: Figures showing the spread of the statistical uncertainty for the toy data fit with $\Delta\Gamma_s = 84.5\text{ns}^{-1}$, $S_{val} = 0$ and $\sigma_t = 50\text{fs}$.

other two models in the case where $S_{val} = 0$ but for the toy data with $S_{val} = -0.7$ and $\Delta\Gamma_s = 84.5\text{ns}^{-1}$ it became apparent that both the ULF and BAF method performed worse than the ROF method with the largest bias being 0.033. This can not be caused solely by the statistical uncertainty of the fit as the value of this parameter for S is one order of magnitude smaller. Looking at the results for the fit value of C tells us that that all three methods performed similar when $S_{val} = 0$ with the ULF method being slightly less biased. For the case where $S_{val} = 0.7$ it becomes clear that the ULF and BLF methods had a very large bias compared to the ROF

method both when $\Delta\Gamma_s$ was zero and nonzero. We found that the size of the bias in C is 0.010 for the ULF model and it is close to the statistical uncertainty of the fit in C , $\sigma_C = 6.30e-3$. Having completed the toy data analysis we can now move on to the LHCb run 2 data analysis.

$\Delta\Gamma_s$	S_{val}	σ_t	Param	ULF	BAF	ROF
0	0	1	$\langle S \rangle$	0.000	0.000	0.000
0	0	50	$\langle S \rangle$	0.000	0.000	0.000
0	0	100	$\langle S \rangle$	-0.001	-0.001	-0.001
84.5	0	1	$\langle S \rangle$	0.000	0.000	0.000
84.5	0	50	$\langle S \rangle$	0.000	0.000	0.000
84.5	0	100	$\langle S \rangle$	-0.002	-0.002	-0.002
0	-0.7	1	$\langle S \rangle$	-0.700	-0.670	-0.700
0	-0.7	50	$\langle S \rangle$	-0.699	-0.699	-0.699
0	-0.7	100	$\langle S \rangle$	-0.699	-0.699	-0.699
84.5	-0.7	1	$\langle S \rangle$	-0.667	-0.667	-0.700
84.5	-0.7	50	$\langle S \rangle$	-0.667	-0.667	-0.700
84.5	-0.7	100	$\langle S \rangle$	-0.667	-0.667	-0.700

Table 3.2: Table containing the toy decay time data fit results $\langle S \rangle$ with varying parameters.

$\Delta\Gamma_s$	S_{val}	σ_t	Param	ULF	BAF	ROF
0	0	1	$\langle RMS(\mathcal{P}(S)) \rangle$	1.102	1.107	1.112
0	0	50	$\langle RMS(\mathcal{P}(S)) \rangle$	0.987	1.020	1.030
0	0	100	$\langle RMS(\mathcal{P}(S)) \rangle$	1.012	0.991	1.042
84.5	0	1	$\langle RMS(\mathcal{P}(S)) \rangle$	0.872	0.816	0.818
84.5	0	50	$\langle RMS(\mathcal{P}(S)) \rangle$	0.874	0.876	0.880
84.5	0	100	$\langle RMS(\mathcal{P}(S)) \rangle$	1.011	0.942	0.953
0	-0.7	1	$\langle RMS(\mathcal{P}(S)) \rangle$	0.904	0.729	0.890
0	-0.7	50	$\langle RMS(\mathcal{P}(S)) \rangle$	0.893	0.784	0.911
0	-0.7	100	$\langle RMS(\mathcal{P}(S)) \rangle$	0.947	0.928	1.021
84.5	-0.7	1	$\langle RMS(\mathcal{P}(S)) \rangle$	0.908	0.750	0.946
84.5	-0.7	50	$\langle RMS(\mathcal{P}(S)) \rangle$	0.898	0.758	0.837
84.5	-0.7	100	$\langle RMS(\mathcal{P}(S)) \rangle$	0.879	0.889	0.902

Table 3.3: Table containing the toy decay time data fit results of the pull width of $\langle S \rangle$ with varying parameters.

