Measurement of CP-violation in decays of strange beauty mesons

Lukashenko, Valeriia

2024

DOI (link to publisher)
10.5463/thesis.677

document version
Publisher's PDF, also known as Version of record

VRIJE UNIVERSITEIT

MEASUREMENT OF CP-VIOLATION IN DECAYS OF STRANGE BEAUTY MESONS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op dinsdag 28 mei 2024 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Valeriia Lukashenko

geboren te Kyiv, Oekraïne

promotor:	prof.dr. H.G. Raven
copromotor:	dr. W.D. Hulsbergen
promotiecommissie:	prof.dr. A. Freise prof.dr. M.P. Decowski dr.dr.-habil Y. Amhis dr. K.K. Vos dr. L. Brenner prof.dr. I.B. van Vulpen

Contents

1 Introduction

The Standard Model (SM) of elementary particles is a theory that describes all known fundamental particles and three out of the four known interactions: the electromagnetic, the strong, and the weak interaction [1]. Despite its overwhelming success in explaining all the experimental results obtained so far, it is not a complete theory: the gravitational force is not included; there is no natural explanation for the size of the masses of particles; there are no particles in the SM with properties matching the dark matter required to explain astronomical observations [2] and finally, the SM can not explain the abundance of matter in the Universe. This last question is the central topic of this thesis.

The SM predicts approximately the same amount of matter and anti-matter, but our everyday observation shows that almost everything around us, including ourselves, predominantly consists of matter: electrons, protons, and neutrons. In 1967, to explain this abundance of matter in the Universe, Sakharov formulated a list of conditions, required for a matter-abundant Universe generated with baryogenesis [3]:

1. baryon number violating processes;
2. interactions outside of the thermal equilibrium;
3. processes that violate $\mathcal{C}$- and $\mathcal{CP}$-symmetry[1].

Baryon number is predicted to be violated non-perturbatively in the SM by a chiral anomaly [4]. The second condition would require a first-order

[1] $\mathcal{CP}$ stands for the combination of parity transformation, which reverts the direction of spatial coordinates, and charge conjugation, which flips the sign of the quantum charges of particles.

electroweak phase transition, but more likely a continuous state crossover has happened [5]. The weak interaction is left-chiral, which means that it violates the charge conjugation symmetry because it only couples to the left-handed particles and right-handed anti-particles. In the SM, $\mathcal{CP}$-violation is present in the weak interaction.

However, the amount of the $\mathcal{CP}$-violation in the SM is by far too small to explain the observed baryon asymmetry η [2],

$$\eta = \frac{n_b - n_{\bar{b}}}{n_\gamma} \approx 10^{-10}, \tag{1.1}$$

where n_b is the number density of baryons; $n_{\bar{b}}$ is the number density of anti-baryons; n_γ is the number density of the photons. Hence, extra sources of $\mathcal{CP}$-violation are necessary to explain the matter abundance in the Universe. These sources have yet to be discovered.

Two approaches are used to look for new phenomena that possibly contribute to $\mathcal{CP}$-violation: searches at high energies and precision measurements of the SM. Additional heavier particles can be produced by colliding more luminous and more energetic beams. If these heavier particles decay within the detector body it is possible to observe them, which constitutes the first approach. The second approach is to measure known SM processes and parameters extremely accurately. If a new particle or force contributes to these processes, it will modify the observation with respect to the SM prediction.

In this thesis, the path of precision measurement is taken. The $\mathcal{CP}$-violating phase ϕ_s is used as a probe for new $\mathcal{CP}$-violating phenomena. The SM predicts ϕ_s to be close to zero. If a relatively large ϕ_s is observed, the difference could be attributed to unknown phenomena [6]. The $\mathcal{CP}$-violating phase ϕ_s is measured with $B_s^0 \to J/\psi\,\phi(1020)$ decays [7]. The $B_s^0 \to J/\psi\,\phi(1020)$ decay is dominated by transition amplitudes that are reasonably well known in the SM, which makes this decay a "golden" mode for the ϕ_s measurement.

