Search for CPT and Lorentz violation in B^0 mixing with $B^0 \rightarrow J/\psi\,K^0_{\mathrm{S}}$ at LHCb

M.C. van Veghel

Universiteit Utrecht, Utrecht, the Netherlands

Supervisors:
dr. J. van Tilburg[†] ,
prof. dr. R. Snellings[‡]

Abstract

Effects of CPT and Lorentz violation in neutral meson mixing, in particular the sensitivity of the LHCb detector, have been studied. To parametrize CPT and Lorentz violation, the Standard Model Extension has been used. At LHCb, B^0 mixing in $B^0 \rightarrow J/\psi\,K^0_{\mathrm{S}}$ has the highest sensetivity to Standard Model Extension parameter $\Delta a_\mu^{B^0}$. Using a decay-time dependent analysis and the periodogram method, B^0 mixing has been measured to be consistent with no CPT and no Lorentz violation. The likelihood fit, without systematic uncertainties, resulted in

$$\Delta a_0^{B^0} - 0.38\Delta a_Z^{B^0} = (-0.8 \pm 1.1\ (stat.)) \times 10^{-15}\,\mathrm{GeV},$$
$$\Delta a_X^{B^0} = (-0.7 \pm 1.7\ (stat.)) \times 10^{-15}\,\mathrm{GeV},$$
$$\Delta a_Y^{B^0} = (+0.6 \pm 1.6\ (stat.)) \times 10^{-15}\,\mathrm{GeV},$$

making this the world's best measurement of $\Delta a_\mu^{B^0}$ to date.

† *Nikhef, Amsterdam, the Netherlands*
‡ *Universiteit Utrecht, Utrecht, the Netherlands*

Contents

1　Introduction

Throughout the history of physics, symmetries play an important role. The highly succesful theory describing subatomic particles and their interactions, the Standard Model, is no exception to this. Not only are the forces in the Standard Model manifestations of underlying symmetries, the Standard Model is built on certain fundamental symmetries as well. For example, the Lorentz transformation and the combination of charge conjugation, parity transformation and time reversal (CPT) are exact symmetries in the Standard Model. In quantum field theories, CPT and Lorentz violation are deeply connected.

Despite its succes, the Standard Model cannot describe gravity. At the high energies at the Planck scale, gravity becomes a significant force for subatomic particles. Hypothetical theories that are able to describe Planck scale physics, like string theory, are not necessarily symmetric under a CPT or Lorentz transformation. This way, CPT and Lorentz violation searches provide a test of fundamental properties of the Standard Model and may give signs of Planck scale physics. In addition, it might even contribute to baryogenesis [1].

A good method to investigate CPT and Lorentz violation is neutral meson mixing. In neutral meson mixing, particle and antiparticle states mix. This interference is very sensitive to small property differences between particle and antiparticle, including CPT violation. LHCb is one of the main experiments looking into neutral meson mixing. As an example of the power of these kind of interference experiments, the mass of the top quark could be constrained with neutral meson interferometry, long before the top quark was discovered and provided important information in the direct search for the top quark. LHCb has not yet performed any CPT violation study, but has great opportunities to improve existing sensitivities.

This thesis discusses how CPT and Lorentz violation effect neutral mesons and where LHCb can gain the most with respect to CPT and Lorentz violation, in particular for $B^0_{(s)}$ mesons. The data analysis has been performed on the 2011 and 2012 LHCb data set with the decay channel $B^0 \rightarrow J/\psi\, K^0_{\mathrm{s}}$. A likelihood fit and in addition a Fourier analysis, the periodogram, will be presented.

2 Theory

In the coming section, the Standard Model and the mechanisms available for introducing CPT and Lorentz violation will be presented. From this, the effect of these violations on neutral meson systems and in particular on the B^0 system will be discussed. To set the stage, the effect of symmetry transformations $\mathcal{C}$, P, T and combinations of these in the Standard Model will be presented.

2.1 The Standard Model and CPT

Interactions of particles, in particlar sub-atomic particles, are described by the Standard Model. It is a local gauge-invariant quantum field theory with $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{SU}(3)_C$ gauge symmetries. A rotation under $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_\mathrm{L}$ corresponds to an electroweak interaction with quantum numbers weak hypercharge Y and weak ispospin I. For the symmetry group $\mathrm{SU}(3)_C$, a rotation corresponds to a strong interaction and has the quantum number C called colour. The gauge bosons of $\mathrm{SU}(3)_C$ are the 8 gluons. The particle content of the Standard Model consists of coloured particles called quarks and not-coloured particles called leptons. These particles are grouped in three families, consisting each of an up- and down-type quark, charged lepton and a chargeless neutrino. To each doublet of an up- and down-type quark or lepton and neutrino is a quantum number associate called flavour. In addition, the Standard Model contains the scalar Higgs field with a non-zero expectation value of the vacuum, which breaks the $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_\mathrm{L}$ symmetry. The so-called Higgs mechanism mixes the $\mathrm{U}(1)_Y$ and $\mathrm{SU}(2)_\mathrm{L}$ gauge-boson fields to form the massless photon γ of the electromagnetic force and the Z, W^+ and W^- vector bosons of the weak force, which aquire mass due to the symmetry breaking [2]. Interactions with W^+ and W^- transform left-handed up- to down-type fermions and vice versa and have the ability to change flavour. γ and Z couple to both left-handed and right-handed fermions, but do not change flavour. From the Higgs mechanism, fundamental fermions gain mass by the coupling of the Higgs field to the fundamental fermions. The resulting couplings lead to a difference in mass and interaction eigenstates due to the interaction terms between different fermion families. To get the mass eigenstates of these fermions, these couplings have to be diagonalized to form proper mass terms in the Lagrangian. The unitary transformation between the interaction q^I and the mass q eigenstates is given by,

$$\begin{pmatrix} d^I \\ s^I \\ b^I \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1}$$

where the complex-valued matrix is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM} and the up-type quarks are given by $u_i^I = u_i$ [3].

2.1.1 C, P and T operators

Consider a fermionic field, which is a spinor, $\psi(\vec{x}, t, \vec{c})$, with spacial coordinates $\vec{x}$, time coordinate t and a set of additive quantum numbers $\vec{c}$. Being a fermion, it has to obey the

Dirac equation, which has the following Lagrangian,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \tag{2}$$

where γ^μ are the Dirac matrices and m is the mass. The Dirac equation also has solutions with negative energy. These are identified as the corresponding antiparticles and given by $\bar{\psi} = \psi^\dagger \gamma^0$. While keeping the form of the Dirac equation, the effect of charge conjugation C, parity transformation P, time translation T and combinations of these are given:

1. **C.** Charge conjugation translates additive quantum numbers by $\vec{c} \to -\vec{c}$ and for a Dirac spinor,

$$C\psi = i\gamma^2 \psi^*(\vec{x}, t, -\vec{c}),$$

   and for example changes the electric charge of a state.

2. **P.** Parity transformation is the mirroring of space by $\vec{r} \to -\vec{r}$ and for a Dirac spinor,

$$P\psi = \gamma^0 \psi(-\vec{x}, t, \vec{c}),$$

   and conserves spin, but flips helicity.

3. **T.** Time translation reverses time by $t \to -t$,

$$T\psi = i\gamma^1 \gamma^3 \psi^*(\vec{x}, -t, \vec{c}),$$

   and flips momentum and spin, therefore helicity is conserved.

With these transformation, CP and CPT transformations can be derived:

1. **CP.** In terms of Dirac spinors, the CP transformation is given by,

$$CP\psi = i\gamma^2 \gamma^0 \psi^*(-\vec{x}, t, -\vec{c}).$$

2. **CPT.** For a Dirac spinor, the CPT transformation is given by,

$$CPT\psi = \gamma^5 \psi(-\vec{x}, -t, -\vec{c}),$$

   where $\gamma^5 \equiv i\gamma^0 \gamma^1 \gamma^2 \gamma^3$.

From experiment, for example Ref. [4], it is known that the weak interaction violates parity. For example, there are no right-handed neutrinos or left-handed antineutrinos, only left-handed neutrinos and right-handed antineutrinos in the Standard Model. Therefore the weak interaction violates C and P. Any violation of C and P and therefore CP is not observed in electromagnetism, gravity and the strong interaction. Due to the complex conjugate involved in the transformation, CP violation in the Standard Model can, for example, arise due to differences in the interference of two decay amplitudes by,

$$\begin{aligned}
|A|^2 &= ||A_1|e^{i\phi_1}e^{i\theta_1} + |A_2|e^{i\phi_2}e^{i\theta_2}|^2 & |\overline{A}|^2 &= ||A_1|e^{-i\phi_1}e^{i\theta_1} + |A_2|e^{-i\phi_2}e^{i\theta_2}|^2 \\
|A|^2 &= |A_1|^2 + |A_2|^2 & |\overline{A}|^2 &= |A_1|^2 + |A_2|^2 \\
&\quad + 2|A_1||A_2|\cos(\Delta\phi + \Delta\theta) & &\quad + 2|A_1||A_2|\cos(-\Delta\phi + \Delta\theta),
\end{aligned} \tag{3}$$

where ϕ is a *CP*-odd phase and θ is a *CP*-even phase. Clearly for *CP* violation to occur, there has to be a non-zero *CP*-odd and non-zero *CP*-even phase. Since the CKM matrix, Eq. 1, has complex values and therefore can have a non-zero complex phase, a so-called weak phase can arise that is *CP*-odd. Since the CKM matrix is unitary and one is free to assign a phase to a (quark) field, the CKM matrix has to have at least three dimensions to have a physical phase. In addition, for *CP* violation to occur, a *CP*-even term is needed. Since no *CP* violation has been seen in the strong interaction, any phase coming from the strong interaction, for example, gluon exchange in the final state, can contribute a non-zero *CP*-even phase. This phase is called the strong phase.

2.1.2   *CPT* theorem

Due to Lorentz invariance, fermionic fields appear in so-called bilinear forms in Lagrangians, for example the scalar mass term $m\bar{\psi}\psi$ in Eq. 2. In a similar way, interaction terms are sandwiched between the spinors and antispinors of fermions. Since the *CPT* transformation does not involve complex conjugates, in contrast to *CP*, any *CPT* violation in a Lagrangian should come from a *CPT*-odd bilinear. The transformation under *CPT* for bilinears is given in Table. 1.   For a theory to be Lorentz invariant, terms in the Lagrangian should be Lorentz contracted. So, for example, the vector and axial-vector bilinears, that are *CPT*-odd, cannot exist in this form in a Lorentz-invariant Lagrangian. They should be Lorentz contracted by another four-vector and for local gauge-invariant quantum field theories, these terms are contracted by a field that is a vector or an axial-vector field. As an example, take the electromagnetic vector-field A_μ. Under parity transformation, the electric field gains a minus sign, while the magnetic field does not. For time reversal, the magnetic field gains a minus sign and the electric field does not. Charge conjugation gives both a minus sign. Therefore, the scalar electromagnetic potential only changes sign under charge conjugation, but the vector potential changes sign under all transformations. In general, a vector and an axial-vector field change according to Table 2, using four-vector relations $q^i = -q_i$ and $q^0 = q_0$.   It shows that any bilinear in a Lagrangian that is Lorentz contracted by a vector or axial vector is invariant under *CPT*. With the proof in Ref. [5], the *CPT* theorem is derived and states that any local interacting quantum field theory that violates *CPT*, violates Lorentz invariance as well. As an example for the close relation between Lorentz invariance and *CPT* violation, consider a vector bilinear, but

bilinear		*CPT*
scalar	$\bar{\psi}_1\psi_2$	$\bar{\psi}_2\psi_1$
pseudo scalar	$\bar{\psi}_1\gamma_5\psi_2$	$\bar{\psi}_2\gamma_5\psi_1$
vector	$\bar{\psi}_1\gamma_\mu\psi_2$	$-\bar{\psi}_2\gamma_\mu\psi_1$
axial vector	$\bar{\psi}_1\gamma_\mu\gamma_5\psi_2$	$-\bar{\psi}_2\gamma_\mu\gamma_5\psi_1$
tensor	$\bar{\psi}_1\sigma_{\mu\nu}\psi_2$	$\bar{\psi}_2\sigma^{\mu\nu}\psi_1$

Table 1: Transformations of bilinears under *CPT*.

		C	P	T	CPT
vector	V_μ	$-V_\mu^\dagger$	V^μ	$V^{*\mu}$	$-V_\mu^T$
axial	A_μ	$A_\mu^\dagger$	$-A^\mu$	$A^{*\mu}$	$-A_\mu^T$

Table 2: Transformations of four-vectors and axial four-vectors.

now contracted with a real-valued coupling constant with a Lorentz index, a_μ, by,

$$a_\mu \bar{\psi} \gamma^\mu \psi, \tag{4}$$

which conserves the Lorentz-invariant form of the Lagrangian when one performs an observer transfromation, *i.e.* a coordinate-frame transformation, since both a_μ and the fields transform. Charge conjugation does not change the sign of a real-valued coupling, therefore Eq. 4 transforms under CPT as,

$$(CPT)^{-1} a_\mu \bar{\psi} \gamma^\mu \psi (CPT) = a_\mu (CPT)^{-1} \bar{\psi} \gamma^\mu \psi (CPT) = -a_\mu \bar{\psi} \gamma^\mu \psi, \tag{5}$$

and clearly shows that the term is CPT-odd and therefore CPT violating. The Lorentz violation in this term is made clear when one considers a particle in the same frame, but with different boost. Therefore the boosted particle sees a different a_μ and only a_μ transforms, causing the Lorentz violation.

A consequence of CPT invariance in the Standard Model is that particle and antiparticle have the same mass and lifetime, as, for example, can be seen from the mass term, which is of the form of a bilinear $m\bar{\psi}\psi$. As will be shown later on, CPT violation in mixing of B^0 mesons can be parametrized by a mass difference between particle and antiparticle.

To summarize, CP violation in the Standard Model can occur due to the interference of amplitudes with different phases, but CPT violation cannot be introduced in the Standard Model in any Lorentz-invariant way.

2.2 Neutral mesons and CP/CPT violation

For neutral mesons, a particle can transform to an antiparticle and vice versa. This transformation is called neutral meson mixing. Mixing is drived by small differences in the masses and lifetimes of mass eigenstates of neutral mesons. Effects of CPT, CP violation and mixing are deeply intertwined in neutral mesons and their mixing. Therefore, the time-dependent decay rates of mixing with CPT violation is derived first, after which the effect of different types of CP violation is introduced.

2.2.1 Neutral meson mixing with CPT violation

In the Standard Model, neutral meson mixing is possible, *e.g.* due to a specific type of Feynman diagram, called a box diagram. An example is shown in Fig. 1. It shows that under the exchange of W bosons, a particle can transform into its antiparticle. In terms of its interaction eigenstates, the neutral meson clearly is not constant, therefore the mass eigenstates are different from the interaction eigenstates. In the coming section, the effect of mixing to the measureable decay rates of neutral mesons will be derived. The derivation will closely follow Ref. [6].

The wave function Ψ of the neutral meson can be described as a complex linear combination of the interaction eigenstates. In the mesons rest frame, the time evolution of the wave function is governed by the Schrödinger equation,

$$i\partial_t \Psi = \hat{H}\Psi, \tag{6}$$

where $\hat{H}$ is the effective Hamiltonian and the state is labeled as a B meson, without loss of generality. The effective Hamilitonian is governed by a Hermitian mass term M, for the propagation of the meson plus a general Hermitian decay term $-i\Gamma/2$,

$$\hat{H} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22} \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}. \tag{7}$$
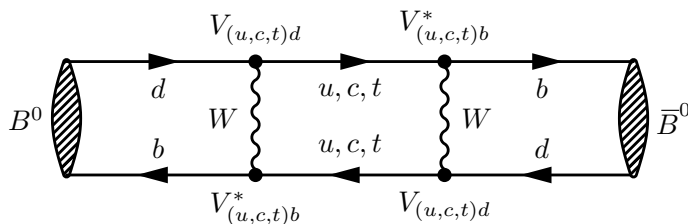

Figure 1: Example of a Feynman diagram for neutral meson mixing in the Standard Model, called a box diagram. Depicted here is the oscillation of a B^0 meson. Note the CKM matrix couplings, which cause CP violation in mixing.

To continue, the following definitions are given,

$$m = \frac{1}{2}(M_{11} + M_{22}), \qquad\qquad \Gamma = \frac{1}{2}(\Gamma_{11} + \Gamma_{22}),$$
$$\delta m = M_{11} - M_{22}, \qquad\qquad \delta\Gamma = \Gamma_{11} - \Gamma_{22}.$$

Diagonalization of Eq. 7 produces two eigenvectors that are mass eigenstates, a heavy $|B_H\rangle$ and a light $|B_L\rangle$ with masses $m_{H,L}$ and decay rates $\Gamma_{H,L}$. The eigenvalues $\lambda_{H,L}$ are given by,

$$\lambda_{H,L} = m - \frac{i}{2}\Gamma \pm \sqrt{\left(M_{12} + \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* + \frac{i}{2}\Gamma_{12}^*\right) + \frac{1}{4}\left(\delta m + \frac{i}{2}\delta\Gamma\right)^2}, \qquad (8)$$

and allow for the definition of,

$$\Delta m = m_H - m_L, \qquad\qquad \Delta\Gamma = \Gamma_H - \Gamma_L,$$
$$\lambda_{H,L} = m_{H,L} - \frac{i}{2}\Gamma_{H,L}, \qquad\qquad \Delta\lambda = \Delta m - \frac{i}{2}\Delta\Gamma.$$

As discussed in section 2.1, any difference in the mass or lifetime of particle and antiparticle would mean CPT violation. For similar reasons, a difference in total decay rate would also mean CPT violation. Therefore, to parametrize CPT violation in mixing, it is sufficient to introduce a parameter for this difference,

$$z = \frac{\delta m - \frac{i}{2}\delta\Gamma}{\Delta m - \frac{i}{2}\Delta\Gamma}, \qquad (9)$$

where the denominator is the eigenvalue difference $\Delta\lambda$, making z dimensionless. By solving for the eigenvectors, *i.e.* mass eigenstates, and with the definition of z, the mass eigenstates are given by,

$$|B_L\rangle = p\sqrt{1-z}|B^0\rangle + q\sqrt{1+z}|\overline{B}^0\rangle,$$
$$|B_H\rangle = p\sqrt{1+z}|B^0\rangle - q\sqrt{1-z}|\overline{B}^0\rangle, \qquad (10)$$

where p and q are given by,

$$\frac{q}{p} = -\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}, \qquad (11)$$

and describe mixing without CPT violation. The time-dependent mass eigenstates are given by solving the Schrödinger equation,

$$|B_L(t)\rangle = e^{-i\lambda_L t}|B_L(0)\rangle,$$
$$|B_H(t)\rangle = e^{-i\lambda_H t}|B_H(0)\rangle. \qquad (12)$$

At the moment of the production of the neutral meson, it is in an interaction eigenstate, so either a $|B^0\rangle$ or $|\overline{B}^0\rangle$. By writing the interaction states in terms of mass eigenstates

and inserting the time dependence, the time evolution of the neutral meson states are obtained,

$$\begin{aligned} |B^0(t)\rangle &= [g_+(t) + zg_-(t)] |B^0\rangle - \sqrt{1-z^2}\frac{q}{p}g_-(t)|\overline{B}^0\rangle, \\ |\overline{B}^0(t)\rangle &= [g_+(t) - zg_-(t)] |B^0\rangle - \sqrt{1-z^2}\frac{p}{q}g_-(t)|\overline{B}^0\rangle, \end{aligned} \tag{13}$$

where $g_\pm(t)$ are given by,

$$g_\pm(t) = \frac{1}{2}\left[e^{-i\lambda_H t} \pm e^{-i\lambda_H t} \right]. \tag{14}$$

With Eq. 13, the amplitude for a decay to final state f at decay time t can be written as,

$$\langle f|T|B^0(t)\rangle = a_+ g_+(t) + a_- g_-(t), \tag{15}$$

where $a_\pm$ are complex coefficients that are determined by the specific decay type. The decay rate is given by,

$$\Gamma_{B^0 \to f}(t) = |\langle f|T|B^0(t)\rangle|^2, \tag{16}$$

and it shows that $g_\pm(t)$ functions appear in products of each other in the decay rate as,

$$\begin{aligned} |g_\pm|^2 &= \frac{1}{2}e^{-\Gamma t}\left[\sinh(\Delta\Gamma t/2) \pm \cos(\Delta mt) \right], \\ g_+^* g_- &= -\frac{1}{2}e^{-\Gamma t}\left[\cosh(\Delta\Gamma t/2) + i\sin(\Delta mt) \right]. \end{aligned} \tag{17}$$

The general decay rate equations are now given by,

$$\begin{aligned} \Gamma_{B^0 \to f}(t) = e^{-\Gamma t}\Bigg[&\frac{1}{2}(|a_+|^2 + |a_-|^2)\cosh(\Delta\Gamma t/2) + \frac{1}{2}(|a_+|^2 - |a_-|^2)\cos(\Delta mt) \\ &- \mathcal{R}e(a_+^* a_-)\sinh(\Delta\Gamma t/2) + \mathcal{I}m(a_+^* a_-)\sin(\Delta mt) \Bigg]. \end{aligned} \tag{18}$$

To proceed to calculating more specific decay rates, two types of decay have to be considered: decay to flavour-specific final states and to CP eigenstates.