$\Delta\Gamma_s$	S_{val}	σ_t	Param	ULF	BAF	ROF
0	0	1	$\langle C \rangle$	0.000	0.000	0.000
0	0	50	$\langle C \rangle$	0.001	0.000	0.000
0	0	100	$\langle C \rangle$	0.001	0.002	0.001
84.5	0	1	$\langle C \rangle$	0.000	0.000	0.000
84.5	0	50	$\langle C \rangle$	-0.001	-0.001	-0.001
84.5	0	100	$\langle C \rangle$	-0.001	-0.002	-0.002
0	-0.7	1	$\langle C \rangle$	-0.089	-0.087	-0.076
0	-0.7	50	$\langle C \rangle$	0.010	0.011	0.000
0	-0.7	100	$\langle C \rangle$	0.018	0.017	0.001
84.5	-0.7	1	$\langle C \rangle$	0.001	0.001	0.000
84.5	-0.7	50	$\langle C \rangle$	0.010	0.010	0.000
84.5	-0.7	100	$\langle C \rangle$	0.014	0.015	0.000

Table 3.4: Table containing the toy decay time data fit results $\langle C \rangle$ with varying parameters.

$\Delta\Gamma_s$	S_{val}	σ_t	Param	ULF	BAF	ROF
0	0	1	$\langle RMS(\mathcal{P}(C)) \rangle$	0.930	0.914	0.913
0	0	50	$\langle RMS(\mathcal{P}(C)) \rangle$	0.902	0.890	0.884
0	0	100	$\langle RMS(\mathcal{P}(C)) \rangle$	0.913	0.923	0.923
84.5	0	1	$\langle RMS(\mathcal{P}(C)) \rangle$	1.095	1.136	1.139
84.5	0	50	$\langle RMS(\mathcal{P}(C)) \rangle$	1.027	1.051	1.057
84.5	0	100	$\langle RMS(\mathcal{P}(C)) \rangle$	1.067	1.049	1.056
0	-0.7	1	$\langle RMS(\mathcal{P}(C)) \rangle$	0.928	0.879	0.928
0	-0.7	50	$\langle RMS(\mathcal{P}(C)) \rangle$	1.004	0.984	1.005
0	-0.7	100	$\langle RMS(\mathcal{P}(C)) \rangle$	1.140	1.075	1.022
84.5	-0.7	1	$\langle RMS(\mathcal{P}(C)) \rangle$	1.020	0.968	1.036
84.5	-0.7	50	$\langle RMS(\mathcal{P}(C)) \rangle$	0.979	0.961	0.990
84.5	-0.7	100	$\langle RMS(\mathcal{P}(C)) \rangle$	1.038	1.043	1.035

Table 3.5: Table containing the toy decay time data fit results of the pull width of $\langle C \rangle$ with varying parameters.

Results

Using the ULF method we fitted the LHCb run 2 data. This data contains the $B_s^0 \rightarrow J/\psi K_s^0$ decays, the $B^+ \rightarrow J/\psi K^+$ decay and the $B^0 \rightarrow J/\psi K^{*0}$ decay where $K^{*0} \rightarrow K^+ \pi^-$ decay. Where the last two decays are needed to calibrate the OS tagger. With this fit we obtained the following results:

$K_S J/\psi \mu^+ \mu^-$ (total yield)	$= 3827 \pm 102$
$K_S J/\psi \mu^+ \mu^-$ (tagged yield)	$= 97.3 \pm 7.63$
$K_S J/\psi \mu^+ \mu^-$ (ϵ_{eff})	$= 0.025 \pm 0.002$
$K_S \psi(2S) \mu^+ \mu^-$ (total yield)	$= 246 \pm 34.7$
$K_S \psi(2S) \mu^+ \mu^-$ (tagged yield)	$= 7.49 \pm 2.65$
$K_S \psi(2S) \mu^+ \mu^-$ (ϵ_{eff})	$= 0.030 \pm 0.010$
$K_S J/\psi e^+ e^-$ (total yield)	$= 488 \pm 56.9$
$K_S J/\psi e^+ e^-$ (tagged yield)	$= 11.9 \pm 5.17$
$K_S J/\psi e^+ e^-$ (ϵ_{eff})	$= 0.024 \pm 0.010$
C	$= 0.08 \pm 0.33$
S	$= -0.11 \pm 0.29$

The first thing we can look at is the yield of all the different decay processes. We see that the $B_s^0 \rightarrow J/\psi \mu^+ \mu^-$ makes up most of the data. Combining all of the different processes we are left with a tagged yield of 116.9 events which is much smaller than the amount of events we used in the toy data analysis. The fact that the total amount of events is not a round number is caused by two factors. The first factor being the background

subtraction process and the second factor is the tagging efficiency ε_{eff} being included in the 'tagged' yields. This low yield is due to the OS flavor tagging efficiency that, for all of the different processes, is between 2% and 3% which results in most events being filtered out. This efficiency would have been better if the SS tagging method was also used but for this project that was not done for reasons mentioned earlier.