This thesis uses a data sample of $B_s^0 \to J/\psi\,\phi(1020)$ decays collected between 2015 and 2018, known as Run 2, by LHCb [8], one of the four main experiments at the LHC. Chapter 2 describes the $B_s^0 \to J/\psi\,\phi(1020)$ decays theoretically. Chapter 3 describes the LHCb detector, which was used to collect the data used in this thesis. The analysis of the collected data sample is described in detail in Chapter 4. There the most recent LHCb measurement of ϕ_s is presented. Chapter 5 describes the preparation for future LHCb data-taking. Lastly, Chapter 6 discusses the future of the ϕ_s measurement.

2 $\mathcal{CP}$-violation and $B_s^0 \to J/\psi\, K^+K^-$ decays

The Standard Model describes visible matter particles and their known interactions (excluding gravity) [1]. The matter part of the SM consists of spin-1/2 fermions split into two groups: quarks and leptons, see Fig. 2.1. Both are organized into three groups, known as "generations". Their interactions are mediated by spin-1 particles called vector bosons

The strong interaction confines the quarks together into two- or three-quark-bound states known as hadrons. It is mediated by eight massless gluons, which couple to a quantum number called colour: red, green, or blue. The electromagnetic interaction is mediated by photons, which couple to the electric charge. The weak interaction is mediated by Z and $W^{\pm}$ bosons and couples to the weak isospin. The last particle of the SM is the Higgs boson, which is the only spin-0 (or scalar) boson in the SM. The Higgs field gives masses to the Z and $W^{\pm}$ bosons and masses to the matter particles via the Higgs mechanism.

2.1 Symmetries of the SM

The SM is a Lorentz-invariant theory and is symmetric under the $SU(3) \times SU(2) \times U(1)$ gauge groups, where $SU(3)$ corresponds to the strong interaction; $SU(2) \times U(1)$ corresponds to the electroweak interaction.

The Lorentz-invariance of the SM is linked to the discrete $\mathcal{CPT}$-symmetry. $\mathcal{CPT}$-symmetry stands for the symmetry under three simultaneous discrete transformations: charge conjugation $\mathcal{C}$, parity $\mathcal{P}$, and time $\mathcal{T}$. In the quantum field theory, fermions are described by the Dirac spinors $\psi(\vec{x}, t)$.

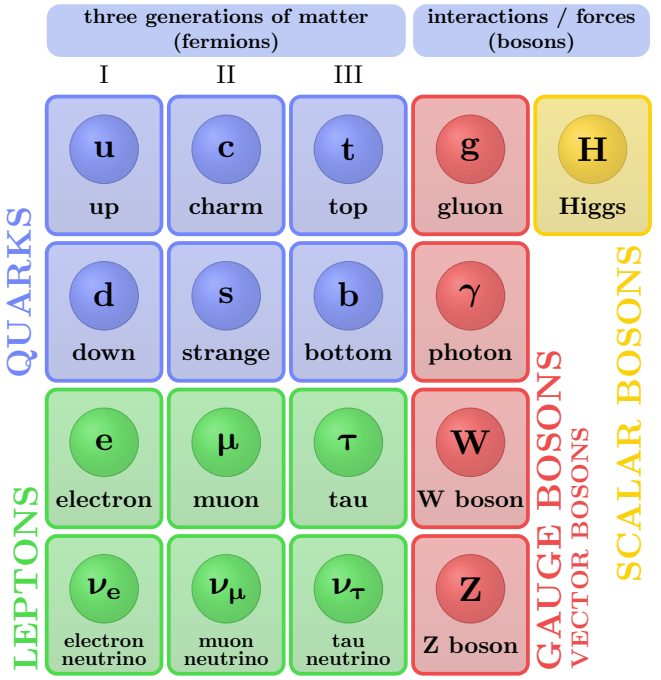


Figure 2.1: The content of the Standard Model.