2.2.2 Decay to a flavour eigenstate

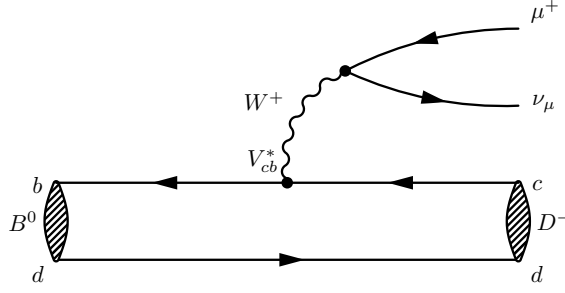Take for example the semileptonic decay $B^0 \to D^- \mu^+ \nu_\mu$ pictured in Fig. 2.

Figure 2: Feynman diagram for semileptonic decay $B^0 \to D^- \mu^+ \nu_\mu$.

It clearly is flavour, *i.e.* interaction state, specific. In case of a B^0, one would exchange a W^+ and in case of a $\overline{B}^0$ a W^-. Therefore, $B^0 \to D^- \mu^+ \nu_\mu$ and $\overline{B}^0 \to D^+ \mu^- \bar{\nu}_\mu$. This allows us to set the following decay amplitudes,

$$\langle f|T|B^0\rangle = A_f \qquad\qquad \langle \overline{f}|T|\overline{B}^0\rangle = \overline{A}_{\overline{f}}$$
$$\langle \overline{f}|T|B^0\rangle = A_{\overline{f}} = 0 \qquad\qquad \langle f|T|\overline{B}^0\rangle = \overline{A}_f = 0, \tag{19}$$

and with Eq. 13 and 19, the values for $a_\pm$ are given in Table. 3. Since *CPT* violation is expected to be small, only terms to first order in z are considered and with Eq. 18 and Table. 3, the decay rates are given by,

$$\begin{aligned}
\Gamma_{B^0 \to f}(t) = |A_f|^2 e^{-\Gamma t}\Big[&\frac{1}{2}\cosh(\Delta\Gamma t/2) + \frac{1}{2}\cos(\Delta m t) \\
&- \mathcal{R}e(z)\sinh(\Delta\Gamma t/2) + \mathcal{I}m(z)\sin(\Delta m t)\Big] \\[4pt]
\Gamma_{\overline{B}^0 \to f}(t) = |A_f|^2 e^{-\Gamma t}\left|\frac{p}{q}\right|^2 &\left[\frac{1}{2}\cosh(\Delta\Gamma t/2) - \frac{1}{2}\cos(\Delta m t)\right] \\[4pt]
\Gamma_{B^0 \to \overline{f}}(t) = |\overline{A}_{\overline{f}}|^2 e^{-\Gamma t}\left|\frac{q}{p}\right|^2 &\left[\frac{1}{2}\cosh(\Delta\Gamma t/2) - \frac{1}{2}\cos(\Delta m t)\right] \\[4pt]
\Gamma_{\overline{B}^0 \to \overline{f}}(t) = |\overline{A}_{\overline{f}}|^2 e^{-\Gamma t}\Big[&\frac{1}{2}\cosh(\Delta\Gamma t/2) + \frac{1}{2}\cos(\Delta m t) \\
&+ \mathcal{R}e(z)\sinh(\Delta\Gamma t/2) - \mathcal{I}m(z)\sin(\Delta m t)\Big]
\end{aligned} \tag{20}$$

	$B^0(t) \to f$	$\overline{B}^0(t) \to f$	$B^0(t) \to \overline{f}$	$\overline{B}^0(t) \to \overline{f}$
a_+	A_f	0	0	$\overline{A}_{\overline{f}}$
a_-	zA_f	$-\sqrt{1-z^2}\frac{p}{q}A_f$	$-\sqrt{1-z^2}\frac{q}{p}\overline{A}_{\overline{f}}$	$z\overline{A}_{\overline{f}}$

Table 3: Values for $a_\pm$ for flavour specific decays.

2.2.3 Decay to a CP eigenstate

In case a decay is not flavour specific, the mixing will start to interfere with the decay. As will turn out, the following parameter is useful in describing this interference effect,

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{21}$$

Compared to the flavour specific case, the amplitudes $A_{\bar{f}}$ and $\overline{A}_f$ are not zero any more. For this case, the values for $a_\pm$ are given in Table. 4. To calculate the decay rates, z contributions of first order are only considered and to simplify the equations, the following definitions are used,

$$C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \qquad S_f = \frac{2\mathcal{I}m(\lambda_f)}{1 + |\lambda_f|^2}, \qquad D_f = \frac{2\mathcal{R}e(\lambda_f)}{1 + |\lambda_f|^2}. \tag{22}$$

With these definitions, the decay rates are given by,

$$\Gamma_{B^0 \to f_{CP}}(t) \propto e^{-\Gamma t}\Bigg[[1 - D_f\mathcal{R}e(z) - S_f\mathcal{I}m(z)]\cosh(\Delta\Gamma t/2) + [D_f - \mathcal{R}e(z)]\sinh(\Delta\Gamma t/2)$$

$$+ [C_f + D_f\mathcal{R}e(z) + S_f\mathcal{I}m(z)]\cos(\Delta m t) - [S_f - \mathcal{I}m(z)]\sin(\Delta m t)\Bigg],$$

$$\Gamma_{\overline{B}^0 \to f_{CP}}(t) \propto e^{-\Gamma t}\left|\frac{p}{q}\right|^2\Bigg[[1 + D_f\mathcal{R}e(z) - S_f\mathcal{I}m(z)]\cosh(\Delta\Gamma t/2) + [D_f + \mathcal{R}e(z)]\sinh(\Delta\Gamma t/2)$$

$$- [C_f + D_f\mathcal{R}e(z) - S_f\mathcal{I}m(z)]\cos(\Delta m t) + [S_f - \mathcal{I}m(z)]\sin(\Delta m t)\Bigg],$$

$$\tag{23}$$

where the proportionality factor is $|A_f|(1 + |\lambda_f|^2)/2$. In case of no CPT violation, $i.e.$ $z = 0$, the parameters of C_f, D_f and S_f retain their role as coefficients for individual time-dependent functions in the decay rates.

	$B^0(t) \to f$	$\overline{B}^0(t) \to f$
a_+	A_f	$-\overline{A}_f$
a_-	$A_f(z - \lambda_f)$	$\frac{p}{q}A_f(1 + z\lambda_f)$

Table 4: Values for $a_\pm$ for decays to CP eigenstates. Only first order z contributions are considered.

2.2.4 Classification and measurement of CP/CPT violation

As the behaviour of neutral mesons including mixing and decay has now been described, the effect of CP and CPT violation can be properly studied. These violation effects can be described by measurable asymmetries. Since decay rates are expressed in terms of flavour at production and final state at decay time t, two types of decay-time dependent asymmetries are considered. The untagged asymmetry, where the initial flavour is not known, is given by,

$$A_{\text{untagged}}(t) \equiv \frac{\Gamma_f - \Gamma_{\overline{f}}}{\Gamma_f + \Gamma_{\overline{f}}} = \frac{\Gamma_{Bz \to f} + \Gamma_{\overline{B}^0 \to f} - \Gamma_{B^0 \to \overline{f}} - \Gamma_{\overline{B}^0 \to \overline{f}}}{\Gamma_{Bz \to f} + \Gamma_{\overline{B}^0 \to f} + \Gamma_{B^0 \to \overline{f}} + \Gamma_{\overline{B}^0 \to \overline{f}}}, \tag{24}$$

and is used with flavour-specific decays. It is also possible to tag a neutral meson, $i.e.$ determine its initial flavour. For a decay to a CP eigenstate, the tagged asymmetry is given by,

$$A_{CP/CPT}(t) \equiv \frac{\Gamma_{\overline{B}^0 \to f} - \Gamma_{B^0 \to f}}{\Gamma_{\overline{B}^0 \to f} + \Gamma_{B^0 \to f}}, \tag{25}$$

Three categories of CP/CPT violation effects are considered:

1. **Decay.** With Eq. 3, it was made clear that two or more interfering decay amplitudes can cause a difference in the size of an amplitude and its CP conjugated amplitude. This difference is independent of mixing and therefore it is also possible for charged hadrons. In general, this difference translates into the following definition of the direct asymmetry,

$$\mathcal{A}^{\text{dir}} = \frac{|\overline{A}|^2 - |A|^2}{|\overline{A}|^2 + |A|^2}. \tag{26}$$

   If one considers the possibility of direct CPT violation, this is hard to disentangle from direct CP violation, since both effect the size of the amplitude.

2. **Mixing.** For mixing, an asymmetry can occur through the relative complex phase of Γ_{12} and M_{12} of Eq. 7 and is parametrized by,

$$\mathcal{A}^{\text{mix}} = \frac{1 - |q/p|^2}{1 + |q/p|^2} \approx |p/q|^2 - 1, \tag{27}$$

   and since predictions from the Standard Model estimate $\mathcal{A}^{\text{mix}}$ at $\mathcal{O}(10^{-4})$ for B^0 [7], $|p/q|$ is very near unity. Experimental precision has not yet reached this [8]. For other neutral mesons, $\mathcal{A}^{\text{mix}}$ is low as well and for kaons it has been measured at $\mathcal{O}(10^{-3})$ [9]. The experimental determination is usually done by determining the untagged or tagged asymmetry for semileptonic decays, since they have a high branching fraction and are flavour specific.

3. **Decay with mixing.** As shown in section 2.2.3, the behaviour of decays to CP eigenstates are governed by the parameter λ_f. Assuming low direct CP violation, $i.e.$

$C_f \approx 0$, $|\lambda_f|$ is of order unity, but the phase in λ_f has great effect, due to the fact that the sin term in the decay rates factor with $\mathcal{I}m\lambda_f$. The phase of λ_f is the combined phase of the phases of the mixing and decay Feynman diagrams. Experimentally, λ_f is determined using the tagged asymmetry $A_{CP/CPT}$. From Eq. 23, CPT violation occurs in most of the same terms as CP violation. Only the cosh term adds a unique contribution.

Next to these theoretical CP/CPT violation effects, difference in detection efficiencies can cause asymmetries. Take for example flavour specific decays. Due to different efficiencies for the detection of f and $\overline{f}$, a fake direct CP violation can arise. This detection asymmetry is defined,

$$A_D = \frac{\epsilon_{\overline{f}} - \epsilon_f}{\epsilon_{\overline{f}} + \epsilon_f}. \tag{28}$$

In addition, at the LHC, $b\overline{b}$ quark pairs are produced in a hadronic environment, where protons are collided with protons, as depicted in the diagram from Fig. 3. Quarks produced in such an evironment can, for example, hadronize with the remaining valence quarks of the proton and since protons do not contain valence antiquarks, there will be a bias toward hadronizing in baryons with respect to antibaryons. This will reduce the amount of b quarks hadronizing in a $\overline{B}^0$ compared to $\overline{b}$ quarks hadronizing in a B^0. The effect encompassing these biases is called production asymmetry and is parametrized by,

$$A_{\mathrm{P}} = \frac{N_{\overline{B}^0} - N_{B^0}}{N_{\overline{B}^0} + N_{B^0}}, \tag{29}$$

where N_{B^0} and $N_{\overline{B}^0}$ are the number of B^0 and $\overline{B}^0$ mesons at production. These detection effects are taken into consideration in this analysis and will be further discussed in section 4.

Having set a classification of CP/CPT violation, specific neutral mesons have to be studied to determine where CPT violation has the biggest effect. But first, a beyond the Standard Model framework that is able to introduce CPT and Lorentz violation will be discussed.
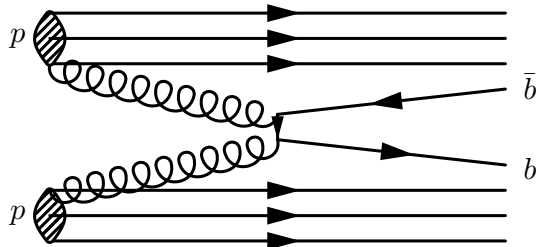

Figure 3: Example of a diagram for $b\overline{b}$ production called gluon-gluon fusion.

2.3 Effective field theory: Standard Model Extension

Despite its big succes in describing interactions at the electroweak scale, the Standard Model has still a number of problems. For example, it cannot describe gravity. Being the weakest force of the four fundamental forces, gravity becomes significant at a much larger energy scale, the Planck scale, roughly $M_{\mathrm{Pl}} \approx 10^{19}\,\mathrm{GeV}$. At this scale, a more fundamental theory is needed for which the Standard Model would be a low-energy approximation. The most promising theory at this time that is able to do this, is string theory. Since string theory has particles that are extended objects called strings, the *CPT* theorem does not apply anymore. In string theory, there can be *CPT* and/or Lorentz violating operators [10]. Therefore, if measured, *CPT* and Lorentz violation would offer an unambiguous signature of physics at the Planck scale. In order to parametrize the effect of possible *CPT* and Lorentz violating operators in the Standard Model, the Standard Model Extension has been developed [11].

The Standard Model Extension is an effective field theory, therefore it is a framework and not a theory and this makes it (string) model independent. Operators in string theory can have non-zero expectation values at low energy due to spontaneous symmetry breaking. Possible forms of operators have been categorized and as expectation values put into the Standard Model Lagrangian. The Standard Model Extension incorporates many such terms while keeping properties of the Standard Model like renormalizability and gauge structure.

2.3.1 Standard Model Extension for neutral mesons

In case of quarks, there appears to be only one *CPT*-odd term in the Standard Model Extension [12]. This is the zeroth component of the before mentioned bilinear from Eq. 4,

$$a_0^q \bar{\psi} \gamma^0 \psi, \tag{30}$$

where q denotes the valence-quark flavour, added since there is no a priori reason for equal expectation values for quarks. Since a_0 is real, $\delta\Gamma = 0$. Eq. 30 is quark-flavour diagonal, therefore it effects the flavour diagonal terms in the effective Hamiltonian of neutral meson mixing, Eq. 7. Ignoring small effects of quark binding and normalization, the diagonal mass difference δm is given by,

$$\delta m \equiv \Delta a_0 \simeq a_0^{q_1} - a_0^{q_2}. \tag{31}$$

Since Eq. 30 and 31 hold for the rest frame of the meson, transforming to the lab frame is trivial, due to the Lorentz invariance under coordinate frame transformations. Another particle in the same frame, related to the other particle by a boost $\beta^\mu = \gamma(1, \vec{\beta})$, has a coupling of $\beta^\mu \Delta a_\mu$. This boost dependence clearly shows the Lorentz violating nature of *CPT* violation. The dimensionless z parameter can now be written as,

$$z = \frac{\beta^\mu \Delta a_\mu}{\Delta m - i\Delta\Gamma/2}. \tag{32}$$

Since Δa_μ is real, the following relation for z holds,

$$\mathcal{R}e(z)\Delta\Gamma = 2\mathcal{I}m(z)\Delta m, \qquad (33)$$

and is called the Standard Model Extension constraint. Due to Eq. 31, the following approximate relation between Δa_μ of different neutral mesons holds,

$$\Delta a_\mu^K - \Delta a_\mu^{B^0} + \Delta a_\mu^{B_s^0} \approx 0. \qquad (34)$$

The boost is the biggest advantage of LHCb with respect to B-factories. Most B-factories have $\gamma\beta$ factors around 0.5 [13]. In case of LHCb, B mesons are created at a much higher energy and the B mesons have a boost of about $\langle\gamma\beta\rangle \approx 20$. The meson direction dependence shows that in the analysis to determine Δa_μ, the frame of LHCb is needed in a inertial frame, *i.e.* a frame with respect to fixed stars, called a sidereal frame. As LHCb rotates along with the earth, the mesons, which are collimated along the beam of the accelerator, will precess around the rotational axis of the earth and therefore $\beta^\mu\Delta a_\mu$ can be written as,

$$\beta^\mu\Delta a_\mu = \gamma\left[\Delta a_0 + \beta\cos(\chi)\Delta a_Z + \beta\sin(\chi)\left[\Delta a_Y\sin(\Omega\hat{t}) + \Delta a_X\cos(\Omega\hat{t})\right]\right], \qquad (35)$$

where $\beta = |\vec{\beta}|$, $\Delta a^{X,Y,Z} = -\Delta a_{X,Y,Z}$ and χ is an effective angle of the mesons with respect to the rotational axis of the earth. It shows that z has a time-dependent part and a constant part that are both boosted by γ. The time dependence is the result of the rotation of the earth with frequency Ω. The determination of χ, Ω and the timekeeping will be extensively discussed and derived in section 3.2.

With this and previous sections, a general description of *CP* and in particular *CPT* violation has been presented. Next, the $B_{(s)}^0$ meson systems and the specific decay channel, analyzed in this thesis, will be discussed, in order to show where most sensitivity can be gained.

2.4 Comparison of neutral mesons systems

The neutral meson system that has been most intensively studied is the kaon system. This includes CPT violation. $\Delta a_\mu^{K^0}$ has been determined to be consistent with zero at a precision of $\mathcal{O}(10^{-18}\,\text{GeV})$ [14]. Considerations with respect to CPT and Lorentz violation measurements at LHCb have been discussed in Ref. [15][1] and show that LHCb is not able to improve CPT measurements with kaons. In case of the D^0 meson, LHCb is able to improve measurements, but the focus of this thesis is on $B^0_{(s)}$ mesons.

2.4.1 The B^0 and B^0_s systems

From the decay rate equations, Eq. 20 and 23, it shows that Δm and $\Delta\Gamma$ play an important role. The experimental values of these parameters are given for B^0 and B^0_s in Table 5. Since in the definition of z the denominator is $\Delta m + i\Delta\Gamma/2$, it is favourable to have low Δm and/or $\Delta\Gamma$. This means that B^0 is more sensitive to CPT violation than B^0_s with a factor of around 30, but the measurement of B^0_s can be used to test Eq. 34. In the decay rate equations, z comes in the form of $\mathcal{R}e(z)$ and $\mathcal{I}m(z)$. Due to the values from Table 5 and the Standard Model prediction for $\Delta\Gamma = -0.0027 \pm 0.0007\,\text{ps}^{-1}$ for B^0 [16], the Standard Model Extension constraint from Eq. 33 results in the relations for $\mathcal{R}e(z)$ and $\mathcal{I}m(z)$ listed in Table 6. It makes clear that, with a small z, $\mathcal{I}m(z)$ can be ignored when determining Standard Model Extension parameters. In case of flavour-specific decays, $\mathcal{R}e(z)$ appears in the decay rates multiplied with $\sinh(\Delta\Gamma t/2) \approx \Delta\Gamma t/2$. Due to the low value of $\Delta\Gamma$ for B^0 mesons, $\Delta\Gamma_d$, flavour-specific decays become unfavourable, despite that they are more abundant. Take for example the measurement of $\mathcal{A}^{\text{mix}}$ for B^0 at LHCb [8], where uncertainties on the untagged asymmetry are around 0.2%, resulting in uncertainties on $\mathcal{R}e(z)$ of 7% with the Standard Model prediction of $\Delta\Gamma_d$. In case of a CP eigenstate, where decay rate equations from Eq. 20 hold, uncertainties on the asymmetry are of the order of $(5-10)\%$ at LHCb [17] and with the full data set will go down to a few percent and do not need to use Standard Model predictions. Therefore, the neutral meson chosen for this analysis is B^0 in the decay channel $B^0 \to J/\psi\, K^0_\text{S}$, which is the most abundant final state that is a CP eigenstate.