Since we have fewer events than in the toy data we expect the statistical uncertainty from the fit to be much larger than the bias resulting from the toy data fitting. We know that the statistical uncertainty depends on the amount of events in the following manner $\sigma_{stat} \propto 1/\sqrt{N_{events}}$. For the toy data measurements we had 100 data sets with 49000 events each, after applying the efficiency. Plugging in these numbers tells us that the statistical uncertainty is expected to be about 60 times larger for the LHCb run 2 data.

When we look the value of the fit results for S and C this is exactly what we see, as we obtain a value of 0.08 ± 0.33 for S and -0.11 ± 0.29 for C . The decay time distribution of the weighted B_s^0 data from which we obtained these results is shown in figure 4.1 where we can see that the data has a shape similar to that of figure 3.4. Two plots showing the asymmetry of the LHCb run 2 $J/\psi \rightarrow \mu^+\mu^-$ decay data with fits of the complete decay data are shown in figure 4.2. The unfolded asymmetry shows multiple oscillations and the folded asymmetry shows only one for the same reason as mentioned in section 3.4.1. Note also, that the tagging calibration has not been taken into account for these asymmetry plots.

From theory it is expected that value of both S and C are very close to zero[2]. We see that results of both S and C differ within one σ from zero, therefore the result agrees with what the theory predicts. If we compare the errors of these results with those obtained from the toy data analysis done in the previous section we can indeed see that they are larger than the bias for both S and C which, in the worst case, was around ± 0.033 for S and ± 0.010 for C . Hence the biggest contribution to the total error magnitude in S and C is due to statistics and not the bias resulting from the fit model(ULF) that was used. It also possible to explore the statistical uncertainty of the model by generating toy data samples with a number of events similar to that of the actual data but this would take significantly more time to do which was not possible for this project of short duration.

An additional way to judge the results is to compare them to results for

S and C that others obtained. We can compare it to an independent analysis of the same decay process but for LHCb run 1 data done by R. Aaij et al. 2013[1] where they found values of $-0.08 \pm 0.40(\text{stat}) \pm 0.08(\text{syst})$ for S and $-0.28 \pm 0.41(\text{stat}) \pm 0.08(\text{syst})$ for C . These results are comparable to ours as both their results for S and C are within one σ different from zero and have an uncertainty of a similar magnitude. The fact that our uncertainty is smaller can be explained by the fact that our data set was larger than theirs was. But our statistical uncertainty would have been even smaller had we been able to calibrate and use the SS tagger.