For a fermion $\psi(\vec{x},t)$, the charge transformation $\mathcal{C}$ is defined as,

$$\psi(\vec{x},t) \xrightarrow{C} \hat{\mathcal{C}}\bar{\psi}^T(\vec{x},t) \equiv i\gamma^0\gamma^2\bar{\psi}^T(\vec{x},t), \tag{2.1}$$

where $\vec{x}$ are the spacial coordinates; t is the time coordinate; $\hat{\mathcal{C}}$ is the charge conjugation operator; γ^0 is the time-like Dirac gamma-matrix; γ^2 is the space-like the Dirac gamma-matrix (one of three); Then the symmetry under charge transformation is,

$$|\hat{\mathcal{C}}\bar{\psi}^T(\vec{x},t)|^2 = |\psi(\vec{x},t)|^2, \text{ iff } \mathcal{C} \text{ is conserved}, \tag{2.2}$$

where $\bar{\psi}(\vec{x},t) = \psi^\dagger(\vec{x},t)\gamma^0$ is the Dirac adjoint spinor. The parity $\hat{\mathcal{P}}$ transformation is defined as,

$$\psi(\vec{x},t) \xrightarrow{P} \hat{\mathcal{P}}\psi(-\vec{x},t) \equiv \gamma^0\psi(-\vec{x},t), \tag{2.3}$$

where $\hat{\mathcal{P}}$ is the parity operator. The symmetry under parity $\hat{\mathcal{P}}$ implies that,

$$|\mathcal{P}\psi(-\vec{x},t)|^2 = |\psi(\vec{x},t)|^2 \text{ iff } \mathcal{P} \text{ is conserved}. \tag{2.4}$$

The time reflection $\mathcal{T}$ is defined as,

$$\psi(\vec{x},t) \xrightarrow{T} \hat{\mathcal{T}}\psi^*(\vec{x},-t) \equiv i\gamma^1\gamma^3\psi^*(\vec{x},-t), \tag{2.5}$$

where $\hat{\mathcal{T}}$ is the time operator; $\psi^*(\vec{x},t)$ is the complex-conjugate of $\psi(\vec{x},t)$; γ^1 and γ^3 are space-like the Dirac gamma-matrices. The symmetry under time reflection $\mathcal{T}$ is,

$$|\hat{\mathcal{T}}\psi^*(\vec{x},-t)|^2 = |\psi(\vec{x},t)|^2 \text{ iff } \mathcal{T} \text{ is conserved.} \tag{2.6}$$

The $\mathcal{CPT}$theorem states that any local Lorentz-invariant quantum field theory obeys $\mathcal{CPT}$symmetry, which makes SM $\mathcal{CPT}$-invariant. However, the SM is not invariant under the individual transformations, and the Eq. 2.2, 2.4, 2.6 do not hold. In particular, Wu et al. [9] observed that the parity symmetry is broken, which is because in the SM the weak interaction only couples to the left-chiral particles and right-chiral anti-particles.

2.2 Weak interactions and $\mathcal{CP}$-violation

In the SM the fermion states interacting weakly are not the same as the mass eigenstates. As a result, when expressed in mass eigenstates, interactions change the type of a quark, known as flavour. The quarks have six flavours: up, down, charm, strange, top, and beauty. Figure 2.2 shows an example of a weak interaction vertex, where the incoming beauty (b) quark is transformed into an outgoing up (u) quark. The strength of the flavour-changing transition is described by a single coupling constant g_W and an element of the unitary complex Cabibbo-Kobayashi-Maskawa matrix, for example, V_{ub} as in Fig. 2.2.

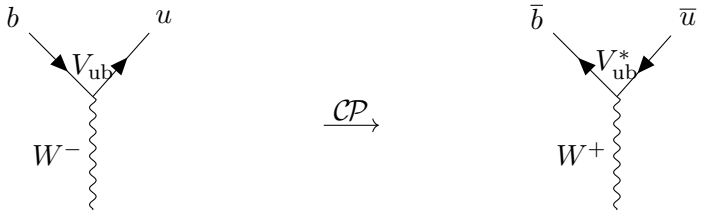


Figure 2.2: On the left: a $b-u$ weak transition, which is mediated by W^-. On the right: a $\overline{u}-\overline{b}$ weak transition, which is mediated by W^+. The coupling strength V_{ub} is also shown. Diagrams are created with the `TikZ-Feynman` package [10].