2.4.2 $B^0 \to J/\psi\, K^0_\text{S}$

The J/ψ resonance consist of a $\bar{c}c$-pair and has spin-1. K^0_S is the short-lived neutral kaon mass eigenstate, $|K^0_\text{S}\rangle = p|K\rangle + q|\overline{K}\rangle$, and has spin-0. Clearly, if one performs the CP

---

[1]Authors are the supervisor and author of this thesis.

	$\tau[\,\text{ps}]$	$\Delta m[\,\text{ps}^{-1}]$	$\Delta\Gamma[\,\text{ps}^{-1}]$
B^0	1.519 ± 0.005	0.510 ± 0.03	0.002 ± 0.015
B^0_s	1.512 ± 0.007	17.761 ± 0.022	0.091 ± 0.008

Table 5: Experimental values for mixing parameters for B^0 and B^0_s [9].

	$\lvert \mathcal{R}e(z)/\mathcal{I}m(z) \rvert = \left\lvert \frac{2\Delta m}{\Delta \Gamma} \right\rvert$
B^0	380
B_s^0	390

Table 6: Relative approximate sizes of $\mathcal{R}e(z)$ and $\mathcal{I}m(z)$ for B^0 and B_s^0.

(a) Leading order diagram for $B^0 \to J/\psi\, K_{\mathrm{S}}^0$.

(b) Penguin diagram contributing to direct CP violation in $B^0 \to J/\psi\, K_{\mathrm{S}}^0$.
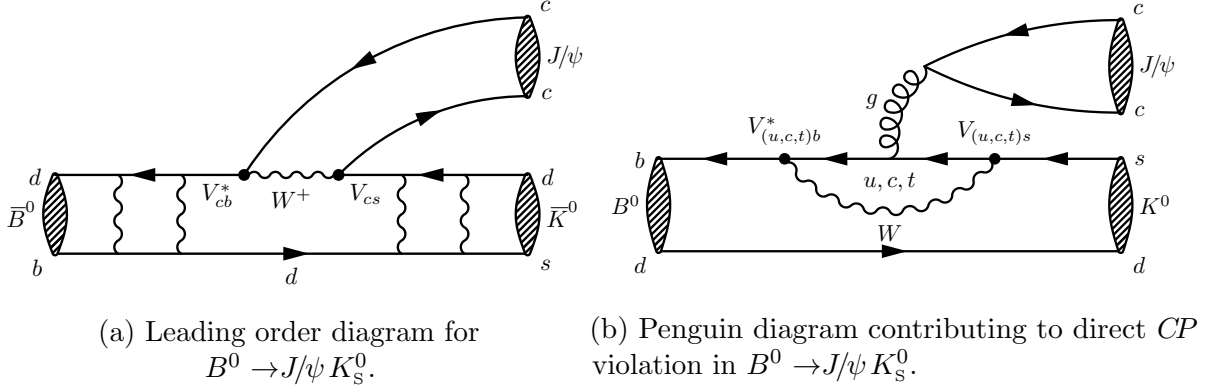
Figure 4: Feynman diagrams contributing to $B^0 \to J/\psi\, K_{\mathrm{S}}^0$.

operation, the J/ψ is CP-even and since for kaons $\lvert p/q \rvert \approx 1$, K_{S}^0 is (almost) CP-even. Since the B^0 is spin-0, J/ψ and K_{S}^0 have relative orbital angular momentum $l = 1$. Since the angular momentum part of the wave function gets a $(-1)^l$ phase under charge conjugation, $J/\psi\, K_{\mathrm{S}}^0$ changes sign under charge conjugation and therefore, $J/\psi\, K_{\mathrm{S}}^0$ is CP-odd.

With the leading order Feynman diagram, Fig. 4a, and the CKM matrix, Eq. 1, $\lambda_{J/\psi\, K_{\mathrm{S}}^0}$ can be determined. Important to note is that both the mixing of the B^0 and the K^0 has to be accounted for. Since the decay and the CP conjugated decay only differ in terms of CKM elements, $\lambda_{J/\psi\, K_{\mathrm{S}}^0}$ is given by,

$$\lambda_{J/\psi\, K_{\mathrm{S}}^0} = - \left(\frac{q}{p} \right)_{B^0} \left(\frac{\overline{A}_{J/\psi\, K_{\mathrm{S}}^0}}{A_{J/\psi\, K_{\mathrm{S}}^0}} \right) \left(\frac{p}{q} \right)_{K^0} = - \left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \right) \left(\frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*} \right) \left(\frac{V_{cd}^* V_{cd}}{V_{cs} V_{cd}^*} \right), \tag{36}$$

where the minus sign comes from the fact that the final state is CP-odd. To be able to write $\lambda_{J/\psi\, K_{\mathrm{S}}^0}$ in terms of CKM phases, the Wolfenstein parameterization is used,

$$V_{CKM} = \begin{pmatrix} \lvert V_{ud} \rvert & \lvert V_{us} \rvert & \lvert V_{ub} \rvert e^{-i\gamma} \\ -\lvert V_{cd} \rvert & \lvert V_{cs} \rvert & \lvert V_{cb} \rvert \\ \lvert V_{td} \rvert e^{-i\beta} & -\lvert V_{ts} \rvert e^{-i\beta_s} & \lvert V_{tb} \rvert \end{pmatrix} + \mathcal{O}(\lambda^{-5}), \tag{37}$$

where the Wolfenstein parameter $\lambda \approx 0.23$ [18]. This results in the following Standard Model prediction for the first order tree contribution,

$$\lambda_{J/\psi\, K_{\mathrm{S}}^0} = -e^{-2i\beta}, \tag{38}$$

16

and with the Standard Model fit from the Heavy Flavour Averaging Group (HFAG) [19] for $\beta = (21.5^{+0.8}_{-0.7})°$ results in,

$$S_{J/\psi K^0_{\rm S}} = \sin(2\beta) = 0.682^{+0.020}_{-0.018}, \quad D_{J/\psi K^0_{\rm S}} = -\cos(2\beta) = -0.731^{+0.019}_{-0.016}, \quad C_{J/\psi K^0_{\rm S}} = 0. \quad (39)$$

As shown in section 2.1, for direct CP violation, multiple amplitudes with different weak phases have to occur. For example, Fig. 4b adds to the amplitude of $B^0 \to J/\psi K^0_{\rm S}$ and carries different weak and strong phases and therefore direct CP violation can occur. These so-called penguin contributions are expected to be of the order of $C_{J/\psi K^0_{\rm S}} \approx 1\%$ [20].

With the low value for $\Delta\Gamma_d$, $\mathcal{A}^{\rm mix}$ and $\mathcal{I}mz$, the decay rate equations can be approximated to,

$$\Gamma_{B^0 \to J/\psi K^0_{\rm S}}(t|d) \propto e^{-\Gamma t}\bigg[1 - dD_f \mathcal{R}e(z) + d[C_f + D_f \mathcal{R}e(z)]\cos(\Delta mt)$$
$$- dS_f \sin(\Delta mt)\bigg], \quad (40)$$

and leads to the following asymmetry,

$$A_{CPT/CP}(t) = D_f \mathcal{R}e(z) - [C_f + D_f \mathcal{R}e(z)]\cos(\Delta mt) + S_f \sin(\Delta mt), \quad (41)$$

where d is the tag of the B^0 meson. For a B^0 tag, $d = 1$ and for a $\overline{B}^0$ tag, $d = -1$.

2.5 Conclusion

CPT and Lorentz symmetry are exact in the Standard Model. Any violation suggests low-energy effects of Planck scale physics. At LHCb, there is a promising opportunity to test these symmetries in neutral meson mixing. Particularly decays to *CP* eigenstates, where these violating effects are not suppressed by the low decay rate difference $\Delta\Gamma$ of B^0, are a good opportunity. Due to the high boost compared to e^+e^- B-factories, LHCb is able to measure Standard Model Extension parameter Δa_μ to high precision. The chosen final state is $B^0 \to J/\psi K^0_{\rm S}$ and requires a decay-time dependent analysis. In the coming sections, the detector and the coordinate frames used to do the Lorentz-violation analysis will be described. After this, the decay-time dependent analysis, using a likelihood fit, and in addition a Fourier analysis, called the periodogram, will be presented.

3 Detector

The data used for this analysis are collected at the LHCb detector, where protons, supplied by the Large Hadron Collider (LHC), are collided at $\sqrt{s} = 7\,\text{TeV}$ in 2011 and $\sqrt{s} = 8\,\text{TeV}$ in 2012. The total data set corresponds to an integrated luminosity of $3\,\text{fb}^{-1}$. The LHC is a two-ring proton synchrotron and collider [21]. In section 3.1 the LHCb detector will be described and in section 3.2 the coordinate frames used in the Lorentz-violating analysis will be presented.

3.1 LHCb

The LHCb detector [22] is one of the four big experiments at the LHC along with ATLAS, CMS and ALICE and its purpose is to study bottom and charm physics, in particular *CP* violation in *B* decays and rare decays. Due to the shape of gluon-momentum distributions in protons, *b*-hadrons have a high cross-section in the forward and backward regions. Therefore, LHCb is build as a forward spectrometer with a pseudorapidity acceptance between $2 < \eta < 5$. The schematical overview of the detector is given in Fig. 5. The tracking system constists of a silicon-strip Vertex Locator (VELO) around the interaction region, a silicon-strip tracker (TT) before the magnet and after the magnet, three tracking stations (T1-T3), consisting of a straw-tube Outer Tracker (OT) and a silicon-strip Inner Tracker (IT) around the beam pipe. The combination of the dipole magnet with a bending power of $4\,\text{Tm}$ and the tracking system results in relative uncertainties on momenta of 0.4% at $5\,\text{GeV}$ to 0.6% at $100\,\text{GeV}$. The particle identification (PID) system contains two ring-imaging Cherenkov detectors (RICH1 and RICH2), a scintillating-pad detector
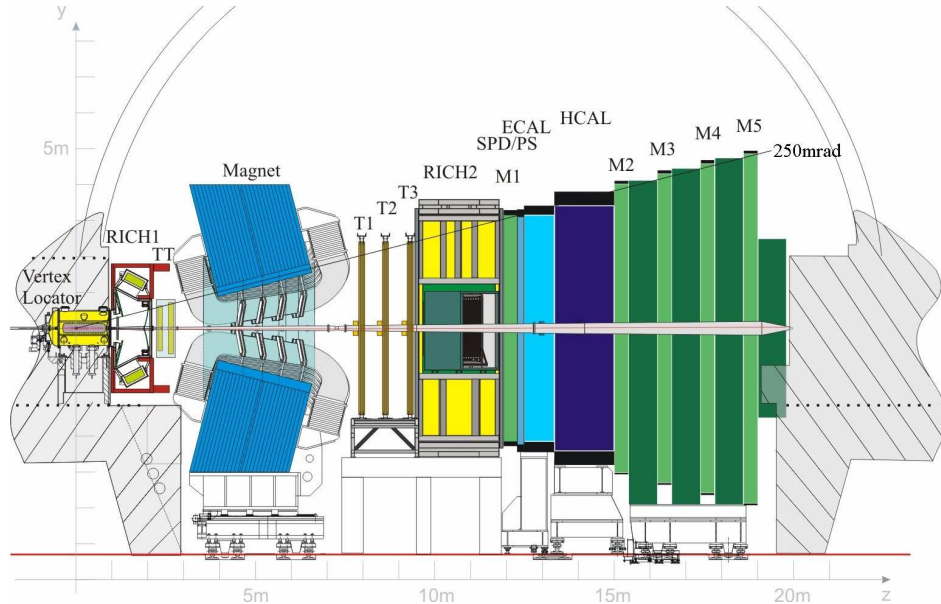

Figure 5: LHCb detector [22].

18

(SPD) and pre-shower detector (PS), an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL) and a muon system (M1-M5), consisting of alternating layers of multiwire proportional chambers and iron. The PID system is designed to distinguish charged hadrons with the RICH1 and RICH2, identify electron, photon and hadron candidates with the ECAL and HCAL and muons with the muon system. Triggering is done in two stages: a fast hardware stage called Level-0 (L0) and a sofware stage, using full event reconstruction, called the High Level Trigger (HLT). The L0 stage uses only information of the calorimeters and the muon system. Candidate selection for this analysis will be discussed in section 4.1.

3.2 LHCb in the sun-centered frame

To perform the sidereal analysis, the direction of the B mesons is needed with respect to a non-rotating frame, *i.e.* a frame constant with respect to fixed stars. The appropriate choice is the sun-centered frame, as is used in the definition of Δa_μ [23]. First, the lab frame of LHCb is introduced, after which the definition of the sun-centered frame and the rotation from the lab frame to the sun-centered frame is given.

3.2.1 LHCb lab frame

The LHCb lab frame is given by the unit vectors $\{\hat{x}, \hat{y}, \hat{z}\}$, where the $\hat{z}$-axis points along the clockwise beam of the LHC. The $\hat{x}$-axis lies in the geodetic horizontal plane and points out of the LHC ring. The plane spanned by the $\hat{x}$ and $\hat{z}$ axes is inclined with respect to the geodetic plane with an angle α, rotated over the $\hat{x}$-axis. The $\hat{y}$-axis points away from the earth, perpendicular to the inclined $\hat{x}\hat{z}$-plane and together, the set $\{\hat{x}, \hat{y}, \hat{z}\}$ forms the right-handed LHCb lab frame, depicted in Fig. 6 [24]. The geodetic position (longitude l and latitude λ) of interaction point 8 (IP8) in the World Geodetic System 84 (WGS84)
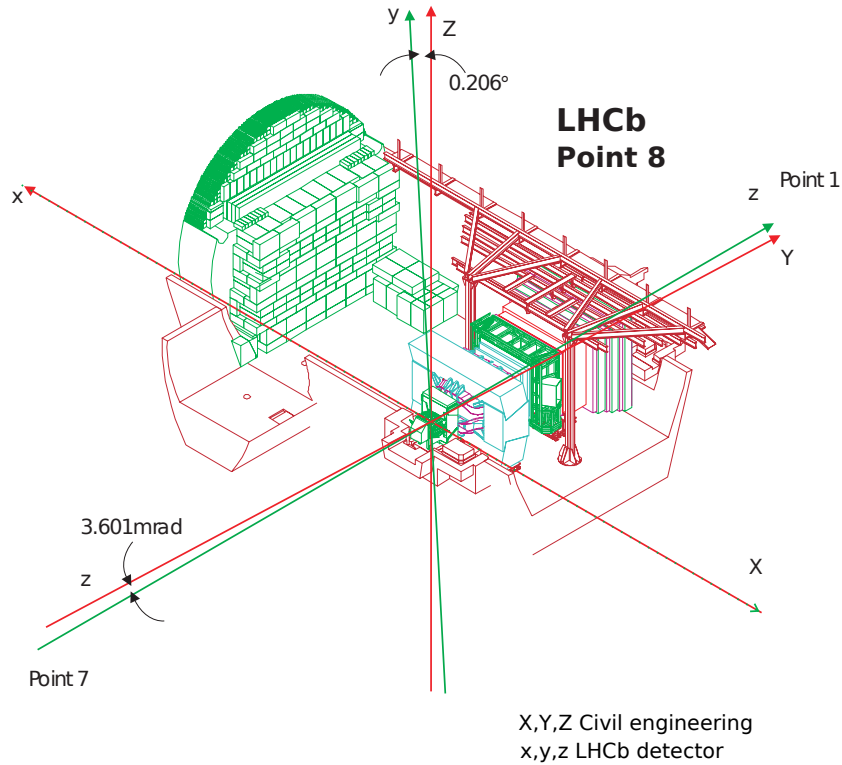

Figure 6: LHCb coordinate system [24]. The LHCb detector coordinates define the lab frame. In the background, the cavern and the detector are depicted. The coordinate system used by the civil engineering is also shown. The difference between the two is the tilt α.

Table 7: Positional angles of LHCb [25]. Geodetic position in the WGS84 system. Note the dominating uncertainty in the azimuth θ.

	Angle		[°]
longitude	λ	46.2414	$\pm$ 0.0001
azimuth	θ	236.296	$\pm$ 0.003
latitude	l	6.0963	$\pm$ 0.0001
tilt	α	0.20632	$\pm$ 0.00003

has been reported in Ref. [25], with an accuracy of about 1.5 m. The collision point of LHCb is located rougly 13 m from IP8, but this difference is negligible in terms of global coordinates and is translated into the error reported in Table 7 for the geodetic position of LHCb. The azimuth θ, defined by the angle of the beam east of north, is determined using the geodetic coordinates of LHCb and an additional point at the end of the long straight-section in the LHC ring [25]. The azimuth has the biggest contributing error to the direction of the beam of about 10 arc seconds.

3.2.2 Sun-centered frame

The sun-centered frame (SCF), depicted in Fig. 7, is defined by the unit vectors $\{\hat{X}, \hat{Y}, \hat{Z}\}$, where $\hat{Z}$ points along the rotational axis of the earth and $\hat{X}$ points from the sun to the vernal equinox. $\hat{Y}$ complements $\hat{X}$ and $\hat{Z}$ to form a right-handed coordinate system. The
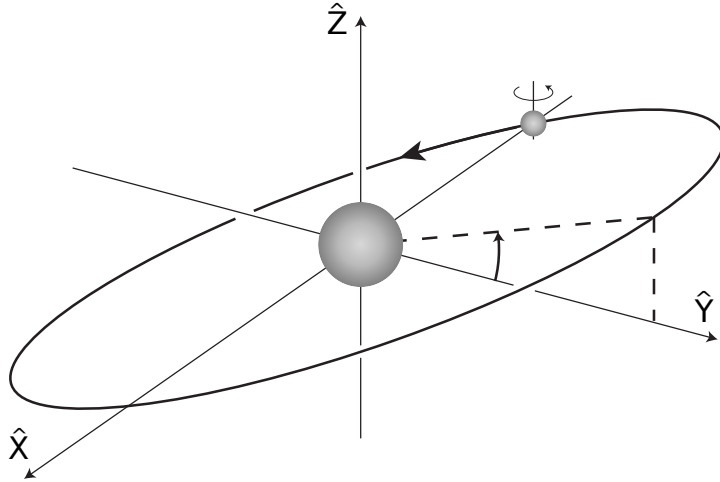

Figure 7: The sun-centered frame. $\hat{X}$ points to the vernal equinox, $\hat{Z}$ along the rotational axis of the earth and $\hat{Y}$ complements the right-handed coordinate system [23].

equinoctes are defined by the points where the ecliptic intersects with the equator of the earth. In Fig. 7 the ecliptic is given by the plane in which the earth rotates around the sun and the equator is given by the plane spanned by $\hat{X}$ and $\hat{Y}$. At these points, the earth does not incline away from or to the sun, giving rise to the name 'equinox' (equal night). There are two equinoctes, the vernal (in spring) and the autumnal (in fall). Due to small variations in the orientation of the earth, the equinoctes are not constant. Therefore the SCF is defined with the vernal equinox at a specific moment in time: the epoch J2000, defined at January 1, 2000, 12^{h} UT1 [26]. The timekeeping will be discussed in more detail later on.

This definition of the SCF results in a non-rotating coordinate system. The main contributions that make coordinate systems on earth not inertial are the rotation of the earth around its own axis and the rotation of the earth around the sun. There are contributions from other rotations as well. For example there is the larger-scale rotation of the solar system around the galactic center, which is negligible on our time scales. The rotational axis of the earth itself has small variations as well. Precession causes the rotational axis to change orientation by roughly $1°$ in 72 years [26]. Additional perturbations to the rotational axis, like nutation and aberration, are so small that they are not considered. The solar day is defined by the rotation of the earth such that it regains its orientation with respect to the sun. The sidereal day is defined by one rotation of the earth in an inertial frame, which has a frequency Ω. Due to the motion of the earth around the sun, it takes roughly 4 minutes longer for the earth to regain its orientation toward the sun, than to rotate around its axis in an inertial frame. Therefore a year contains roughly one sidereal day more than solar days.