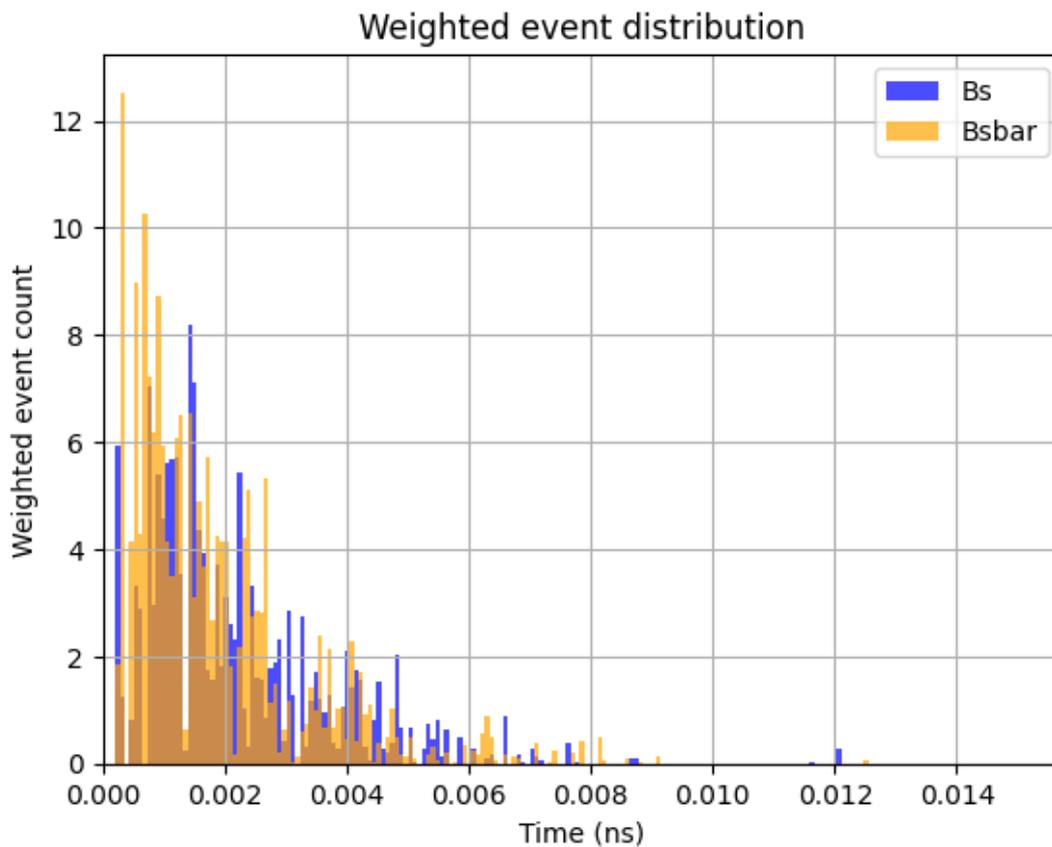
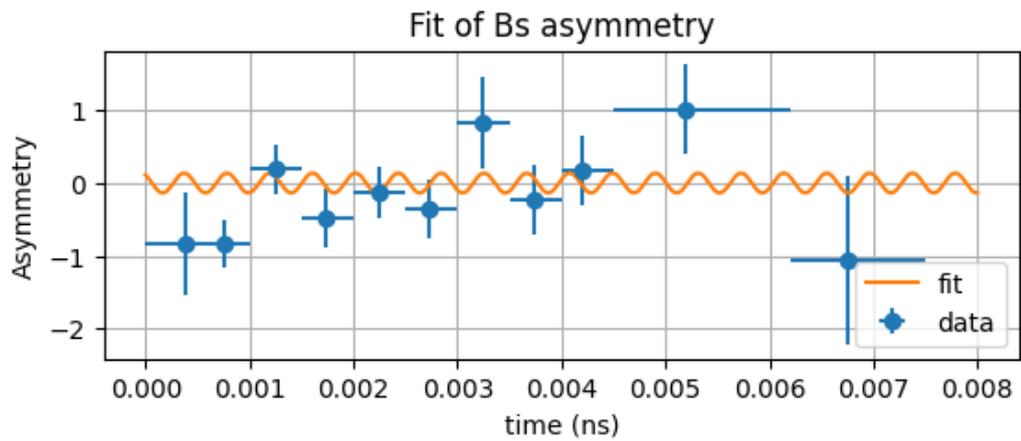
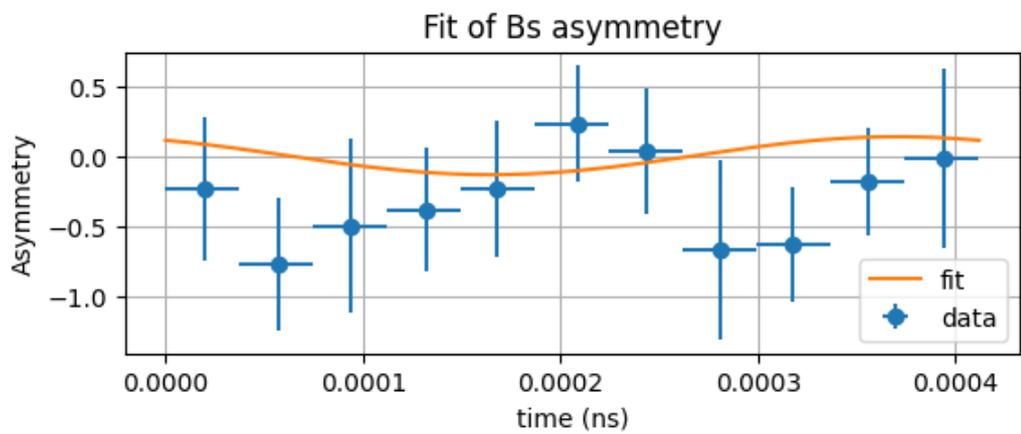


Figure 4.1: Weighted B_s^0 and \bar{B}_s^0 event distribution of LHCb run 2 data.