The Cabbibo-Kobayashi-Maskawa matrix V_{CKM} or the *CKM matrix* is a

3×3 matrix [11, 12],

$$V_{\rm CKM} = \begin{pmatrix} V_{\rm ud} & V_{\rm us} & V_{\rm ub} \\ V_{\rm cd} & V_{\rm cs} & V_{\rm cb} \\ V_{\rm td} & V_{\rm ts} & V_{\rm tb} \end{pmatrix}, \tag{2.7}$$

where $V_{\rm ij}$ is a matrix element for the $i-j$ quark transitions.

The unitarity condition imposes the following relation between the CKM matrix elements,

$$\sum_{\rm i} V_{\rm ij} V_{\rm ik}^* = \delta_{\rm jk}. \tag{2.8}$$

The relations for $i \neq k$ can be represented in the form of triangles. Two of them have approximately equal-sized terms, [13]:

$$V_{\rm ud}V_{\rm ub}^* + V_{\rm cd}V_{\rm cb}^* + V_{\rm td}V_{\rm tb}^* = 0, \tag{2.9}$$

$$V_{\rm td}V_{\rm ud}^* + V_{\rm ts}V_{\rm us}^* + V_{\rm tb}V_{\rm ub}^* = 0. \tag{2.10}$$

Figure 2.3 shows the experimental status of the triangle from the first relation on the $(\bar{\rho}, \bar{\eta})$ plane [14], where the real parameters $\bar{\rho}$ and $\bar{\eta}$ are defined as,

$$\bar{\rho} + i\bar{\eta} \equiv \frac{V_{\rm ud}V_{\rm ub}^*}{V_{\rm cb}V_{\rm cb}^*} \tag{2.11}$$

This triangle is known as *the* Unitary Triangle with angles defined as:

$$\alpha \equiv \arg\Big(-\frac{V_{\rm td}V_{\rm tb}^*}{V_{\rm ud}V_{\rm ub}^*}\Big), \tag{2.12}$$

$$\beta \equiv \arg\Big(-\frac{V_{\rm cd}V_{\rm cb}^*}{V_{\rm td}V_{\rm tb}^*}\Big), \tag{2.13}$$

$$\gamma \equiv \arg\Big(-\frac{V_{\rm ud}V_{\rm ub}^*}{V_{\rm cd}V_{\rm cb}^*}\Big). \tag{2.14}$$

The second triangle from Eq. 2.10 is shown in Fig. 2.4 on the $(\bar{\rho}_{sb}, \bar{\eta}_{sb})$ plane, where $\bar{\rho}_{sb} + i\bar{\eta}_{sb} \equiv -\frac{V_{\rm ub}V_{\rm us}^*}{V_{\rm cb}V_{\rm cs}^*}$. It is characterized by the angle β_s defined as,

$$\beta_s \equiv \arg\Big(-\frac{V_{\rm ts}V_{\rm tb}^*}{V_{\rm cs}V_{\rm cb}^*}\Big) \tag{2.15}$$

A unitary complex matrix of size $n \times n$ has $(n-1)^2$ free parameters, which are divided into the $\frac{1}{2}n(n-1)$ rotation angles and $\frac{1}{2}(n-1)(n-2)$ complex phases. The size of the CKM matrix is defined by the number of

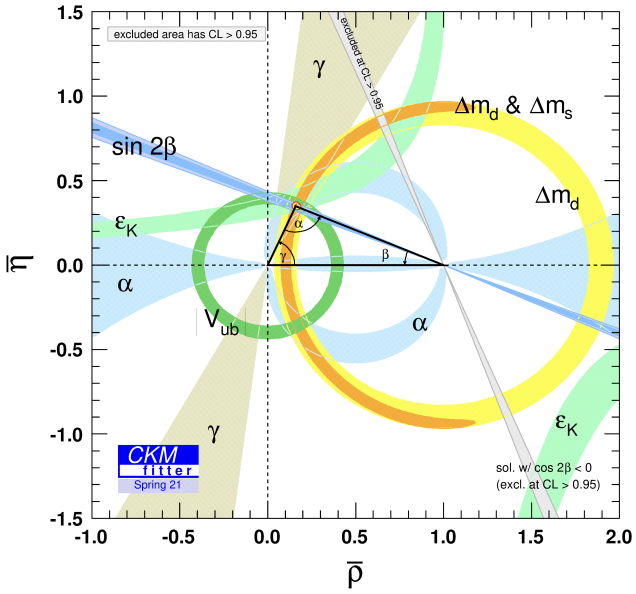


Figure 2.3: The Unitarity triangle of Eq. 2.9. From [14].