A rotation from the SCF to the LHCb lab frame is needed to get the lab frame in terms of the SCF. Four well identifiable rotations are performed: the latitude λ, the azimuth θ, the tilt of the lab frame axes α and the sidereal phase Ωt. The complete description of the rotations is given in appendix A. The first three rotations are performed with proper Euler angles. The last angle, the sidereal phase, is performed using a standard rotation matrix over the $\hat{Z}$-axis. First the following terms are defined,

$$\sin(\lambda) \equiv s_1 \qquad \sin(\theta) \equiv s_2 \qquad \sin(\alpha) \equiv s_3 \qquad \sin(\Omega t) \equiv s_t$$
$$\cos(\lambda) \equiv c_1 \qquad \cos(\theta) \equiv c_2 \qquad \cos(\alpha) \equiv c_3 \qquad \cos(\Omega t) \equiv c_t,$$

with which the rotation matrix $R \equiv R(\lambda, \theta, \alpha, \Omega t)$ can be written as,

$$R = \begin{pmatrix} -c_t s_1 s_2 + s_t c_2 & c_t(c_3 c_1 + s_1 c_2 s_3) + s_t s_2 s_3 & -c_t(s_1 c_2 c_3 - c_1 s_3) - s_t c_3 s_2 \\ -s_t s_1 s_2 - c_t c_2 & s_t(c_3 c_1 + s_1 c_2 s_3) - c_t s_2 s_3 & -s_t(s_1 c_2 c_3 - c_1 s_3) + c_t c_3 s_2 \\ c_1 s_2 & s_1 c_3 - c_2 c_1 s_3 & s_1 s_3 + c_2 c_3 c_1 \end{pmatrix}, \quad (42)$$

such that $\vec{v}_{\mathrm{SCF}} = R \cdot \vec{v}_{\mathrm{lab}}$.

With the rotation matrix from Eq. 42, the vector for the beam can be determined in the SCF. This vector gives the average direction to where the B^0 mesons are

collimated to. For simplification, the effective angles χ and $\tilde{\lambda}$ are introduced. χ is the angle between the beam and the rotational axis of the earth and $\tilde{\lambda}$ is the effective latitude between the $\hat{y}$-axis of the LHCb lab frame to the rotational axis of the earth. $\tilde{\lambda}$ is defined in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and χ in $[0, \pi]$. The unit vector for the beam in the SCF is given by,

$$\hat{z}_{beam} = \begin{pmatrix} -c_t(s_1 c_2 c_3 - c_1 s_3) - s_t c_3 s_2 \\ -s_t(s_1 c_2 c_3 - c_1 s_3) + c_t c_3 s_2 \\ s_1 s_3 + c_2 c_3 c_1 \end{pmatrix} \tag{43}$$

and with $\hat{z}_{beam} \cdot \hat{Z} = \cos(\chi)$, gives the following value for $\cos(\chi)$,

$$\cos(\chi) = \cos(\theta)\cos(\lambda)\cos(\alpha) + \sin(\alpha)\sin(\lambda). \tag{44}$$

With the $\hat{y}$-axis of the lab frame in the SCF, the effective latitude is given with $\hat{y} \cdot \hat{Z} = \sin(\tilde{\lambda})$ by,

$$\sin(\tilde{\lambda}) = \sin(\lambda)\cos(\alpha) - \sin(\alpha)\cos(\theta)\cos(\lambda). \tag{45}$$

The simplification leads to the following definition of the rotation matrix,

$$R_{\text{simple}} \equiv \begin{pmatrix} \cos(\Omega t) & -\sin(\Omega t) & 0 \\ \sin(\Omega t) & \cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\chi) & 0 & \sin(\chi) \\ 0 & 1 & 0 \\ -\sin(\chi) & 0 & \cos(\chi) \end{pmatrix}, \tag{46}$$

such that $z_{beam}^{\text{SCF}} = R_{\text{simple}} \cdot z_{beam}^{\text{lab}}$.

3.2.3 Spherical coordinates

For a non-collimated analysis, right-handed spherical coordinates, θ_{lab} and ϕ_{lab}, are introduced. A boost, $\gamma\vec{\beta}$, in the lab frame is given by,

$$\gamma\vec{\beta}_{\text{lab}} = \gamma|\vec{\beta}| \begin{pmatrix} \cos(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) \\ \sin(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) \\ \cos(\theta_{\text{lab}}) \end{pmatrix}. \tag{47}$$

Eq. 46 simplifies the coordinate transformation by writing $\vec{\beta}$ in the SCF as $\vec{\beta} = R_{\text{simple}} \cdot \vec{\beta}_{\text{lab}}$,

$$\vec{\beta} = |\vec{\beta}| \begin{pmatrix} \cos(\Omega t)[\cos(\chi)\cos(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) + \sin(\chi)\cos(\theta_{\text{lab}})] - \sin(\Omega t)\sin(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) \\ \sin(\Omega t)[\cos(\chi)\cos(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) + \sin(\chi)\cos(\theta_{\text{lab}})] + \cos(\Omega t)\sin(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) \\ \cos(\chi)\cos(\theta_{\text{lab}}) - \sin(\chi)\cos(\phi_{\text{lab}})\sin(\theta_{\text{lab}}) \end{pmatrix}. \tag{48}$$

3.2.4 Spherical coordinate phase

With the simpler χ rotation from Eq. 46, the $\hat{x}$ and $\hat{y}$ axes are wrongly orientated in the SCF, but in the right plane. To correct for this, the substitution $\phi_{\text{lab}} \to \phi_{\text{lab}} + \phi_0$ is introduced. From Eq 45, the effective latitude $\tilde{\lambda}$, at which the $\hat{y}$ vector should be with

respect to the $\hat{Z}$-axis in the SCF, is known. By setting the spherical coordinates ϕ_{lab} = $\pi/2$ and $\theta_{\text{lab}} = \pi/2$, the $\hat{y}$ lab-frame axis is retained. Using Eq. 43 and Eq. 48, ϕ_0 is determined by,

$$\sin(\tilde{\lambda}) = -\sin(\chi)\cos(\pi/2 + \phi_0) = \sin(\chi)\sin(\phi_0)$$

$$\sin(\phi_0) = \frac{\sin(\tilde{\lambda})}{\sin(\chi)} = \frac{\sin(\lambda)\cos(\alpha) - \sin(\alpha)\cos(\theta)\cos(\lambda)}{\sin\chi}. \tag{49}$$

Eq. 49 gives two solutions, one with an $\hat{x}$-axis pointing with positive $\hat{Z}$ and one with negative $\hat{Z}$ component. Since the LHCb $\hat{x}$-axis points out of the LHC, roughly towards the southeast and LHCb is on the northern hemisphere, the $\hat{x}$-axis has a negative $\hat{Z}$ component. This leaves one solution for the spherical coordinate phase,

$$\phi_0 = 51.512 \pm 0.003°$$

3.2.5 Sidereal phase

Finally the sidereal phase of the beam axis of LHCb with respect to the vernal equinox at the J2000 epoch has to be determined. Until now, in Eq. 42, $\hat{X}$ points at 0° longitude (Greenwich meridian) at $t = 0$. However, it should point at the vernal equinox at J2000.

To be able to proceed, timekeeping methods need to be discussed. Due to the two main rotations of the earth, three types of timekeeping are considered: rotational (1), solar (2) and fixed (3).

1. The rotational timekeeping used is Greenwich mean sidereal time (GMST), given in sidereal seconds such that one rotation, a sidereal day T_{sid}, is 86400 GMST seconds. GMST is based on observations of distant stars and quasars and is defined as the angle with respect to the vernal equinox at J2000.

2. Solar timekeeping is based on the mean sun crossing a certain meridian at a fixed time. The mean is used, since a solar day varies periodically by the order of a minute due to the eccentricity of the earth's orbit. Used here is universal time (UT1), given by the mean sun crossing the Greenwich meridian at exactly 12:00 hours and a mean solar day, T_{sol}, takes 86400 UT1 seconds. UT1 is formally defined by a fixed transformation from GMST.

3. The problem with these timekeepings is that they depend on rotations that are not exactly constant and seconds of UT1 and GMST are therefore not constant. To resolve this, fixed timekeeping is introduced with the introduction of the SI second, using atomic clocks. This way sidereal and mean solar days can be expressed in SI seconds. A hybrid timekeeping used throughout the world is universal coordinated

time (UTC), which evolves in SI seconds and a UTC day consists normally of 86400 SI seconds. To get the solar properties of UT1, UTC is kept within one SI second of UT1, using leap seconds, roughly adding one SI second to a UTC day once a year.

From Aoki *et al.* [26], the durations of the sidereal, including precession, and mean solar day are given in Table 8, including a summary of the timekeeping methods. The timekeeping at LHCb is done using a GPS receiver that gives a UTC time stamp, t_{GPS}, in microseconds since 00:00:00 at January 1$^{\text{st}}$, 1970 UTC, called the Unix epoch. To go from t_{GPS}, in UTC, to GMST in radians, a linear transformation is needed, consisting of a radial rate of change Ω and a phase t_0, given by,

$$\Omega t = \Omega t_{\text{GPS}} + t_0. \tag{50}$$

The rate of change Ω is expressed in radians per UT1 microseconds as,

$$\Omega = \frac{2\pi}{T_{\text{sid}}} = 7.29211515 \cdot 10^{-11} \, \text{rad} \cdot \text{µs}^{-1}(\text{UT1}), \tag{51}$$

and can be used accurately, since UT1 $\approx$ UTC. With the reported significance, Ωt is accurate over 10 years (since J2000) up to a few seconds, due to the accuracy of T_{sid} up to irregular variations of a few microseconds [26]. This error is translated into the error reported on t_0.

A simple transformation is possible from the Unix epoch to the J2000 epoch. Often used in astronomy is the Julian date (JD), to give a date as one linearly-evolving day number. This way, J2000 is defined at JD 2451545.0 UT1 [26]. Counting back to the Unix epoch gives 30 Julian years consisting of 27 years of 365 days and 7 leap years and

Temporal definitions	
GMST	Hour angle of Greenwich meridian w.r.t. vernal equinox,
	i.e. following the rotation of the earth.
	$T_{\text{sid}} = 86164.0989$ s (UT1)
UT1	12:00:00 UT1 given by mean sun crossing Greenwich meridian,
	based on GMST.
	$T_{\text{sol}} \equiv 86400$ s (UT1) ≈ 86400.002 s (SI)
UTC	Similar to UT1, but days defined at default by 86400 s (SI).
	Occasionally leap second added/subtracted to keep $|\text{UTC} - \text{UT1}| < 1$ s.
Epoch definitions	
J2000	12:00:00 UT1, January 1$^{\text{st}}$, 2000 (JD 2451545.0 UT1)
Unix epoch	00:00:00 UTC, January 1$^{\text{st}}$, 1970 (JD 2440587.5 UTC)

Table 8: List of used temporal definitions [26].

half a day, resulting in -10957.5 Julian days, setting the Unix epoch at JD 2440587.5. The transformation from UT1 to GMST, in GMST seconds, defining UT1, is given by,

$$\text{GMST of } 0^{\text{h}} \text{ UT1} = 24110.54841 + 8640184.812866 T_{\text{U}} + 0.093104 T_{\text{U}}^2 - 6.2 \cdot 10^{-6} T_{\text{U}}^3, \quad (52)$$

where $T_{\text{U}} = (\text{JD} - 2451545.0)/36525$, where JD is the number of Julian date days of UT1 [26]. T_{U} is essentially Julian centuries since the J2000 epoch and shows that quadratic corrections are very small, within one century. The linear term shows roughly the extra day the earth has to turn in a year to regain its orientation towards the sun. The resulting sidereal phase of the Greenwich meridian at the Unix epoch is then given by,

$$t_{\text{Unix}} = 6.6820 \pm 0.0014 \, \text{h (GMST)}, \quad (53)$$

where the error is estimated to be about 5 seconds. This leaves us to find the effective longitude of LHCb with respect to the Greenwich meridian. The effective longitude, l_0, defined as the angle of the beam with respect to the Greenwich meridian is given by the $\hat{x}$ and $\hat{y}$ components of $\hat{z}_{beam}$ from Eq. 43 at $\Omega t = l$,

$$\tan(l_0) = \frac{\hat{z}_y}{\hat{z}_x}, \quad (54)$$

and using the parameters for LHCb from Table 7 gives,

$$l_0 = -58.041 \pm 0.003°. \quad (55)$$

This way, the sidereal phase at $t_{\text{GPS}} = 0$, t_0, will be l_0 added to the angle, t_{Unix}, of the Greenwich meridian with respect to the vernal equinox at J2000, at the Unix epoch. The sidereal phase, Ωt, is then given by,

$$\Omega t = \Omega t_{\text{GPS}} + l_0 + t_{\text{Unix}} = \Omega t_{\text{GPS}} + t_0. \quad (56)$$

The resulting parameters for LHCb are given in Table 9.

Sidereal-phase parameters for LHCb			
l_0	-3.8694	± 0.0002	h (arc)
t_{Unix}	6.6820	± 0.0014	h (GMST)
t_0	2.8126	± 0.0014	h (GMST)
t_0	0.7363	± 0.0004	rad

Table 9: Effective longitude and sidereal phases. t_0 can be expressed in an (earth rotation) angle, therefore radians, as well.

3.2.6 Parameters for LHCb

With the positional angles of LHCb from Table 7, Eq. 44, Eq. 49, Eq. 51 and the calculation for the sidereal phase from Table 9, the relevant parameters needed in this analysis are derived and are summarized in Table 10.

	SCF parameters for LHCb		
χ	112.40702	$\pm$ 0.00003	°
ϕ_0	51.512	$\pm$ 0.003	°
t_0	0.7363	$\pm$ 0.0004	rad
Ω	$7.29211515 \cdot 10^{-11}$		rad $\cdot\, \mu s^{-1}$ (UT1)
$\cos(\chi)$	-0.38118	$\pm$ 0.00003	
$\sin(\chi)$	0.92450	$\pm$ 0.00001	

Table 10: Positional angles of LHCb in the SCF.

4 Likelihood Fit

To subtract the *CPT* and Lorentz violating parameter $\Delta a_\mu^{B^0}$ from the 2011 and 2012 data set of LHCb, a decay time dependent analysis will be used. Essentially, this comes down to maximizing the likelihood of the probability density functions (PDF) for B^0 and $\overline{B}^0$ decay rates, provided by Eq. 40. The final state $J/\psi \, K_s^0$ is reconstructed in the $J/\psi \rightarrow \mu^+ \mu^-$ and $K_s^0 \rightarrow \pi^+ \pi^-$ channels. Due to the small decay time of the J/ψ of $\mathcal{O}(10^{-20}\,\mathrm{s})$ and the excellent detectability of a muon, the vertex of the B^0 decay can be reconstructed. In case of proton-proton interactions, one can pinpoint the position of the collision, *i.e.* the primary vertex (PV), due to the many tracks originating from it. Since the lifetime of the B^0 is about $1.5\,\mathrm{ps}$ and $\gamma\beta \approx 20$, the average traveling distance of the B^0 is about $9\,\mathrm{mm}$ and this distance can be detected by the VELO. The B^0 decay distance and B^0 momentum provide the measurement of the decay time. The K_s^0 travels on average $90\,\mathrm{cm}$ before decay and this long flight distance causes a part of the K_s^0 mesons to travel outside the VELO before decaying. The decay topology of $B^0 \rightarrow J/\psi \, K_s^0 \rightarrow \mu^+ \mu^- \pi^+ \pi^-$ is depicted in Fig. 8. In order to determine in what interaction state the B^0 was at production, a method called flavour tagging is used. Flavour tagging is based on clues left behind from surrounding particles that shared quark pairs during hadronization.

This section on the likelihood fit is arranged as follows. First, the data set is discussed in section 4.1, including cuts on the data and triggers. Then the mass fit and their subsequent weights in section 4.2. This is followed by the decay time acceptance, *i.e.* a decay time dependent efficiency caused by the trigger, which will be presented in section 4.3. This is followed by a discussion on the decay time resolution in section 4.4 and the flavour tagging in section 4.5. Leading all to the full decay time dependent PDF presented in section 4.6. After the full PDF description, the strategy of fitting this PDF will be discussed in section 4.7. The fit results will be shown in section 4.8 and finally, section 4.9 will describe the systematics qualitatively. A quantative determination of the systematics still has to be done. The data set used by the analysis, the mass fit and the calibration of the acceptance, decay time resolution and flavour tagging is the same as
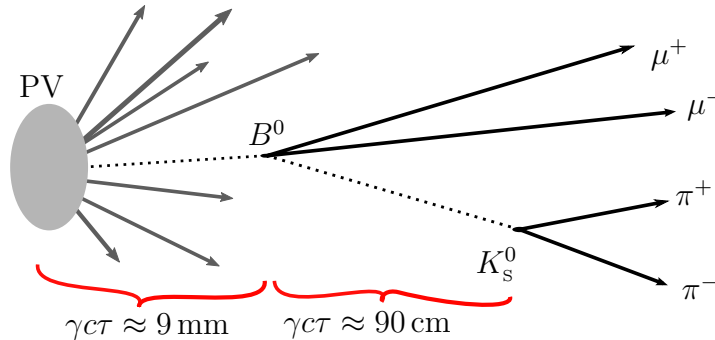

Figure 8: Decay topology with the, from the PV detached, decay vertices of the B^0 and the K_s^0.

used in the LHCb analysis measuring $\sin(2\beta)$ with the same decay [27]. The code of the likelihood fit, including the additional CPT violating terms, is written by the author.

4.1 Data set

From the LHCb run 1 in 2011 and run 2 in 2012, data is collected corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. This analysis uses hardware triggers on muons with a $p_T > 1.48\,\text{GeV}$. Software triggers, using event reconstruction, use so called trigger lines. These trigger lines have certain requirements, for example that the events contain a dimuon detached from the PV with an invariant mass around the J/ψ mass. In addition, requirements on track and PID quality are used in the software triggers. The trigger will be discussed more in section 4.3. After the trigger, a method called stripping is used: Data output after triggering is still to high and additional cuts to enhance purity and in some cases pre-scaling are applied. The decay chain of $B^0 \to J/\psi\,K_S^0 \to \mu^+\mu^-\pi^+\pi^-$ is fitted using the DecayTreeFitter (DTF) algorithm [28]. This algorithm fixes the origin of the tracks of the muons and of the pions and sets the masses of the J/ψ and the K_S^0 to their PDG values of $m_{J/\psi} = 3096.916 \pm 0.011\,\text{MeV}$ and $m_{K_S^0} = 497.614 \pm 0.024\,\text{MeV}$. In addition, the PV is reconstructed from other tracks. The momenta of the muons and pions are obtained from the fitting of the tracks with the DecayTreeFitter. To correct for biases, momentum scaling with data is used. With the DTF, values for the invariant mass, $m_{J/\psi K_S^0}$, momentum, $\vec{p}_{J/\psi K_S^0}$, decay time t and decay time uncertainty estimate σ_t of the B^0 are obtained. Flavour tagging algorithms return a tag decision d and a prediction of the probability that this tag decision is wrong, called η. For the Lorentz violation study, GPS time stamps are used and they are supplied by a GPS receiver.

Next to the trigger and stripping, offline cuts will be applied to enhance purity. With the PDG mass of the B^0, $m_{B^0} = 5279.58 \pm 0.17\,\text{MeV}$, a window of $5220 < m_{J/\psi K_S^0} < 5330\,\text{MeV}$ is used. In order to reduce background from the many prompt decays that are in the events, the decay time minimum is $0.3\,\text{ps}$. The maximum is set at 18.3. As for the decay time uncertainty estimate, the maximum is set at $0.2\,\text{ps}$. For flavour tagging, only tagged events are used in the likelihood fit. A list of all the cuts, including stripping and triggers, is documented in an internal analysis note of LHCb for the $\sin(2\beta)$ measurement [27]. A summary of the observables and their cuts is given in Table 11.

Possible backgrounds are inclusive $J/\psi X$ final states and misidentifications of kaons to pions from $B^0 \to J/\psi\,K^{*0}$ and protons to pions from $\Lambda_b^0 \to J/\psi\,\Lambda$. After trigger, stripping and offline cuts applied to a Monte Carlo data set with a few million of these inclusive decays, no background events remain of the misidentification events. From the inclusive J/ψ sample, a negligible amount remains. This shows that only combinatorial background is left.

The data set will be divided into 16 disjoint categories to distinguish different properties of the data for each category. The 16 categories consist of combinations of the

Observable	Description	Cut
$m_{J/\psi K^0_S}$	invariant mass of $J/\psi K^0_S$	$5220 - 5330\,\mathrm{MeV}$
t	decay time reconstructed from discplacement and momenta	$0.3 - 18.3\,\mathrm{ps}$
σ_t	decay time uncertainty prediction	$0.0 - 0.2\,\mathrm{ps}$
d_{OS}	tag decision from opposite side (OS)	$+1, -1$
d_{SS}	tag decision from same side pion (SS)	$+1, -1$
η_{OS}	mistag probability prediction from OS	$0.0 - 0.5$
η_{SS}	mistag probability prediction from SS	$0.0 - 0.5$
$p_{J/\psi K^0_S}$	momentum of $J/\psi K^0_S$	
t_{GPS}	GPS time stamp	

Table 11: Observables used in analysis with cuts applied to the data.

following categories:

1. **Trigger.** Due to the use of triggers that give a bias towards higher decay times, two categories are considered: almost unbiased (AU) and exclusively biased (EB). These two categories will be discussed more in section 4.3.