(a) Unfolded LHCb run 2 $J/\psi \rightarrow \mu^+\mu^-$ decay time data asymmetry with full asymmetry data fit.



(b) Folded LHCb run 2 $J/\psi \rightarrow \mu^+\mu^-$ decay time data asymmetry with full asymmetry data fit.

Figure 4.2: Figures showing the unfolded(a) and folded(b) LHCb run 2 $J/\psi \rightarrow \mu^+\mu^-$ decay data uncorrected for the tagging calibration, with a fit of the complete B_s meson asymmetry data.

Conclusion

Overall we see that we are able to successfully extract S and C from the LHCb run 2 data by using the ULF method where we obtained that $S = -0.11 \pm 0.29(\text{stat}) \pm 0.03(\text{syst})$ and $C = 0.08 \pm 0.33(\text{stat}) \pm 0.01(\text{syst})$. When comparing these results to their theoretical predictions we find no inconsistencies and if we look at another older analysis for the same parameters the size of the statistical uncertainty in both S and C was smaller in our analysis but the estimates were comparable. During this project we also looked into the effect that the value of S_{val} and $\Delta\Gamma_s$ have on the toy data fit results of S and C . We found that there is a bias in both S and C caused by the fit method(ULF), being ± 0.033 for S and ± 0.010 for C , even though this is small compared to the statistical uncertainties of S and C . The bias can not be discarded as it is independent from the statistical uncertainty of the fitting method.

To improve the efficiency of future analysis of the same LHCb run 2 data there is much to be gained by reducing the statistical uncertainty. A key factor in this improvement is the flavor tagging efficiency which is about 2%-3% in our analysis since we only used the OS flavor tagging procedure. Using the SS tagging procedure will increase the efficiency of the flavor tagging procedure. Doing so in future projects one would need to generate more toy data samples, or look at the $B_s^0 \rightarrow J/\psi \phi$ as the angle β_s can be measured here with more precision[9].

Bibliography

- [1] R. Aaij et al. Measurement of the time-dependent cp asymmetry in $B^0 \rightarrow J/\psi K_S^0$ decays. *Physics Letters B*, 721(1):24–31, 2013.
- [2] Kristof De Bruyn and Robert Fleischer. A roadmap to control penguin effects in $B_d^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi \phi$. *Journal of High Energy Physics*, 2015(3), March 2015.
- [3] R. Aaij et al. First observation of cp violation in the decays of B_s^0 mesons. *Physical Review Letters*, 110(22), May 2013.
- [4] J. van Leerdam. *Measurement of CP Violation in Mixing and Decay of Strange Beauty Mesons*. Phd-thesis - research and graduation internal, Vrije Universiteit Amsterdam, 2016.
- [5] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973.
- [6] Wouter Hulsbergen. Constraining new physics in b_s^0 meson mixing. *Modern Physics Letters A*, 28(27):1330023, August 2013.
- [7] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Phys. Rev. D*, 2:1285–1292, Oct 1970.
- [8] Claudio O. Dib, Isard Dunietz, Frederick J. Gilman, and Yosef Nir. Standard-model predictions for CP violation in B^0 -meson decay. *Phys. Rev. D*, 41:1522–1530, Mar 1990.
- [9] K.A.M. De Bruyn. *Searching for Penguin Footprints: Towards High Precision CP Violation Measurements in the B Meson Systems*. Phd-thesis - research and graduation internal, Vrije Universiteit Amsterdam, 2015.

- [10] R. Aaij et al. Opposite-side flavour tagging of b mesons at the lhcb experiment. *The European Physical Journal C*, 72(6), June 2012.
- [11] Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data. 2012. Linked to LHCb-ANA-2011-103.
- [12] M. Pivk and F.R. Le Diberder. : A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 555(1â2):356â369, December 2005.
- [13] Yuehong Xie. sfit: a method for background subtraction in maximum likelihood fit, 2009.
- [14] H.-G Moser and A Roussarie. Mathematical methods for $b_0b_1^0$ oscillation analyses. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 384(2):491–505, 1997.
- [15] LHCb collaboration et al. Improved measurement of cp violation parameters in $b_s^0 \rightarrow j/\psi k^+ k^-$ decays in the vicinity of the $\phi(1020)$ resonance, 2023.