quark generations in the SM, which is three. Three angles and one complex phase parameterize a unitary complex matrix of size 3×3. The presence of a complex phase means that, in general, $V_{ij} \neq V_{ij}^*$, so that the left and right amplitude in Fig. 2.2 differ. This complex phase is known as the $\mathcal{CP}$-violating phase $\delta^{\mathcal{CP}}$. It is the source[1] of $\mathcal{CP}$-violation in the quark sector of the SM [13]. The angles γ, β, and β_s of the Unitary triangle are all connected to this one phase $\delta^{\mathcal{CP}}$. Note that if there were fewer than three generations, the complex phase would not be present, which means that to have $\mathcal{CP}$-violation in the SM, at least three generations of matter are needed.

Observing the $\mathcal{CP}$-violation is a non-trivial task. The $\mathcal{CP}$-violating phase introduces the difference between the left and the right sides of Fig. 2.2. However, the probability of both transitions is the same, as it is the squared magnitude of the amplitude. So, to reveal the $\mathcal{CP}$-violating (weak) phase in the experiment, it is not enough to compare two processes that differ by the $\mathcal{CP}$ transformation, as in Fig. 2.2. The first conditions to observe $\mathcal{CP}$-violation are:

[1]The lepton sector also has $\mathcal{CP}$-violation, although not observed yet.

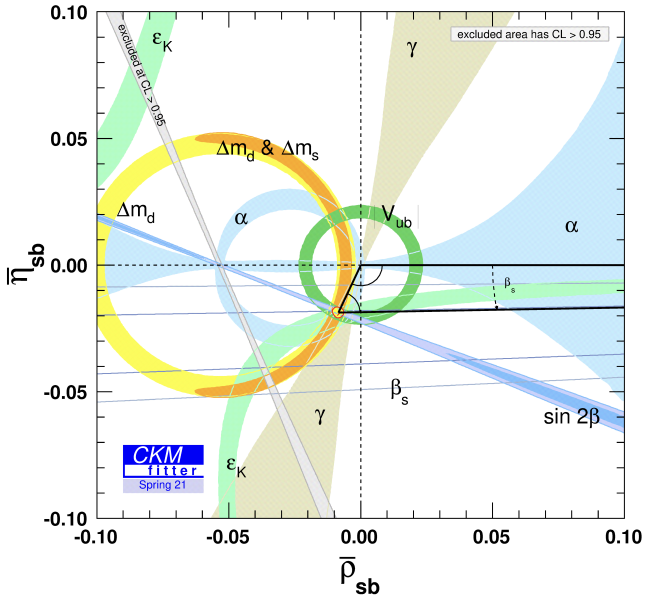


Figure 2.4: A unitary triangle of Eq. 2.10. From [14].

1. There must be two amplitudes from the initial state to the final state f in the processes; see Fig. 2.5.

2. There should be both a $\mathcal{CP}$-violating (weak) phase and a $\mathcal{CP}$-conserving phase in the amplitude. The first phase flips the sign under $\mathcal{CP}$ operation, but the second one does not. Therefore, the sum of the two phases will change when the $\mathcal{CP}$ transformation is applied.

Choosing the two amplitudes as A_1 and A_2 with the weak $\mathcal{CP}$-violating phases ϕ_1, ϕ_2 and strong $\mathcal{CP}$-conserving phases δ_1, δ_2 respectively,

$$\begin{aligned} A_1 &= |A_1| e^{i\phi_1} e^{i\delta_1}, \\ A_2 &= |A_2| e^{i\phi_2} e^{i\delta_2}. \end{aligned} \tag{2.16}$$

The amplitude of the process A to which both amplitudes A_1 and A_2 contribute is then,

$$A_f = |A_1| e^{i\phi_1} e^{i\delta_1} + |A_2| e^{i\phi_2} e^{i\delta_2}. \tag{2.17}$$