2. **Track.** As mentioned earlier, due to the long flight distance of the K^0_S, the K^0_S can decay outside the VELO. Decaying outside the VELO changes the momentum resolution significantly, therefore the following categories are used: downstream track (DD), *i.e.* outside the VELO, and long track (LL), *i.e.* inside the VELO.

3. **Tagging.** As will be explained in section 4.5, flavour tagging will be divided into two categories: an exclusive opposite side (OS) tag or a same side (SS) pion tag that may also include an OS tag. Although the tuple used in this analysis contains untagged data, in the likelihood fit, they are not used.

4. **Year.** Due to different center-of-mass energies, properties, for example production asymmetry, are different per run. Therefore, for the year of the run, categories 2011 and 2012 are introduced.

4.2   Mass fit and sWeights

To distinguish background from signal, an invariant mass fit of the $J/\psi K^0_S$ is performed. As discussed earlier, the background is combinatorial. The distribution used for the background is exponential. For the signal, the Ipathia function is used, which is a generalization of the Crystal Ball function [29]. The result of this fit for all 16 categories combined is given in Fig. 9.

From the mass fit, sWeights are obtained. sWeights are used to subtract background in sums over data [30]. This will be done for the sum over the log-likelihood as well and will be discussed further in section 4.7. sWeights have the following properties,

$$\sum_j^{N_s} w_j = N_s \langle w_j \rangle_s = N_s \qquad \qquad \sum_j^{N_b} w_j = N_b \langle w_j \rangle_b = 0, \qquad (57)$$

where w_j is the sWeight and N_s and N_b are the number of signal and background events. The sums are over the signal and background events, respectively. This shows that if a certain variable X is not correlated to w_j, than $\langle w_j X \rangle = \langle w_j \rangle \langle X \rangle$, and therefore the background cancels and the signal remains.
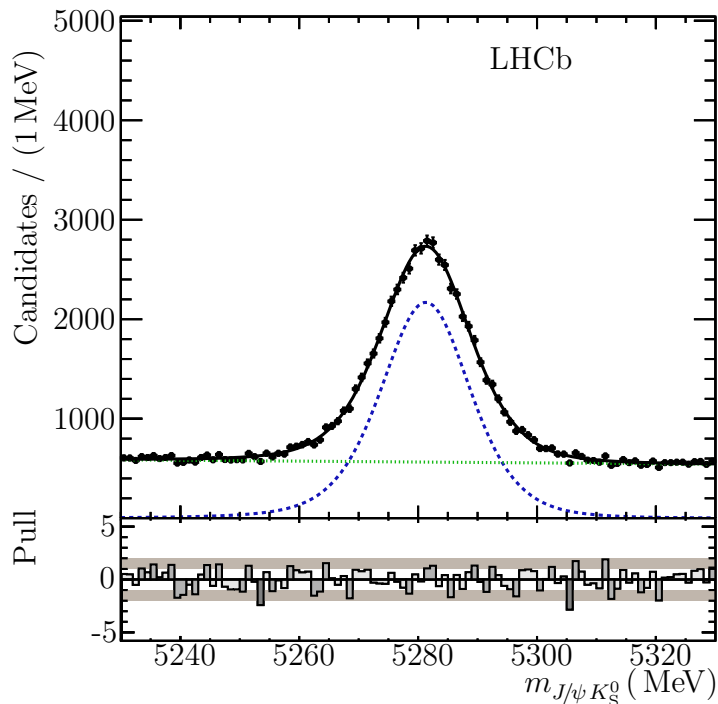

Figure 9: Invariant-mass fit of all 16 categories [27].

4.3 Acceptance

Triggers used in this analysis cause a bias towards higher lifetimes of the B^0. To correct for this bias, an efficiency function, called acceptance is used. Two types of triggers are considered: `Hlt1` and `Hlt2`. `Hlt1` is the first level of the software trigger. The trigger lines from `Hlt1` used in this analysis are `Hlt1DiMuonHighMass` and `Hlt1TrackMuon`. `Hlt1DiMuonHighMass` selects dimuon candidates on track momentum and quality and an invariant mass higher than 2.7 GeV. This line is not biased, since there are no selection requirements that depend on decay time. The other `Hlt1` trigger line is `Hlt1TrackMuon`. In addition to track momentum and quality cuts, it also requires an impact parameter with respect to a primary vertex to be higher than 0.1 mm. This introduces a bias towards higher decay times. For the second level of the software trigger, `Hlt2`, there are again two lines used in this analysis: `Hlt2DiMuonJPsi` and `Hlt2DiMuonDetachedJPsi`. The trigger line `Hlt2DiMuonJPsi` requires the dimuon to have an invariant mass within $m_{J/\psi}^{\mathrm{PDG}} \pm 0.12$ GeV. A random pre-scaling of 0.2 is used to reduce the output. Since no decay time dependent cuts are used, there is no bias in decay time. The second `Hlt2` line, `Hlt2DiMuonDetachedJPsi`, has the same cuts as `Hlt2DiMuonJPsi`, but adds one additional cut that requires the flight distance (FD) to have $\mathrm{FD}\chi^2 > 9$. This cut introduces a bias towards higher decay times. Any further cuts of the stripping and offline selection do not introduce any bias. With before mentioned trigger lines, two types of trigger efficiencies are introduced: almost unbiased (AU) and exclusively biased (EB). The definition of the efficiencies is given by:

1. **AU.**

$$\varepsilon_{\mathrm{AU}} = \frac{\texttt{Hlt1DiMuonHighMass \&\& Hlt2DiMuonDetachedJPsi \&\& Hlt2DiMuonJPsi}}{\texttt{Hlt1DiMuonHighMass \&\& Hlt2DiMuonJPsi}},$$

(58)

where the denominator is the combination of the two unbiased trigger lines. The numerator adds an additional line, a detached J/ψ, and therefore introduces a slight bias.

2. **EB.**

$$\varepsilon_{\mathrm{EB}} = \frac{\texttt{Hlt1TrackMuon \&\& Hlt2DiMuonDetachedJPsi \&\& !Hlt1DiMuonHighMass}}{\texttt{Hlt1DiMuonHighMass \&\& Hlt2DiMuonJPsi}},$$

(59)

where, as the name suggests, the trigger lines in the numerator are biased and unbiased lines are explicitly left out. The same unbiased denominator is used.

From the definition of these efficiencies, an acceptance function $a(t)$ can be constructed by binning these efficiencies in decay time with the bin boundaries chosen such that each bins contains the same amount of AU+EB data. After splitting in AU and EB, the signal yield in these bins is obtained from a mass fit on the data per. The obtained histogram is not directly used in the likelihood fit, but an interpolation method, using cubic splines [31]. The resulting efficiencies per bin in decay time for AU and EB are used as spline parameters

and each bin center is used as knot. The resulting cubic-spline functions are depicted in Fig. 10. One clearly sees that the categories AU and EB are appropriately named. For low decay times, EB has a lower efficiency. For AU, there is a slight drop in efficiency for low decay times. After the trigger, due to reconstruction inefficiencies and selection effects, an
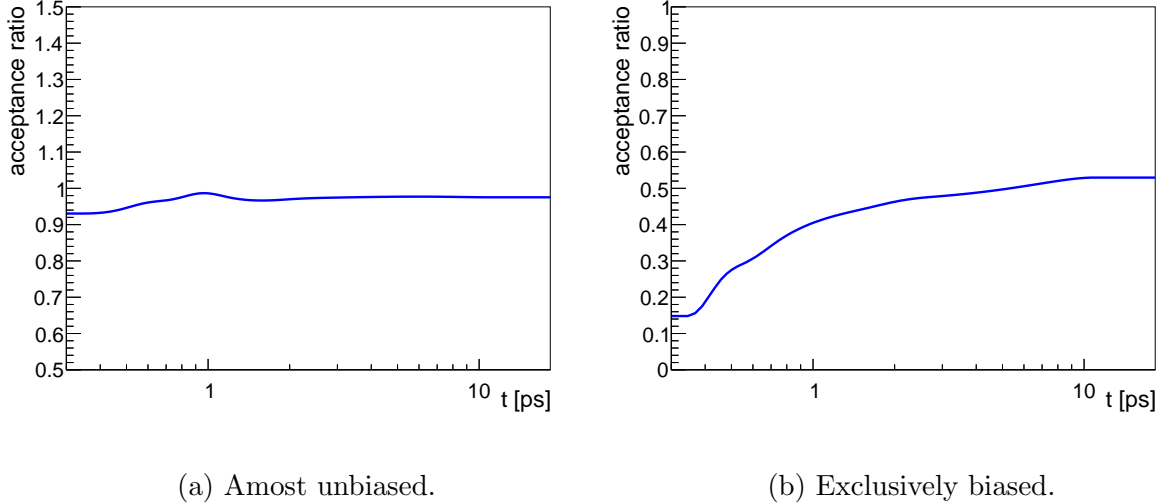

(a) Amost unbiased. (b) Exclusively biased.

Figure 10: Qubic-spline acceptance functions.

additional bias towards higher lifetimes is introduced in the data sample. To correct for this, a modified lifetime is used,

$$\tilde{\tau} = \frac{\tau}{1 + \beta_\tau \tau}, \tag{60}$$

where β_τ is the correction factor. This factor is obtained from applying the same selection to a signal MC data set of $B^0 \to J/\psi K_{\mathrm{s}}^0$. Since the true lifetime is known, the correction factor can be calculated. β_τ is categorized into track type (DD/LL) and year (2011/2012) and its values are given in Table 12.

	2011	2012
DD	0.036 ± 0.0029	0.0084 ± 0.0032
LL	0.018 ± 0.004	0.0035 ± 0.005

Table 12: β_τ correction factors for high lifetime per year and track type. Obtained from MC signal data for $B^0 \to J/\psi K_{\mathrm{s}}^0$.

4.4 Decay time resolution

To determine the decay time resolution, a sample of prompt J/ψ resonances is used. This data sample is obtained by using unbiased trigger lines without cuts on decay time. This

introduces a large background of J/ψ resonances created at the PV. Additional background of fake J/ψ are filtered out using sWeights. By determining the decay time the same way as the data set for the main likelihood fit, a good representation of the resolution is obtained, since the J/ψ resonances from this background are smeared the same way as signal B^0 in the decay time distribution, due to the finite resolution of the VELO and momentum resolution. The shape of this distribution is used as the decay time resolution in the main likelihood fit and is given by three Gaussians according to,

$$
\begin{aligned}
\mathcal{R}(t|\sigma_t) = \sum_{i=1}^{2} g_i \cdot \frac{1}{\sqrt{2\pi}(b_i\sigma_t + c_i)} \exp\left(-\frac{(t-\mu_t)^2}{2(b_i\sigma_t + c_i)}\right) \\
+ f_{\mathrm{PV}} \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{PV}}} \exp\left(-\frac{(t-\mu_t)^2}{2\sigma_{\mathrm{PV}}^2}\right),
\end{aligned}
\tag{61}
$$

where the first two Gaussians are used to model the prompt J/ψ peak and the last one, which contributes very little, is added to account for wrongly associated PVs. In the LHC, multiple pp collisions can happen per bunch crossing and therefore multiple PVs can occur. In addition, the decay time uncertainty estimate from the DTF, σ_t, is calibrated to data by binning data in σ_t and then determining the average resolution per bin. This calibration is linear, with parameters b_i and c_i. The mean of the resolution is roughly $3\,\mathrm{fs}$, but is left out due to problems in implementing this in code. This can later be investigated as possible systematic uncertainty. The values of the parameters in Eq. 61 are given in Table 13.

In the code of the likelihood fit, a combination of acceptance function and decay time resolution is used to speed up the fitter [31]. This uses analytical convolution integrals, which is possible due to the use of qubic splines for the acceptance and Gaussians for the resolution.

		DD	LL
μ_t	[ps]	-0.0291 ± 0.00026	-0.00169 ± 0.00026
b_1		0.88 ± 0.09	1.04 ± 014
c_1	[ps]	0.0077 ± 0.0028	0.0045 ± 0.0028
b_2		1.33 ± 0.33	1.8 ± 0.4
c_2	[ps]	0.019 ± 0.008	0.007 ± 0.005
g_2		0.251 ± 0.020	0.24 ± 0.023
σ_{PV}	[ps]	1.6 ± 0.7	1.4 ± 0.14
f_{PV}		0.048 ± 0.004	0.0488 ± 0.0024

Table 13: List of values for parameters of the decay time resolution. Categorized in track type. The parameter g_1 is given by $g_1 = 1 - g_2 - f_{\mathrm{PV}}$.

4.5 Flavour tagging

B^0 mesons are predominantely produced from hadronizing $b\bar{b}$ pairs. During hadronization, a $\bar{b}$ quark can pick up a d quark, which in turn is coming from a $d\bar{d}$ pair. Since the $\bar{d}$ quark is likely to hadronize into a π^+, the appearance of a π^+ in the vicinity of the signal B^0 enhances the probability that the B^0 is a B^0 and not a $\overline{B}^0$. In the same way, a π^- enhances the probability of a B^0 being a $\overline{B}^0$. This method is called the same side (SS pion) tagger. In addition, the other quark from the $b\bar{b}$ pair can reveal the flavour of the signal B^0. This opposite side (OS) tagging can for example be done by the other b quark hadronizing into a B meson that decays into a flavour specific final state, for example with a charged lepton. Multiple OS tagging methods are combined to form one OS tagging decision. To summarize, in this analysis flavour tagging is done the combined OS and the SS pion tagger. A schematic overview of the tagging principles is given in Fig. 11.

Flavour tagging is quantified by the mistag probability ω defined by,

$$\omega = \frac{N_W}{N_W + N_R}, \tag{62}$$

where N_W and N_R are the number of wrong and right tagged particles respectively. The range of ω is chosen between $0 - 0.5$, otherwise the tagging decision d would be flipped. The tagging algorithms base their output on a neural net, trained on a sample of B^+ $\rightarrow J/\psi\,K^+$, which due to their charge is self tagging. It looks at charges, PID and geometric and kinematic properties of events to determine a tag decision d and a prediction of the
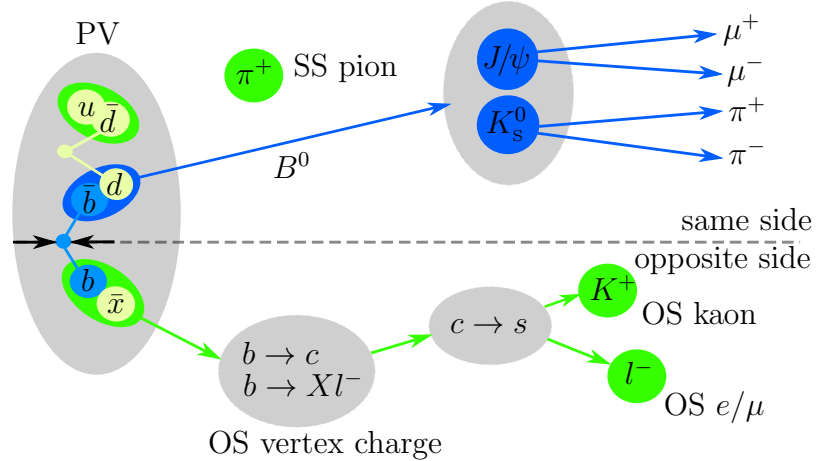

Figure 11: Flavour tagging. Schematic representations of opposite and same side taggers. Same side uses $d\bar{d}$ resulting in π^+ or π^- as tagger. Opposite side uses the other b or $\bar{b}$ quark from $b\bar{b}$ that forms a b-hadron or $\bar{b}$-hadron as tagger.

mistag probability η. This prediction in turn is calibrated to data with self taggers like $B^+ \to J/\psi K^+$ by the Flavour Tagging Group in LHCb. The calibration is a first order correction according to,

$$\omega = p_1(\eta - \langle\eta\rangle) + p_0, \tag{63}$$

where p_1 and p_0 are real constants and $\langle\eta\rangle$ is the average of η of the signal. Tagging can have different performances for a specific tag. To account for these, a mistag calibration per tag is introduced,

$$
\begin{aligned}
\omega &= (p_1 + \frac{\Delta p_1}{2})\,[\eta - \langle\eta\rangle] + p_0 + \frac{\Delta p_0}{2} \\
\overline{\omega} &= (p_1 - \frac{\Delta p_1}{2})\,[\eta - \langle\eta\rangle] + p_0 - \frac{\Delta p_0}{2},
\end{aligned}
\tag{64}
$$

where Δp_0 and Δp_1 parametrize the asymmetry between tags in tagging calibration parameters p_0 and p_1 respectively. Another asymmetry can arise if one tag compared to the other tag is tagged more often. This causes a tagging efficiency asymmetry. This asymmetry is in this analysis so small that it is neglible.

4.5.1 Decay rates with flavour tagging

With these definitions, the effect of tagging in the decay rate equations can be included. As mentioned in section 2.2.4, there is a production asymmetry of B^0 mesons with proton-proton collisions. This asymmetry will be included in these effects by introducing a general normalisation $N = N^{B^0} + N^{\overline{B}^0}$, such that $N_{B^0} = N(1 - A_p)/2$ and $N_{\overline{B}^0} = N(1 + A_p)/2$. If there was a tagging efficiency asymmetry, one would need to account for that similar to production asymmetry, but as mentioned earlier, this is negligible. Suppose there are the following general decay rate equations, Γ for B^0 and $\overline{\Gamma}$ for $\overline{B}^0$. In the measured decay rate equations, Γ^{meas}, there are contributions of both real tags due to mistagging. The measured decay rates are given by,

$$
\begin{aligned}
\Gamma^{meas} &\propto (1 - A_p)(1 - \omega)\Gamma + (1 + A_p)\overline{\omega}\overline{\Gamma} \\
\overline{\Gamma}^{meas} &\propto (1 + A_p)(1 - \overline{\omega})\overline{\Gamma} + (1 - A_p)\omega\Gamma.
\end{aligned}
\tag{65}
$$

With these equations and without production asymmetry, the tagged asymmetry,

$$A_{CP/CPT} \propto (1 - 2\overline{\omega})\overline{\Gamma} - (1 - 2\omega)\Gamma, \tag{66}$$

introduces a dilution of the asymmetry, $D = 1 - 2\omega$. With the dilution, the tagging power can be derived, which is the effective efficiency of tagging. Suppose one has an amplitude S in the asymmetry. The uncertainty on S scales with $\sqrt{\varepsilon_{\text{tag}}}$, which is the square root of the fraction of tagged events. In addition, it gets diluted by mistagging with D. This results in a tagging power of,

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}D^2 = \varepsilon_{\text{tag}}(1 - 2\omega)^2. \tag{67}$$

The tagging power of this analysis is 3%. Compared to e^+e^- B-factories, this is quite low. This is due the hadronic environment of pp collisions.

4.5.2 Calibration for OS and SS

For the calibration of the OS tagger. $B^+ \to J/\psi\,K^+$ is used and as a cross-check, the B^0 $\to J/\psi\,K^{*0}$ decay is used. The charge of the K^+ or K^- provides the true tag and that way, one can compare the tagging algorithm with the true tag. The tagging calibration parameters for OS are given in Table 14.

	OS		
p_0	0.3815	± 0.0011 (*stat.*)	± 0.0016 (*syst.*)
p_1	0.978	± 0.012 (*stat.*)	± 0.009 (*syst.*)
Δp_0	0.0148	± 0.0016 (*stat.*)	± 0.0008 (*syst.*)
Δp_1	0.070	± 0.018 (*stat.*)	± 0.004 (*syst.*)
$\langle \eta \rangle$		0.3786	

Table 14: List of calibration parameters for OS.

For the SS pion tagger, the calibration is performed using data from $B^0 \to J/\psi\,K^{*0}$, which is flavour specific. The calibration parameters for SS are given in Table 15.