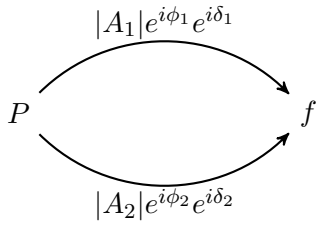


Figure 2.5: A particle P decaying into the final state f via the two amplitudes A_1 and A_2. The amplitudes are defined as $A_1 = |A_1|e^{i\phi_1}e^{i\delta_1}$ and $A_2 = |A_2|e^{i\phi_2}e^{i\delta_2}$, where $\phi_{1,2}$ are the weak $\mathcal{CP}$-violating phases and $\delta_{1,2}$ are the strong $\mathcal{CP}$-conserving phases.

Its transition rate $\mathcal{T}$ is proportional to the $|A_f|^2$,

$$\mathcal{T} \propto |A_f|^2 = |A_1|^2 + |A_2|^2 + 2|A_1||A_2|(\cos(\phi_1 - \phi_2)\cos(\delta_1 - \delta_2) - \sin(\phi_1 - \phi_2)\sin(\delta_1 - \delta_2)). \tag{2.18}$$

The amplitude for the $\mathcal{CP}$-conjugate process, labeled by $\bar{A}_{\bar{f}}$, is then given by

$$\bar{A}_{\bar{f}} = |A_1|e^{-i\phi_1}e^{i\delta_1} + |A_2|e^{-i\phi_2}e^{i\delta_2}, \tag{2.19}$$

with transition rate proportional to,

$$|\bar{A}_{\bar{f}}|^2 = |A_1|^2 + |A_2|^2 + 2|A_1||A_2|(\cos(\phi_1 - \phi_2)\cos(\delta_1 - \delta_2) + \sin(\phi_1 - \phi_2)\sin(\delta_1 - \delta_2)). \tag{2.20}$$

The $|A_f|^2$ and $|\bar{A}_{\bar{f}}|^2$ are the same up to the sign in front of the sin terms. The sin term disappears in case either weak or strong phases are the same. This leads to the last condition of observing $\mathcal{CP}$-violation:

3. The weak and the strong phases of A_1 and A_2 must be different: $\phi_1 \neq \phi_2$ and $\delta_1 \neq \delta_2$.

If the above conditions are met, then the $|A_f| \neq |\bar{A}_{\bar{f}}|$. This type of $\mathcal{CP}$-violation is known as the *direct $\mathcal{CP}$-violation* or the $\mathcal{CP}$-violation in decay [13]. The direct $\mathcal{CP}$-violation results in the observable $\mathcal{CP}$-asymmetry A_{CP}, which is defined as,

$$A_{CP} \equiv \frac{\Gamma_{P\to f} - \Gamma_{\bar{P}\to\bar{f}}}{\Gamma_{P\to f} + \Gamma_{\bar{P}\to\bar{f}}}, \tag{2.21}$$

where $\Gamma_{P\to f}$ is the decay rate of particle P to the final state f and $\Gamma_{\bar{P}\to\bar{f}}$ is the decay rate of the $\mathcal{CP}$-conjugate process.

The first observation of $\mathcal{CP}$-violation was done in the sixties in the kaon system [15]. It took almost forty years more to observe $\mathcal{CP}$-violation in the B-meson systems, which was done by Babar and Belle [16], [17] in the $B^0 \to J/\psi K^0$ decays. Only recently LHCb discovered $\mathcal{CP}$-violation in the charm system [18].

The total amount of $\mathcal{CP}$-violation present in the SM quark sector is quantified by the Jarlskog invariant $\mathcal{J}$, which is equal to twice the area of the unitary triangle in Fig. 2.3. The matter–antimatter asymmetry generated from the quark sector of the SM is estimated to be [19],

$$\begin{aligned}\mathcal{CP}^{SM} &= \frac{2\cdot\mathcal{J}\cdot(m_t^2-m_c^2)(m_t^2-m_u^2)(m_c^2-m_u^2)(m_b^2-m_s^2)(m_b^2-m_d^2)(m_s^2-m_d^2)}{\Lambda_{EW}^{12}} \\ &\approx 10^{-20},\end{aligned} \tag{2.22}$$

where m_t is the top quark mass; m_c is the charm quark mass; m_u is the up quark mass; m_b is the beauty quark mass; m_s is the strange quark mass and m_d is the down quark mass; $\Lambda_{EW} \approx 160$ GeV is the electroweak symmetry breaking scale, which is the temperature at which the electroweak phase transition happened. Interestingly, if quarks had degenerate masses or we could not distinguish between the quarks generation, no matter–antimatter asymmetry would be generated in the SM quark sector.

Cosmological observations show that the matter–antimatter asymmetry of the Universe is $\mathcal{O}(10^{-10})$, see Eq. 1.1, which is ten orders more than can be generated by the $\mathcal{CP}$-violation in the quark sector of the SM. Such a huge difference motivates the search for new $\mathcal{CP}$-violating phenomena that could explain the observed matter–antimatter asymmetry.

The term "beyond SM physics" is an umbrella term for the hypothetical extensions of the SM that add new phenomena to the model. Those effects can include extra $\mathcal{CP}$-violating effects. From Eq. 2.22, one can think of possible extensions to the SM in the quarks sector to increase the amount of $\mathcal{CP}$-violation: either those that modify the amount of $\mathcal{CP}$-violation seen in the Jarlskog invariant $\mathcal{J}$ or new $\mathcal{CP}$-violating interactions outside of the CKM mechanism. To modify the amount of $\mathcal{CP}$-violation in the Jarlskog invariant an additional heavy generation is needed. The number of generations is however strongly constrained by the number of light neutrino types estimated from the Z-boson production measurements in e^+-e^- colliders. The world average on the number of light neutrino types is 2.9963±0.0074 [20]. Assuming that each generation has two quarks and two leptons as in Fig. 2.1 - "full" generation, the observation of 3 neutrino types implies that there

are only three generations of quarks in the SM. However, if the requirement on the "full" generation is lifted, then there is a possibility of having a single-fourth heavy quark with a mass above the top quark mass[2] or a heavy neutrino with a mass above half of the Z-boson mass.

The new $\mathcal{CP}$-violating interactions can also be found by precisely measuring the $\mathcal{CP}$-violating angles of the unitarity triangles and comparing them to their SM predictions. If a discrepancy is observed, it will be a sign of the new $\mathcal{CP}$-violating phenomena, which will contribute to the solution of the matter abundance puzzle.

This abundance can also be generated via the lepton sector. In analogy to the CKM matrix in the quarks sector, in the lepton sector, there is a Pontecorvo-Maki-Nagakawa-Sakata matrix or the *PNMS* matrix [21]. The $\mathcal{CP}$-violation has not been observed in the lepton sector yet, but to get the observed matter–antimatter asymmetry, the lepton sector would require a heavy neutral lepton, so that analogously to Eq. 2.22, the mass splits are substantial. For example, a hypothetical heavy right-handed Majorana neutrino could explain the observed asymmetry [22].

The amount of $\mathcal{CP}$-violation in the second triangle on Fig. 2.4 is defined by a single angle β_s, see Fig. 2.4. It is predicted to be relatively small (almost zero) in the SM and it approximately equals $-\phi_s/2$, which is the $\mathcal{CP}$-violating phase generated in the interference between $b \to c\bar{c}s$ transitions with neutral meson mixing and without neutral meson mixing. To define ϕ_s exactly, first, neutral meson mixing has to be introduced.

2.3 B_s^0-$\overline{B}_s^0$ oscillations

The B_s^0 meson is a neutral meson comprised of anti-beauty and strange quarks. Its antiparticle is $\overline{B}_s^0$, composed of anti-strange and beauty quarks. The B_s^0 and $\overline{B}_s^0$ are flavour basis eigenstates of B_s^0 meson and have a definite quark composition, which defines the interactions of B_s^0 and $\overline{B}_s^0$. The B_s^0 meson has two mass eigenstates: heavy B_H and light B_L, which have definite masses and lifetimes. The mass eigenstates are eigenstates of a full Hamiltonian. The CKM matrix is the result of the fact that the meson mass eigenstates do not correspond to the states that are part of the doublets that take part in the weak gauge interaction. The flavour eigenstate is not an eigenstate of the full Hamiltonian, but it can be represented as a superposition of mass eigenstates and vice versa. The B_s^0 meson is produced as a flavour eigenstate,