	SS		
p_0	0.4232	± 0.0029 (*stat.*)	± 0.0028 (*syst.*)
p_1	1.011	± 0.064 (*stat.*)	± 0.031 (*syst.*)
Δp_0	-0.0026	± 0.0042 (*stat.*)	± 0.0027 (*syst.*)
Δp_1	-0.17	± 0.10 (*stat.*)	± 0.04 (*syst.*)
$\langle \eta \rangle$		0.425	

Table 15: List of calibration parameters for SS.

Since in the data category SS, there can also be an OS tag, a combined tag has to be applied. Since OS and SS tags have very different tagging principles, correlation between the two taggers is expected to be negligible. These correlations have been studied and are indeed negligible. Therefore, the combination is given by multiplying the mistag probabilities. This way, decay rates can be written with tags d_{OS} and d_{SS} as,

$$
\begin{aligned}
\Gamma^{meas}(t, d_{OS}, d_{SS}) \propto\; & (1 - A_p)\left(\frac{1 + d_{OS}}{2} - d_{OS}\omega_{OS}\right)\left(\frac{1 + d_{SS}}{2} - d_{SS}\omega_{SS}\right)\Gamma(t) \\
& + (1 + A_p)\left(\frac{1 - d_{OS}}{2} + d_{OS}\overline{\omega}_{OS}\right)\left(\frac{1 - d_{SS}}{2} + d_{OS}\overline{\omega}_{SS}\right)\overline{\Gamma}(t),
\end{aligned}
\tag{68}
$$

where d_{OS} and d_{SS} can also take the untagged value of 0.

4.6 Decay time probability density function

With the description of the acceptance, decay time resolution and flavour tagging, a full decay rate PDF can be determined. Each PDF has the subscript s, which stands for the 16 data categories, to emphasize the fact that there are 16 full PDFs used. First, the theoretical PDF with the addition of flavour tagging is given. Using Eq. 40 and Eq. 68,

$$\mathcal{P}^s(t, d_{OS}, d_{SS}|\eta_{OS}, \eta_{SS}, p_{B^0}, t_{GPS}) =$$
$$\frac{e^{-t/\tilde{\tau}}}{N} \Bigg[(1 - A_p) \left(\frac{1 + d_{OS}}{2} - d_{OS}\omega_{OS} \right) \left(\frac{1 + d_{SS}}{2} - d_{SS}\omega_{SS} \right)$$
$$\left(1 - D_f \mathcal{R}e(z) + (C_f + D_f \mathcal{R}e(z))\cos(\Delta m t) - S_f \sin(\Delta m t) \right)$$
$$+ \ (1 + A_p) \left(\frac{1 - d_{OS}}{2} + d_{OS}\overline{\omega}_{OS} \right) \left(\frac{1 - d_{SS}}{2} + d_{OS}\overline{\omega}_{SS} \right)$$
$$\left(1 + D_f \mathcal{R}e(z) - (C_f + D_f \mathcal{R}e(z))\cos(\Delta m t) + S_f \sin(\Delta m t) \right) \Bigg],$$

(69)

where for $\mathcal{R}e(z)$ the collimated Eq. 35 is used with $\beta = 1$, due to the high boost, and the values for the sun-centered frame parameters from Table 10. The boost is determined by $\gamma = \sqrt{1 + p_{B^0}^2/m_{B^0}^2}$. This results in,

$$\mathcal{R}e(z) = \frac{1}{\Delta m_{B^0}} \sqrt{1 + \frac{p_{B^0}^2}{m_{B^0}^2}} \Bigg(\Delta a_0 + \cos(\chi)\Delta a_Z$$
$$+ \sin(\chi)\left[\Delta a_Y \sin(\Omega t_{GPS} + t_0) + \Delta a_X \cos(\Omega t_{GPS} + t_0)\right] \Bigg).$$

(70)

Applying the decay time resolution is done by convoluting Eq. 69 with $R(t|\sigma_t)$, since the distribution of each true decay time t' is smeared out by $R(t|\sigma_t)$. The acceptance $a(t)$ is used as an efficiency function, therefore multiplied with the convoluted PDF by,

$$\mathcal{P}^s_{meas}(t, d_{OS}, d_{SS}|\sigma_t, \eta_{OS}, \eta_{SS}, p_{B^0}, t_{GPS}) =$$
$$a(t) \cdot \int_{-\infty}^{+\infty} \mathcal{P}^s(t', d_{OS}, d_{SS}|\eta_{OS}, \eta_{SS}, p_{B^0}, t_{GPS}) \mathcal{R}(t - t'|\sigma_t) dt'.$$

(71)

With Eq. 71, the full PDF used in the likelihood fit is now described. How this fit will be performed will be discussed next.

4.7 Strategy

Having descibed the full decay rate PDF including *CP*, *CPT* violation and detection effects, it is now important to set a strategy. This means determining which parameters will be floating and what values the fixed or constrained parameters can take. In addition,

it has to be determined if and why a collimated analysis will be used. But first of all, the likelihood is defined by,

$$L = \prod_s \prod_i^{N_s} \mathcal{P}_{meas}^s \left(\vec{x}_i^s, \vec{\lambda} \right),$$ (72)

where $\vec{x}_i^s$ is the set of observables for data category s and event i and $\vec{\lambda}$ is the set of floating parameters. Likelihood maximization is much faster when using the log-likelihood, since products transform into sums. In addition, sWeights can be used to subtract the background,

$$\log(L) = \alpha \sum_s \sum_i^{N_s} w_i \log \left(\mathcal{P}_{meas}^s \left(\vec{x}_i^s, \vec{\lambda} \right) \right).$$ (73)

The log-likelihood is scaled by factor α, since uncertainties calculated by an sWeighted likelihood are underestimated. There is still a background dilution that has to be accounted for. This is done by setting,

$$\alpha = \frac{\sum_{i=1}^{N} w_i}{\sum_{i=1}^{N} w_i^2},$$ (74)

which holds if resolution, tagging and momentum have similar distributions for signal and backgrounds [32]. For resolution and tagging distributions, this is checked by the $\sin(2\beta)$ analysis [27]. In case of momentum, the distribution of the background has a slightly lower average momentum ($\langle \beta\gamma \rangle = 16.9$ for background and $\langle \beta\gamma \rangle = 19.5$ for signal). This difference is of the order of 10%. This is neglected and can be treated as a possible systematic uncertainty.

As a test, the fit has been performed using a non-collimated $\mathcal{R}e(z)$, but this resulted in very high (roughly 99%) correlations between $\Delta a_0^{B^0}$ and $\Delta a_Z^{B^0}$ and therefore a collimated analysis is used. To show how collimated the B^0 mesons are, the sWeighted distribution of spherical coordinate θ_{lab} is depicted in Fig. 12.

What is left to determine is which parameters should be floating, Gaussian constrained or set constant. The parameters that are left floating are, obviously $\Delta a_\mu^{B^0}$, but also the lifetime τ. This is done to ensure no bias is caused by possible higher lifetime effects. The parameters $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$ are set to their PDG values and Gaussian constrained to their uncertainties [9]. This can be done, since any CPT violation effects are expected to be much lower in this result, since the PDG values are dominated by e^+e^- B-factory results from BaBar and Belle and they have a factor 40 lower boost. Further, production asymmetries are Gaussian constrained, per year, since there are possible differences in production due to different center-of-mass energies. Since flavour tagging has the biggest effect on calculating time-dependent asymmetries in B^0 analyses, the flavour tagging parameters are Gaussian constrained with their statistical uncertainty. The systematic uncertainties can be used in the general systematics study. The mass
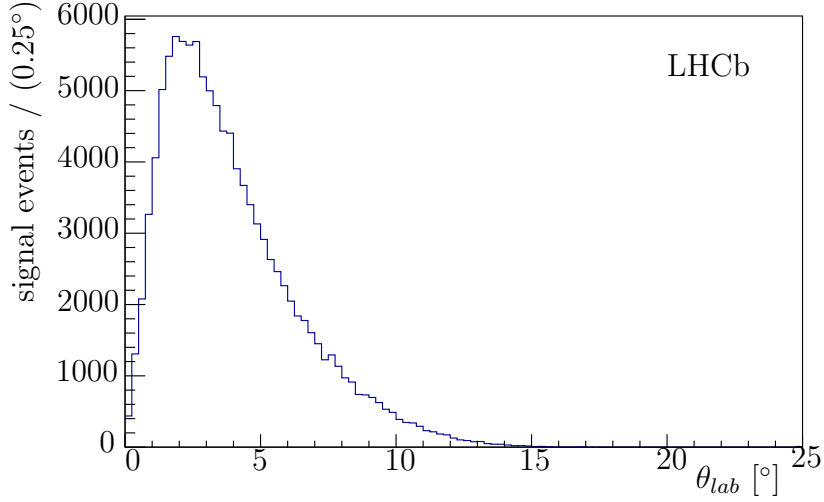
Figure 12: Distribution of the azimuthal angle of the B^0 meson with respect to the LHCb z-axis.

difference Δm_d is Gaussian constrained as well and not fixed, since $\mathcal{R}e(z)$ scales with $1 - \cos(\Delta mt)$ and such a term will stretch the asymmetry, resulting in a fake higher Δm_d. The PDG value for Δm_d is dominated by BaBar and Belle and therefore CPT violation is negligible here compared to LHCb for the same reason as for $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$. Parameters from acceptance, β_τ and resolution are fixed, since they are expected to have negligible systematic uncertainties. Last but not least, the $D_{J/\psi K_S^0}$ parameter has the value $D_{J/\psi K_S^0} = -\cos(2\beta) = -0.731 \pm 0.018$. This value has been computed with the PDG value of $\sin(2\beta) = 0.682 \pm 0.019$. The error can later be used to assign a (most likely negligible) systematic uncertainty. A list of Gaussian constrained variables with their Gaussian width is given in Table 16.

Parameter	Value
$C_{J/\psi K_S^0}$	0.005 ± 0.020
$S_{J/\psi K_S^0}$	0.676 ± 0.021
A_P^{2011}	-0.0066 ± 0.0026
A_P^{2012}	-0.0048 ± 0.0015
$\Delta m_d \ [\mathrm{ps}^{-1}]$	0.510 ± 0.003
p_0^{OS}	0.3815 ± 0.0011
p_1^{OS}	0.978 ± 0.012
Δp_0^{OS}	0.0148 ± 0.0016
Δp_1^{OS}	0.070 ± 0.018
p_0^{SS}	0.4232 ± 0.0029
p_1^{SS}	1.011 ± 0.064
Δp_0^{SS}	-0.0026 ± 0.0043
Δp_1^{SS}	-0.17 ± 0.10

Table 16: List of Gaussian constrained variables in the fit.

4.8 Fit results

The result from maximizing the likelihood of Eq. 73, results in the following values for $\Delta a_\mu^{B^0}$,

$$\Delta a_0^{B^0} - 0.38 \Delta a_Z^{B^0} = (-0.8 \pm 1.1 \ (stat.)) \times 10^{-15} \, \mathrm{GeV}$$
$$\Delta a_X^{B^0} = (-0.7 \pm 1.7 \ (stat.)) \times 10^{-15} \, \mathrm{GeV}$$
$$\Delta a_Y^{B^0} = (+0.6 \pm 1.6 \ (stat.)) \times 10^{-15} \, \mathrm{GeV}.$$

The likelihood fit also includes the results of the following parameters,

$$\tau = \ 1.511 \pm 0.011 \, \mathrm{ps}$$
$$\Delta m_d = \ 0.510 \pm 0.003 \, \mathrm{ps}^{-1} \ \ (constrained)$$
$$C_{J/\psi K_S^0} = -0.002 \pm 0.017 \ \ (constrained)$$
$$S_{J/\psi K_S^0} = \ 0.686 \pm 0.019 \ \ (constrained),$$

The correlations that are above 10% are,

$$\rho(C_{J/\psi K_S^0}, S_{J/\psi K_S^0}) = \ 0.14$$
$$\rho(\Delta a_0^{B^0} - 0.38 \Delta a_Z^{B^0}, S_{J/\psi K_S^0}) = -0.30.$$

What is immediately clear is that the result is consistent with no *CPT* and no Lorentz violation. This is the world's best measurement for $\Delta a_\mu^{B^0}$. Compared to the previous best

Parameter	This analysis	$\sin(2\beta)$ analysis [27]	cFit analysis
$C_{J/\psi K_S^0}$	-0.013 ± 0.036	-0.013 ± 0.027	-0.033 ± 0.032
$S_{J/\psi K_S^0}$	0.712 ± 0.041	0.711 ± 0.031	-0.729 ± 0.035

Table 17: Comparison of non-CPT violation fit results.

from BaBar [13],

$$\Delta a_0^{B^0} - 0.30 \Delta a_Z^{B^0}(\text{BaBar}) = (0.6 \pm 0.5) \times 10^{-12}\,\text{GeV}$$

$$\Delta a_X^{B^0}(\text{BaBar}) = (4.2 \pm 1.3) \times 10^{-12}\,\text{GeV}$$

$$\Delta a_Y^{B^0}(\text{BaBar}) = (2.6 \pm 2.5) \times 10^{-12}\,\text{GeV}$$

the new LHCb values are roughly three orders of magnitude more precise. Keep in mind that, since BaBar used flavour-specific decays, the values are measured with the factor $\Delta m/\Delta\Gamma$. For the modulo values reported here from BaBar, the Standard Model expectation value of $\Delta\Gamma_d = -0.0027 \pm 0.0007\,\text{ps}^{-1}$ is used [16].

As a cross-check for the fitter, $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$ have been fitted without CPT violation. The results are given in Table 17. It shows that the sWeighted fits agree with each other spot on. Notice the difference in uncertainty, they do not account for the background dilution. For the cFit, where they fit the background as well, there is still a discrepancy in central values and uncertainties with respect to the sWeighted fit. Despite that they are statistically compatible, this still has to be understood.

4.9 Systematics

Despite no quantitative study on the systematics has been done yet, an educated guess can be made to what the biggest contributors are. With the parallel analysis of $\sin(2\beta)$, where a quantitative systematics study has been done, an overview for this analysis can be made. For them, there are four not negligible systematics for $S_{J/\psi K_S^0}$: background tagging asymmetry, tagging calibration, decay width difference $\Delta\Gamma$ and kaon regeneration. For $C_{J/\psi K_S^0}$, background tagging asymmetry, tagging calibration and kaon regeneration are the significant ones. Background tagging asymmetry, as the name suggests, is caused by a tagging asymmetry in the background. Since they fit their background as well, where they assume no tagging asymmetry, this can cause the fit results to change for the signal as well. Since in this analysis, only the signal is fitted by subtraction with sWeights, it may not be an issue. Since the CPT violation parameters depend on the tag, the tagging calibration is assumed to be a significant source of systematics. The decay width difference causes the smallest significant contribution, but only for $S_{J/\psi K_S^0}$. Kaon regeneration is a material effect caused by $\overline{d}$ quarks annihilating with d quarks from neutrons and protons in the material. This causes an interference between kaon mixing and material. This effect is only visible in downstream tracks (DD). Regeneration effects

on tagged asymmetries are of the order of 0.5% in DD [33]. Since $\Delta a_0^{B^0} - 0.38 \Delta a_Z^{B^0}$ is for 30% correlated to $S_{J/\psi K_{\mathrm{S}}^0}$, a non-zero $\Delta\Gamma$ may be a systematic as well. As there is a scale uncertainty on the z-axis of the LHCb frame for the VELO, called z-scale, this may be effecting the decay time and is treated as a possible systematic. The parallel analysis did not find any significant contribution, so this analysis is likely not affected as well. Next to the systematics sources from the parallel analysis, there are most likely sources unique to the CPT analysis. Keeping the form of $\mathcal{R}e(z)$ in mind, the error on the momentum scale, the uncertainty on $D_{J/\psi K_{\mathrm{S}}^0}$ and the assumption of collimated B mesons may be a source. The source from the collimation is expected to be much bigger than any determination of LHCb in the sun-centered frame, therefore any contributions from the sun-centered frame parameters are neglected. The slight correlation between the momentum and the invariant mass, from which the sWeights are obtained, may also contribute to the systematics.

As for all the other possible systematics, this means that acceptance parameters, high lifetime correction factor β_τ and decay time resolution parameters do not contribute significantly to any systematics. Also a possible correlation between the mass and decay time is not significant. This also means that sWeights and decay time are not significantly correlated. This strengthens the claim that sWeights may be used. The list of discussed sources is summarized in Table. 18.

Systematics origin	Significance estimate
Tagging calibration	definite
Background tagging asymm.	possible
Mass decay-time correlation	-
z-scale	-
Production asymm.	-
Acceptance	-
Decay time resolution	-
Decay width difference	possible
Kaon regeneration	possible
Specific to $\Delta a_\mu^{B^0}$	
Lower momentum for background	possible
Momentum-sWeight correlation	possible
Momentum scale	possible
$D_{J/\psi K_{\mathrm{S}}^0}$ uncertainty	possible
SCF parameters	-

Table 18: List of possible contributors to the systematics.

5 Periodogram

Since the likelihood fit uses one specific frequency, *i.e.* the sidereal frequency, a complementary analysis is introduced that scans a wide band of frequencies. Generalization of the time-dependent analysis is achieved using Fourier analysis. A type of Fourier analysis often used in observational astronomy is the periodogram method and it is used as an estimator of the spectral power of a time-dependent signal. The spectral power distribution determines the amount that a specific frequency contributes, in amplitude, to the signal.

The periodogram is given by a discrete Fourier transform of the data, where the time domain is given, in our case, by the GPS time stamps of the data and the signal to be transformed is the asymmetry A_{CPT} from Eq. 41. A similar study has been done by the BaBar collaboration in their analysis of CPT violation in B^0 mixing [13]. The discrete Fourier transform is given by,

$$
\begin{aligned}
P(\nu) &= \frac{1}{N\sigma_x^2} \left| \sum_{j=1}^{N} x_j e^{2i\pi\nu t_j} \right|^2 \\
&= \frac{1}{N\sigma_x^2} \left[\left| \sum_{j=1}^{N} x_j \cos 2\pi\nu t_j \right|^2 + \left| \sum_{j=1}^{N} x_j \sin 2\pi\nu t_j \right|^2 \right]
\end{aligned}
\tag{75}
$$

where the x_j is the time-dependent variable provided by A_{CPT} at a given time t_j. Including dilution from flavour tagging, x_j is defined by,

$$
x_j \equiv d_j(1 - 2\omega_j)(1 - \cos(\Delta m \hat{t}_j)),
\tag{76}
$$

where d_j is the tag of the B^0 meson, ω_j the mistag probability, $\hat{t}_j$ the decay time and Δm is the mass difference of B^0. To add additional sensetivity to $\mathcal{R}e(z)$, the weight $1 - \cos(\Delta m \hat{t}_j)$ has been introduced. This definition of x ensures that the value fluctuates around zero. Small offsets like production or tagging asymmetry are removed by subtracting the mean of x_j from x_j in all periodograms mentioned after this.

Since the data set contains backgound events, it is desirable to use sWeights from the mass fit in all summations over the data to subtract the background. The use of sWeights in a periodogram is discussed in appendix B.1. In conclusion, the sWeights cannot be used due to a correlation between x_j and the sWeights and failure of the sWeights to properly subtract the background. Since background events are not expected to contribute to any possible signal and the periodogram statistics discussed in section 5.2 deals with random noise, the background contribution is accounted for.

The discreteness of the sampling of the data results in an important phenomenon. The function $P(\nu)$ is quite noisy. Even with increasing data size the variance of the periodogram

does not go down, but the signal to background ratio increases. This makes $P(\nu)$ not a statistically correct estimator of the spectral power, but due to the increasing signal to background ratio it is still a useful tool to determine significance of frequencies in signals.

In the coming sections a new version of the periodogram is introduced that has better statistical properties, after which the statistics of the periodogram is discussed. Then, a Monte Carlo study is done to determine the behaviour of noise and finally the periodogram of the data will be presented, including a significance study of the peaks using a Monte Carlo study.

5.1   Lomb-Scargle periodogram

Due to a better statistical behaviour, the Lomb-Scargle periodogram [34] is introduced by the definition,

$$P_{LS}(\nu) = \frac{1}{2\sigma_x^2} \left[\frac{\left[\sum_{j=1}^{N} x_j \cos 2\pi\nu(t_j - \tau) \right]^2}{\sum_{j=1}^{N} \cos^2 2\pi\nu(t_j - \tau)} + \frac{\left[\sum_{j=1}^{N} x_j \sin 2\pi\nu(t_j - \tau) \right]^2}{\sum_{j=1}^{N} \sin^2 2\pi\nu(t_j - \tau)} \right], \quad (77)$$

where the translation parameter τ is defined by,

$$\tan(4\pi\nu\tau) = \frac{\sum_{j=1}^{N} \sin(4\pi\nu t_j)}{\sum_{j=1}^{N} \cos(4\pi\nu t_j)}. \quad (78)$$

Although Eq. 77 looks different from Eq. 75, the change in outcome is little, but the different normalisation per real and imaginary part results in a mathematically easier discription of its statistics, which will be discussed in section 5.2.

This redefinition of the periodogram comes with a downside. In a periodogram, the phase of the signal should not matter, since the amplitude of the signal is the only thing that is to be extracted. The simpler definition of the periodogram had the nice and necessary property of time-origin invariance. Time translations, $t_j \rightarrow t_j + T_0$, resulted in a shift of $e^{i\omega T_0}$ and did not result in an amplitude change. With the rescaling of the real and imaginary part in the new defintion, this property was lost. Therefore the translation parameter τ is introduced and due to the property that $\tau \rightarrow \tau + T_0$ the shift gets cancelled in the term $t_j - \tau$. In addition, the τ translation parameter does not alter the statistical properties. In this analysis, only the Lomb-Scargle periodogram is used.

5.2 Periodogram statistics

To interpret the peaks in the periodogram and due to the noisiness of the discrete Fourier transform, a good description of its statistical behaviour is necessary.

The time-dependent variable x, Eq. 76, is based on the binning of A_{CPT} in time. A_{CPT} is a sufficiently large sum of independent random variables, $i.e.$ flavour tags, and therefore the time-binned A_{CPT} has to be Gaussian distributed according to the central limit theorem, if there is no dependence on time. Thus it is expected that the periodogram of a no-signal data set will behave according to a Gaussian distributed variable.

For deriving the statistical no-signal behaviour, a Gaussian distribution with zero mean and constant variance σ^2 is assumed for x. Following the derivation from Ref. [35], the general periodogram function $P(\nu)$ can be rewritten as,

$$P(\nu) = \frac{1}{2} \left[C^2(\nu) + S^2(\nu) \right], \tag{79}$$

where,

$$C(\nu) = A \sum_{j=1}^{N} x_j \cos(2\pi\nu(t_j - \tau)) \qquad S(\nu) = B \sum_{j=1}^{N} x_j \sin(2\pi\nu(t_j - \tau)). \tag{80}$$

With the normally distributed x_j and the zero mean, the variance of $C(\nu)$ and $S(\nu)$ is given by,

$$\sigma_C^2 = A^2 \langle x_j^2 \rangle \sum_{j=1}^{N} \cos^2(2\pi\nu(t_j - \tau)) \qquad \sigma_S^2 = B^2 \langle x_j^2 \rangle \sum_{j=1}^{N} \sin^2(2\pi\nu(t_j - \tau)), \tag{81}$$

where independence between events has been assumed, causing the expectation value of the cross terms in the square to vanish. To give $C(\nu)$ and $S(\nu)$ the variance σ_x^2, A and B are set to follow the Lomb-Scargle definition of the periodogram by defining,

$$A = \frac{1}{\sqrt{\sum\limits_{j=1}^{N} \cos^2(2\pi\nu(t_j - \tau))}} \qquad B = \frac{1}{\sqrt{\sum\limits_{j=1}^{N} \sin^2(2\pi\nu(t_j - \tau))}}. \tag{82}$$

From statistics it is known that the sum z of two squared variables with Gaussian distributions with equal width and zero mean is exponentially distributed,

$$p_z dz = \frac{1}{\sigma^2} e^{-z/\sigma^2} dz, \tag{83}$$

where σ is the shared width of the Gaussians and $p_z dz$ is the probability that z lies between z and $z + dz$. For $P(\nu)$ this means that it is distributed exponentially if its normalized

by the variance of the data, σ_x^2, and has no signal. A deviation in the distribution of $P(\nu)$ from an exponential distribution, *i.e.* much higher peaks, suggests presence of signal. It is instructive to see that the expectation value of $P(\nu)$ is the width of the Gaussian distribution of x_j, if $P(\nu)$ contains no signal, by filling in Eq. 92 with Eq. 81 and Eq. 82,

$$\langle P(\nu) \rangle = \frac{1}{2} \left[\langle C^2(\nu) \rangle + \langle S^2(\nu) \rangle \right] = \frac{1}{2} \left[\sigma^2 + \sigma^2 \right] = \sigma^2. \tag{84}$$

From the periodogram-peak distribution from Eq. 83, a peak significance can be derived. In case of a evenly-spaced time sampling, a so called false-alarm probability function is constructed. This describes how likely it is for a peak to be consistent with no signal. Essentially it is a cumulative distribution function of peak maxima in periodograms. Scargle [34] introduced the following false-alarm probability function,

$$\Pr[Z_{\max} > z]_{\text{Scargle}} = 1 - [1 - e^{-z}]^{N_i}, \tag{85}$$

where $Z_{\max}$ is the periodogram peak maximum. N_i is the number of mutually independent frequencies. In case of evenly spaced data there is well-defined set of frequencies that produce a set of independent random peaks. In case of evenly-spaced data with Gaussian noise, N_i is a well-defined number, given by,

$$N_i = \frac{N}{2}, \tag{86}$$

where N is the number of events. In case of unevenly-spaced time sampling and a non-Gaussian x_j, only sums of x_j behave Gaussian. Therefore, this relation does not hold and a different way of determining the false-alarm probability is needed. Nonetheless, an effective N_i can still be used to estimate the set of frequencies to be used in the periodogram.

5.3   Monte Carlo noise

With the use of Monte Carlo (MC) data, it is possible to construct a false-alarm probability function numerically, allowing to omit the analytical construction of such a function. The MC data is constructed by using the GPS time stamps to match the used time sampling. The variable x_j is determined for the MC data set using a random number generator to sample from a Gaussian distribution with zero mean. A test to confirm that a Gaussian distribution for no signal is correct is given in appendix B.2.

The range of frequencies is set around the mean solar day. Since the time stamps are in UTC and therefore close to UT1, the mean solar day is set at $\nu = 1$ with the GPS time stamps in days of 86400 UT1 seconds. The most interesting frequency, *i.e.* the sidereal frequency, is at $\nu = 1.00274$. The lowest frequency is set at $\nu_{\text{low}} = 0.03$, or roughly a period of month, since lower frequencies will possibly be effected by the periods at which the data is taken and the gaps in between. Similar to the analysis of BaBar [13], the upper limit is set at $\nu_{\text{high}} = 2.10$, to scan a wide range of frequencies around the sidereal frequency.

$\Delta\nu$	$4 \cdot 10^{-4}$ solar-day^{-1}
ν_{low}	0.03 solar-day^{-1}
ν_{high}	2.10 solar-day^{-1}
N_{freqs}	5200
N_i	≈ 2600

Table 19: Set of frequencies for evaluating the used periodograms.

From counting the local maxima and minima in the periodogram, the amount of effective independent frequencies is estimated at roughly 2600. As a cross-check, the effective N_i will be determined later on with a χ^2-fit on the MC constructed false-alarm probability function. With the estimate of N_i, the frequency spacing is set on $\Delta\nu = 4 \cdot 10^{-4}$ solar-day^{-1}. This sets the number of frequencies to 5200 and results in an oversampling of roughly a factor 2 and therefore avoids undersampling. The properties of the sampling frequencies are summarized in Table 19.

As an example the distribution of an MC Lomb-Scargle periodogram is given in Fig 13, where an exponential has been fitted through using χ^2 minimization. With a close look at the highest peak entries, the oversampling of a factor 2 can be seen. From the χ^2/ndf it is clear that the distribution behaves according to the predicted statistics and therefore the sampling frequencies are well chosen. By generating 1000 MC Lomb-Scargle periodogram
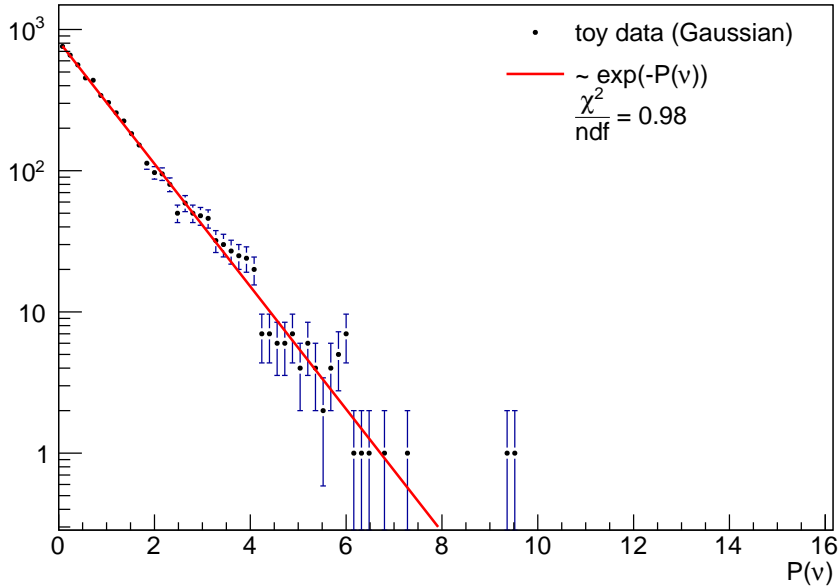
Figure 13: Peak distribution of a Lomb-Scargle periodogram of Gaussian Monte Carlo data. Includes a χ^2-fit of an exponential.
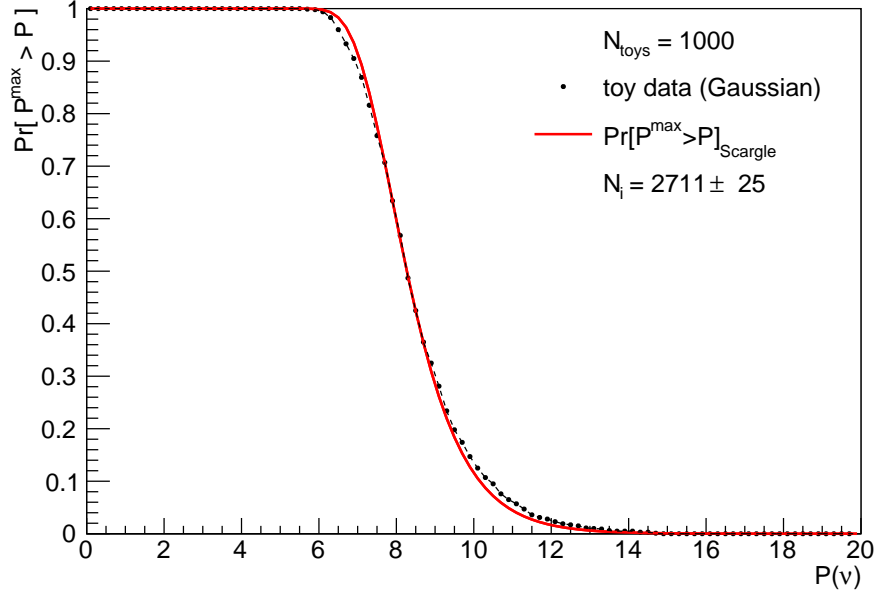
Figure 14: Numerical false-alarm probability function of 1000 Lomb-Scargle periodograms of Gaussian MC data. Including Scargle false-alarm probability function χ^2-fit with resulting N_i and spline interpolating function for the MC data.

toys with a Gaussian distributed variable, the needed false-alarm probability function is obtained and depicted in Figure 14. It is constructed from the cumulative distribution function of the peak maxima of these toys.

To compare with Eq. 85, a χ^2-fit has been performed, giving an estimate of N_i. To determine the numerical function, spline interpolation is used. With a close look at the highest peak entries, the oversampling of a factor 2 can be seen. Since the χ^2/ndf is good, this does not cause any problem. From the χ^2-fit, the effective number of independent frequencies is estimated at $N_i = 2711 \pm 25$. Therefore, the used frequency spacing is correct. The numerical false-alarm probablity function visibly deviates from the analytical Scargle function and indeed suggests that an unevenly spaced data set behaves different from an evenly spaced one.

5.4   Periodogram of $3\,\mathrm{fb}^{-1}$ data set

With the definition of the periodogram and its derived statistics, the Lomb-Scargle periodogram of the full $3\,\mathrm{fb}^{-1}$ data set is given in Fig. 15. A zoom into the interesting range around the solar frequency is given in Fig. 16, which has been achieved by highly oversampling the frequencies in this range, purely used for plotting. From the zoomed in periodogram, Fig. 16, it can be clearly seen that there are so called independent frequencies, which are at the peak minima and maxima. Sampling between those frequencies behaves
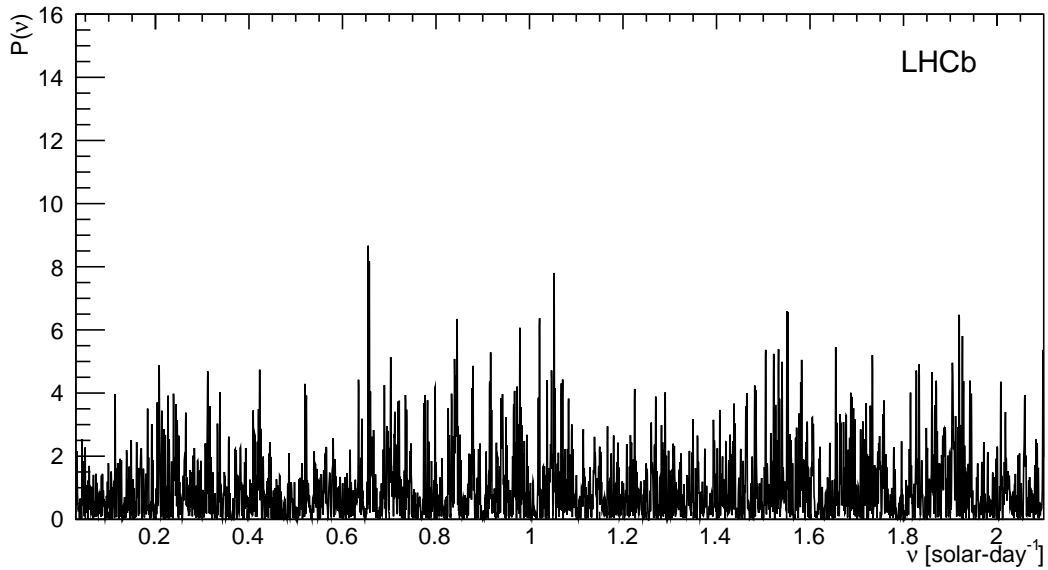
Figure 15: Lomb-Scargle periodogram of $3\,\mathrm{fb}^{-1}$ data set.
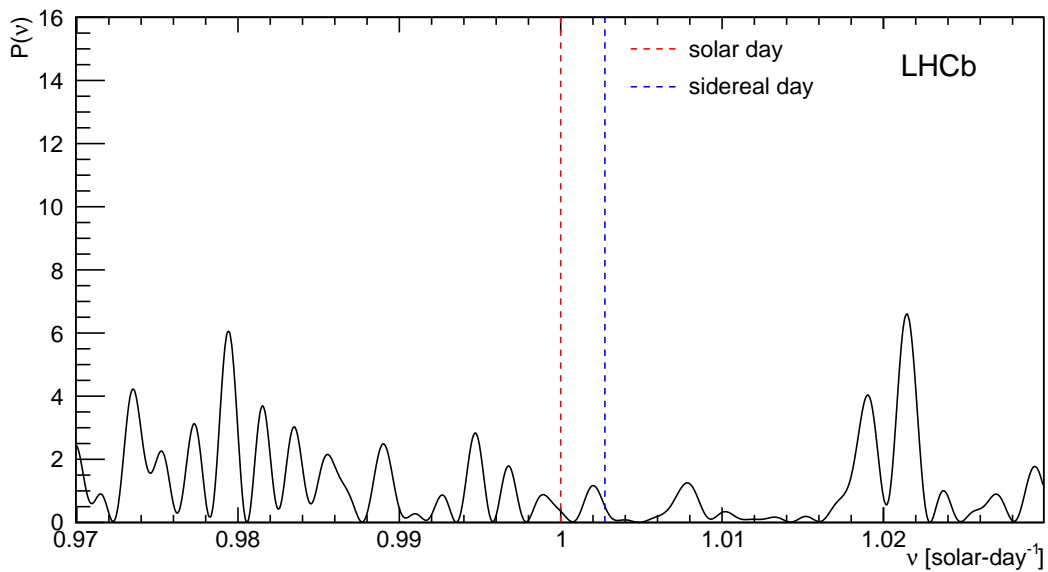

Figure 16: Lomb-Scargle periodogram of $3\,\mathrm{fb}^{-1}$ data set around period of one solar day. The smooth lines between the peaks (local minima and maxima) are used by highly oversampling the frequencies.

as a kind of interpolation between the peaks and therefore are not independent. Since one clearly sees the independent frequencies, the total amount of independent frequencies in
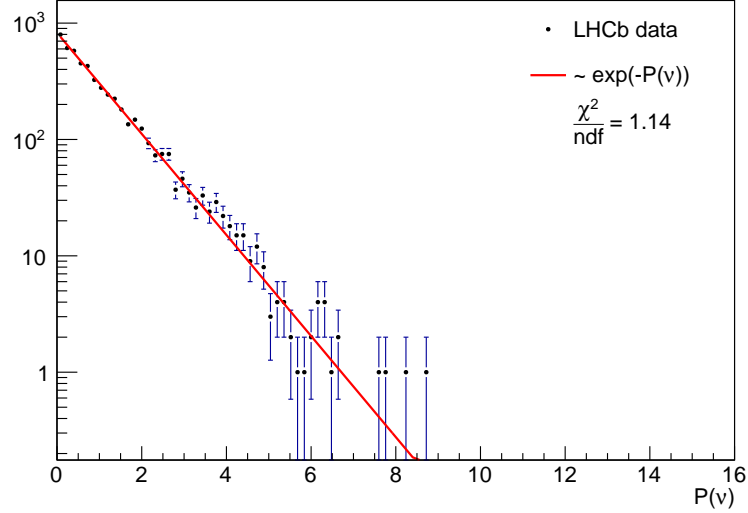
Figure 17: Peak distribution of a Lomb-Scargle periodogram of $3\,\mathrm{fb}^{-1}$ data set.

this frequency window can be estimated by extrapolation. This method was used in the estimation of the indepedent frequencies. The distribution of the peaks is given in Fig. 17. From this, combined with a χ^2-fit of an exponential, it is clear that the periodogram is consistent with no signal.

5.4.1 Periodogram peak significance

Using the Lomb-Scargle periodogram peaks, shown in Fig. 15, the peak maximum of the data is $P_{\mathrm{max}}^{\mathrm{data}} = 8.67$ at $\nu_{\mathrm{max}} = 0.6550$ solar-day^{-1}. This maximum has a false-alarm probability of $\mathrm{Pr}[P_{\mathrm{max}} > P_{\mathrm{max}}^{\mathrm{data}}] = 0.37$, according to the numerical false-alarm probability, depicted in Fig. 14. For the sidereal frequency, the peak height is $P_{\mathrm{sid}} = 0.51$. For the solar frequency, $P_{\mathrm{sol}} = 0.35$. In the periodogram, consisting of 5200 peaks, there are $N_{\mathrm{sid}} = 3112$ peaks higher than the sidereal peak and for the solar peak, there are $N_{\mathrm{sol}} = 3657$ higher peaks.

To conclude the periodogram analysis, the $3\,\mathrm{fb}^{-1}$ LHCb data set is consistent with no time-dependent CPT violation in B^0-mixing in the decay $B^0 \rightarrow J/\psi\, K_{\mathrm{s}}^0$.'

6   Conclusion

CPT and Lorentz symmetry are important fundamental symmetries in the Standard Model. To parametrize low-energy effects of *CPT* and Lorentz violation in the Standard Model, an effective field theory has been used called the Standard Model Extension. LHCb has the possibility to improve measurements of the Standard Model Extension parameter Δa_μ for the D^0, B^0 and B_s^0 meson systems. One of the major contributions to the relative high sensitivity of LHCb is the high boost of the neutral mesons. When comparing B^0 and B_s^0, B^0 has the highest sensitivity due to the smaller mass difference. Therefore, in this thesis, Δa_μ is measured in B^0, $\Delta a_\mu^{B^0}$. The decay channel used in the analysis is $B^0 \to J/\psi\, K_{\mathrm{S}}^0$, which has a final state that is a *CP* eigenstate. For these *CP* eigenstates, larger *CPT* violating effects are expected compared to flavour-specific final states.

To obtain $\Delta a_\mu^{B^0}$ from the LHCb data set corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$, a maximum-likelihood fit is performed on the time- and tag-dependent decay rate distributions of $B^0 \to J/\psi\, K_{\mathrm{S}}^0$. Using the assumption of along-the-beam collimated B^0 mesons, the fit resulted in,

$$\Delta a_0^{B^0} - 0.38\Delta a_Z^{B^0} = (-0.8 \pm 1.1\ (stat.)) \times 10^{-15}\,\mathrm{GeV},$$
$$\Delta a_X^{B^0} = (-0.7 \pm 1.7\ (stat.)) \times 10^{-15}\,\mathrm{GeV},$$
$$\Delta a_Y^{B^0} = (+0.6 \pm 1.6\ (stat.)) \times 10^{-15}\,\mathrm{GeV},$$

which is consistent with no *CPT* and no Lorentz violation. In addition, the periodogram method is used to generalize the search for time-dependent Lorentz violation by scanning a broad range of frequencies. By using Monte Carlo techniques to derive the statistics of the periodogram, the data is determined to be compatible with no time dependence, since the maximum peak in the periodogram has a false-alarm probability of 0.37. Systematics to $\Delta a_\mu^{B^0}$ still has to be determined quantitatively, but likely sources have been identified. Contributions are expected to come from the flavour tagging calibration, kaon regeneration, non-zero decay width difference, non-collimatedness of the B^0 mesons and a possible sWeight-momentum correlation. Since these systematic uncertainties are relatively small scale-uncertainties, it is expected that the statistical uncertainties dominate the systematic uncertainties.

The measurement presented here of $\Delta a_\mu^{B^0}$ is the world's best. Under assumption of the Standard Model prediction $\Delta\Gamma_d = -0.0027$, it is an improvement in precision of three orders of magnitude with respect to the previous best measurement of BaBar.

7   Acknowledgements

Appendices

A Rotation matrix in the SCF

As a reminder, the definition of the sun-centered frame (SCF), depicted in Fig. 7, is given by the unit vectors $\{\hat{X}, \hat{Y}, \hat{Z}\}$, where $\hat{Z}$ points along the rotational axis of the earth and $\hat{X}$ points from the sun to the vernal equinox at the epoch J2000. $\hat{Y}$ complements $\hat{X}$ and $\hat{Z}$ to form a right-handed coordinate system.

To rotate from the SCF to the LHCb lab frame, two types of rotations are considered: extrinsic and intrinsic. An extrinsic rotation is a rotation over one of the fixed axes of the original coordinate system. An intrinsic rotation is a rotation over the rotating coordinate system, *i.e.* the original coordinate system may be reorientated due to previous rotations, resulting in a rotating coordinate system. The three extrinsic rotations are in accordance with proper Euler angles. The rotations are performed such that after the rotations $\{\hat{X}, \hat{Y}, \hat{Z}\}$ corresponds to $\{\hat{z}, \hat{x}, \hat{y}\}$ from the LHCb lab frame respectively. The rotations are all right handed. First there is the latitude of the experiment, which rotates over the $\hat{Y}$-axis with an angle $-\lambda - \pi/2$, where λ is the latitude. Secondly the (new) $\hat{Z}'$-axis is rotated to the correct azimuth of the experiment, with an angle $-\theta + \pi$, where θ is the azimuth. The third angle is the tilt of the experiment. The beam of the experiment lies at an angle of $-\alpha$ with respect to the geodetic horizon. This angle rotates the (new) $\hat{X}''$-axis to the correct tilt. At last, there is the sidereal phase, which can be seen as an effective longitude with respect to the vernal equinox. This is a rotation over the $\hat{Z}$-axis and not the (new) $\hat{Z}'''$-axis, making it an extrinsic rotation. The last angle, the sidereal phase, is performed using a standard rotation matrix over the $\hat{Z}$ axis, since it is an extrinsic rotation. First the following terms are defined,

$$\begin{aligned}
\cos(-\lambda + \pi/2) = \sin(\lambda) \equiv s_1 \quad &\cos(-\theta + \pi) = -\cos(\theta) \equiv -c_2 \quad &\cos(-\alpha) \equiv c_3, \\
\sin(-\lambda + \pi/2) = \cos(\lambda) \equiv c_1 \quad &\sin(-\theta + \pi) = \sin(\theta) \equiv s_2 \quad &\sin(-\alpha) \equiv -s_3,
\end{aligned} \quad (87)$$

which allow us to write the matrix R, provided by the proper Euler angles formalism as,

$$R(\lambda, \theta, \Omega t = 0, \alpha) = \begin{pmatrix} -s_1 c_2 c_3 + c_1 s_3 & -s_1 s_2 & c_3 c_1 + s_1 c_2 s_3 \\ c_3 s_2 & -c_2 & -s_2 s_3 \\ s_1 s_3 + c_2 c_3 c_1 & c_1 s_2 & s_1 c_3 - c_2 c_1 s_3 \end{pmatrix}. \quad (88)$$

The last rotation, over the $\hat{Z}$-axis, is applied to get the correct sidereal phase,

$$R(\lambda, \theta, \Omega t, \alpha) = \begin{pmatrix} \cos(\Omega t) & -\sin(\Omega t) & 0 \\ \sin(\Omega t) & \cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -s_1 c_2 c_3 + c_1 s_3 & -s_1 s_2 & c_3 c_1 + s_1 c_2 s_3 \\ c_3 s_2 & -c_2 & -s_2 s_3 \\ s_1 s_3 + c_2 c_3 c_1 & c_1 s_2 & s_1 c_3 - c_2 c_1 s_3 \end{pmatrix}$$

$$\tag{89}$$

$$= \begin{pmatrix} -c_t(s_1 c_2 c_3 - c_1 s_3) - s_t c_3 s_2 & -c_t s_1 s_2 + s_t c_2 & c_t(c_3 c_1 + s_1 c_2 s_3) + s_t s_2 s_3 \\ -s_t(s_1 c_2 c_3 - c_1 s_3) + c_t c_3 s_2 & -s_t s_1 s_2 - c_t c_2 & s_t(c_3 c_1 + s_1 c_2 s_3) - c_t s_2 s_3 \\ s_1 s_3 + c_2 c_3 c_1 & c_1 s_2 & s_1 c_3 - c_2 c_1 s_3 \end{pmatrix}.$$

$$\tag{90}$$

In order to retain Eq. 42 used in the main text, one only needs to switch columns to get the correspondence $\{\hat{x}, \hat{y}, \hat{z}\} = \{R\hat{X}, R\hat{Y}, R\hat{Z}\}$.

B Periodogram statistics crosschecks

B.1 Periodogram statistics with sWeights

For background subtraction, sWeights can be used and to study the effect of sWeights on a periodogram, the periodogram statistics derivation is repeated including sWeights provided by the mass fit. As a reminder, general characteristics of sWeights w_j are,

$$\sum_j^{N_s} w_j = N_s \langle w_j \rangle_s = N_s \qquad\qquad \sum_j^{N_b} w_j = N_b \langle w_j \rangle_b = 0, \tag{91}$$

where N_s and N_b are the number of signal and background events and the sums are over the signal and background events respectively. The general periodogram function $P(\nu)$ can be rewritten as,

$$P(\nu) = \frac{1}{2} \left[C^2(\nu) + S^2(\nu) \right], \tag{92}$$

where,

$$C(\nu) = A \sum_{j=1}^{N} w_j x_j \cos(2\pi\nu(t_j - \tau)) \qquad S(\nu) = B \sum_{j=1}^{N} w_j x_j \sin(2\pi\nu(t_j - \tau)). \tag{93}$$

If variables X and Y are independent, *i.e.* not correlated, $\langle XY \rangle = \langle X \rangle \langle Y \rangle$. Therefore the mean of $C(\nu)$ and $S(\nu)$ for the signal events are,

$$\langle C(\nu) \rangle_s = A \sum_{j=1}^{N_s} \langle w_j x_j \rangle_s \cos(2\pi\nu(t_j - \tau)) \tag{94}$$

$$= A \langle w_j \rangle_s \langle x_j \rangle_s \sum_{j=1}^{N_s} \cos(2\pi\nu(t_j - \tau))$$

$$= A \langle x_j \rangle_s \sum_{j=1}^{N_s} \cos(2\pi\nu(t_j - \tau)),$$

and for the background,

$$\langle C(\nu) \rangle_b = A \langle w_j \rangle_b \langle x_j \rangle_b \sum_{j=1}^{N_b} \cos(2\pi\nu(t_j - \tau)) = 0. \tag{95}$$

From this it clear that the background cancels. In addition, $\langle C(\nu) \rangle = 0$ if x_j and w_j are independent and $\langle x_j \rangle_s = 0$. With similar definitions, it is trivial that this holds for $\langle S(\nu) \rangle$ as well. With zero mean, the variance of $C(\nu)$ is given by,

$$\langle C^2(\nu) \rangle = \left\langle \left[A \sum_{j=1}^{N} w_j x_j \cos(2\pi\nu(t_j - \tau)) \right] \left[A \sum_{k=1}^{N} w_k x_k \cos(2\pi\nu(t_k - \tau)) \right] \right\rangle, \tag{96}$$

$$= A^2 \sum_{j=1}^{N} \sum_{k=1}^{N} \langle w_j x_j w_k x_k \rangle \cos(2\pi\nu(t_j - \tau)) \cos(2\pi\nu(t_k - \tau)),$$

and because each event is independent from the other events, the cross terms $(j \neq k)$ vanish and result in,

$$\langle C^2(\nu) \rangle = A^2 \langle w_j^2 x_j^2 \rangle \sum_{j=1}^{N} \cos^2(2\pi\nu(t_j - \tau)) \tag{97}$$

$$= A^2 \langle w_j^2 \rangle \langle x_j^2 \rangle \sum_{j=1}^{N} \cos^2(2\pi\nu(t_j - \tau)).$$

To go any further, the behaviour of w_j^2 has to be studied. It is intuitive to convert all sums to sWeighted sums, therefore A and B are set to,

$$A = \frac{1}{\sqrt{\sum_j^N w_j \cos^2(2\pi\nu(t_j - \tau))}} \qquad B = \frac{1}{\sqrt{\sum_j^N w_j \sin^2(2\pi\nu(t_j - \tau))}}, \tag{98}$$

56

and therefore the following equality, $\langle w_j^2 \rangle = \langle w_j \rangle$, is needed. To achieve this, a normalization of the sWeights is introduced by substituting $w_j \to \tilde{w}_j = \alpha w_j$, resulting in,

$$N\langle \tilde{w}_j^2 \rangle = \alpha^2 \sum_j^N w_j^2 \qquad\qquad N\langle \tilde{w}_j \rangle = \alpha \sum_j^N w_j. \tag{99}$$

Therefore, the normalisation set by,

$$\alpha = \sum_j^N w_j / \sum_j^N w_j^2, \tag{100}$$

will achieve the correct variance of $C^2(\nu)$. For the background, the variance is similar,

$$\langle C^2(\nu) \rangle_b = A^2 \langle w_j^2 \rangle_b \langle x_j^2 \rangle_b \sum_{j=1}^{N_b} \cos^2(2\pi\nu(t_j - \tau)), \tag{101}$$

but since $w_j \neq 0$ for all j in the background, $\langle w_j^2 \rangle_b \neq 0$. Despite the correct total variance, the background does not cancel in $\langle P(\nu) \rangle$.

Although the required statistical properties can be met with x_j and w_j uncorrelated and a zero mean for x_j, the background does not cancel and therefore it is not desirable to use sWeights in a periodogram. In conclusion, sWeights in a periodogram are not used.

B.2   Periodogram of data with random tag

As mentioned in section 5, with the use of the central limit theorem, the variable from Eq. 76,

$$x_j = d_j(1 - 2\omega_j)(1 - \cos(\Delta m \hat{t}_j)),$$

is assumed to behave, in case of no signal, like a Gaussian variable in a periodogram. Although the signal periodogram distribution, depicted in Fig. 17, seems to behave like a Gaussian variable, a more robust test can be done. To achieve this test, a Monte Carlo (MC) study with 1000 toys is done with the tag d_j randomized, where each tag has equal probability. The $3\,\mathrm{fb}^{-1}$ data set provides the mistag probability ω_j and the decay time $\hat{t}_j$. An example distribution of one of the toys is given in Fig. 19 and clearly shows the correct exponential distribution, according to the χ^2-fit. To compare all toys, the peak maxima of the random tag and Gaussians toys are plotted in Fig. 19. The distributions of both toys clearly agree. There is no visible deviation between the two, considering the error on the toy data. To conclude, the assumption that the periodogram behaves as if a Gaussian variable has been used, is justified.
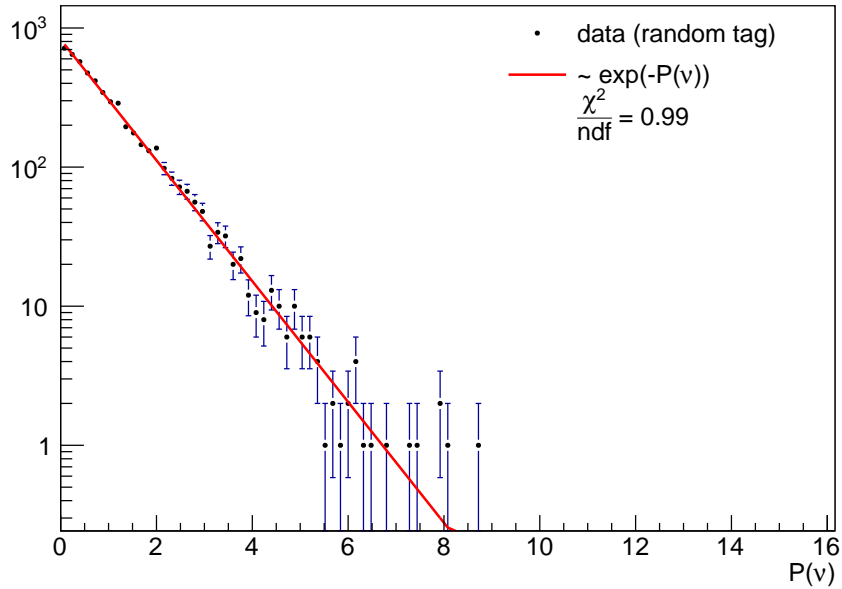
Figure 18: Peak distribution of a Lomb-Scargle periodogram with random tag.
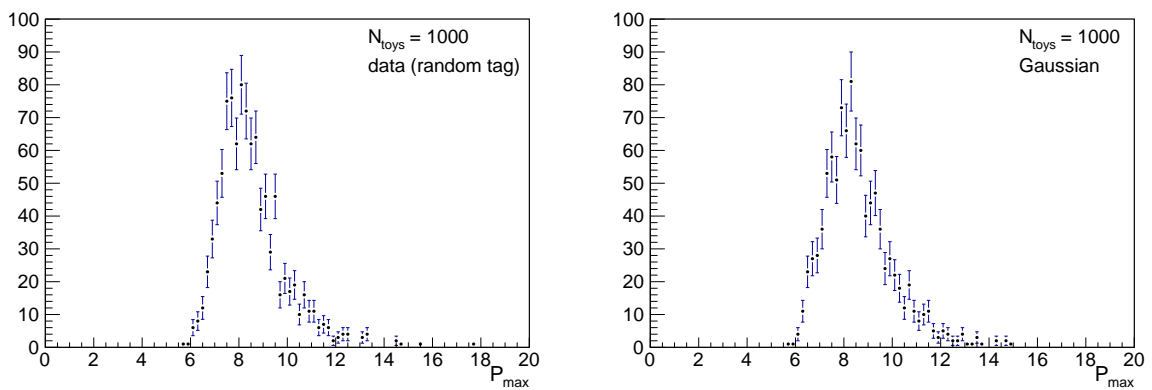

Figure 19: Peak maxima distributions of MC toys of 1000 Lomb-Scargle periodograms each, one with random tag (left), other with Gaussian distributed (right) variable.

References

[1] O. Bertolami, D. Colladay, V. A. Kostelecky, and R. Potting, *CPT violation and baryogenesis*, Phys. Lett. **B395** (1997) 178, `arXiv:hep-ph/9612437`.

[2] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[3] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[4] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[5] O. Greenberg, *CPT violation implies violation of Lorentz invariance*, Phys. Rev. Lett. **89** (2002) 231602, `arXiv:hep-ph/0201258`.

[6] BaBar Collaboration, B. Aubert *et al.*, *Limits on the decay rate difference of neutral$-B$ mesons and on CP, T, and CPT violation in $B^0\overline{B}^0$ oscillations*, Phys. Rev. **D70** (2004) 012007, `arXiv:hep-ex/0403002`.

[7] A. Lenz and U. Nierste, *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons*, `arXiv:1102.4274`.

[8] LHCb collaboration, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in $B^0 - \overline{B}^0$ mixing*, `arXiv:1409.8586`.

[9] Particle Data Group, K. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[10] V. A. Kosteleck and R. Potting, *{CPT} and strings*, Nuclear Physics B **359** (1991), no. 23 545 .

[11] D. Colladay and V. A. Kostelecky, *Lorentz violating extension of the standard model*, Phys. Rev. **D58** (1998) 116002, `arXiv:hep-ph/9809521`.

[12] V. A. Kostelecky, *Sensitivity of CPT tests with neutral mesons*, Phys. Rev. Lett. **80** (1998) 1818, `arXiv:hep-ph/9809572`.

[13] BaBar Collaboration, B. Aubert *et al.*, *Search for CPT and Lorentz violation in B0 - anti-B0 oscillations with dilepton events*, Phys. Rev. Lett. **100** (2008) 131802, `arXiv:0711.2713`.

[14] KLOE-2 Collaboration, D. Babusci *et al.*, *Test of CPT and Lorentz symmetry in entangled neutral kaons with the KLOE experiment*, Phys. Lett. **B730** (2014) 89, `arXiv:1312.6818`.

[15] J. van Tilburg and M. van Veghel, *Status and prospects for CPT and Lorentz invariance violation searches in neutral meson mixing*, Phys. Lett. **B742** (2015) 236, `arXiv:1407.1269`.

[16] A. Lenz and U. Nierste, *Theoretical update of $B_s - \bar{B}_s$ mixing*, JHEP **0706** (2007) 072, `arXiv:hep-ph/0612167`.

[17] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K^0_S$ decays*, Phys. Lett. **B721** (2013) 24, `arXiv:1211.6093`.

[18] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[19] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of early 2012*, `arXiv:1207.1158`, updated results and plots available at `http://www.slac.stanford.edu/xorg/hfag/`.

[20] S. Faller, M. Jung, R. Fleischer, and T. Mannel, *The Golden Modes B0 → J/psi K(S,L) in the Era of Precision Flavour Physics*, Phys. Rev. **D79** (2009) 014030, `arXiv:0809.0842`.

[21] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008), no. 08 S08001.

[22] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[23] V. A. Kostelecky and C. D. Lane, *Constraints on Lorentz violation from clock comparison experiments*, Phys. Rev. **D60** (1999) 116010, `arXiv:hep-ph/9908504`.

[24] R. Lindner, *Definition of coordinate system*, CERN, LHCb (2003), EDMS-372642.

[25] M. Jones, *Computation of WGS84 geodetic coordinates and azimuths at the LHC interaction points*, CERN (2005), EDMS-322747.

[26] S. Aoki *et al.*, *The new definition of universal time*, Astronomy and Astrophysics **105** (1982) 359.

[27] C. Cauet *et al.*, *Measurement of CP violation in the decay $B^0 \to J/\psi K^0_S$ with $3 fb^{-1}$ of LHCb data*, LHCb internal analysis note (2013).

[28] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, `arXiv:physics/0503191`.

[29] D. Martinez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, `arXiv:1312.5000`.

[30] M. Pivk and F. L. Diberder, *: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005), no. 12 356 .

[31] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, `arXiv:1407.0748`.

[32] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, ArXiv e-prints (2009) `arXiv:0905.0724`.

[33] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ decays*, JHEP **1407** (2014) 041, `arXiv:1405.2797`.

[34] J. D. Scargle, *Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data*, The Astrophysical Journal **263** (1982) 835.

[35] J. H. Horne and S. L. Baliunas, *A prescription for period analysis of unevenly sampled time series*, The Astrophysical Journal **302** (1986) 757.