[2] A heavy quark with a mass above the top quark mass would be hard to fit within the SM, as its mass is too heavy to be obtained by the SM Higgs mechanism.

but its evolution is defined by B_H and B_L propagation. The presence of distinctive mass and flavour eigenstates enables B_s^0-$\overline{B}_s^0$ oscillations, also known as the B_s^0-$\overline{B}_s^0$ mixing. Figure 2.6 shows two diagrams that describe the B_s^0-$\overline{B}_s^0$ mixing, so-called "box" diagrams. The transition goes via the exchange of one of three quarks: u, c or t. In the limit of equal quark masses, these amplitudes would destructively interfere and their sum would be zero [23]. Hence it is the contribution of the t quark, which is the outlier in terms of mass, that dominates the process. Ignoring other contributions except for t-quark, the amplitude of Fig. 2.6 is proportional to the V_{ts} and V_{tb}^* CKM elements.

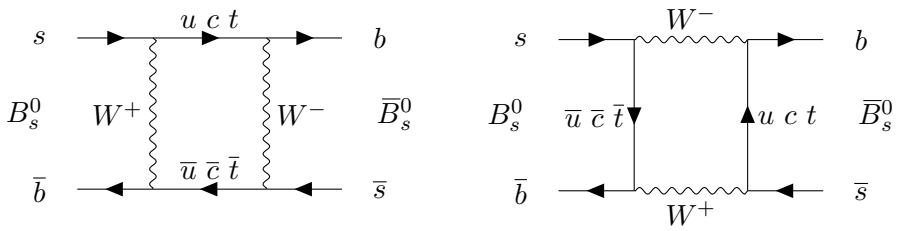


Figure 2.6: The B_s^0-$\overline{B}_s^0$ mixing diagrams. Diagrams are created with the `TikZ-Feynman` package [10].

The B_s^0 meson created at $t = 0$ evolves as a superposition of B_s^0 and $\overline{B}_s^0$ [7],

$$|B_s^0(t)\rangle = a(t)|B_s^0\rangle + b(t)|\overline{B}_s^0\rangle, \tag{2.23}$$

where $|B_s^0(t)\rangle$ is time evolution of B_s^0-meson; $|B_s^0\rangle, |\overline{B}_s^0\rangle$ are flavour eigenstates B_s^0 and $\overline{B}_s^0$; $a(t)$ and $b(t)$ describe time-evolution. The $a(t)$ and $b(t)$ are defined by the Schrödinger equation for an "open" system, which does not describe the decay,

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \overbrace{\left[\underbrace{\begin{pmatrix} M & M_{12} \\ M_{12}^* & M \end{pmatrix}}_{\text{mass matrix}} - \frac{i}{2}\underbrace{\begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix}}_{\text{decay matrix}}\right]}^{\boldsymbol{M}-\frac{\boldsymbol{i}}{\boldsymbol{2}}\boldsymbol{\Gamma}=\text{mass-decay matrix}} \cdot \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{2.24}$$

where $\boldsymbol{M}$ is a mass term and $\boldsymbol{\Gamma}$ is a decay width term. In the absence of B_s^0-$\overline{B}_s^0$ mixing the off-diagonal terms M_{12} and Γ_{12} are zero. M_{12} describes the short-distance transitions that happen off-shell (via u, c, t quarks in Fig. 2.6) and Γ_{12} describes the on-shell contributions via resonances, which contain u, c quarks in Fig. 2.6 [24]. The imaginary i in front of the decay

matrix means that the probability of observing B_s^0 or $\overline{B}_s^0$ is not conserved and decreases with time [13]. Because the SM is $\mathcal{CPT}$invariant, the diagonal terms of both matrices are equal, and $M_{21} = M_{12}^*$, while $\Gamma_{21} = \Gamma_{12}^*$. The eigenstates of $\boldsymbol{M - \frac{i}{2}\Gamma}$ are the two mass eigenstates B_H and B_L with defined mass and lifetime. As a superposition of flavour eigenstates, B_H and B_L are defined as,

$$|B_H\rangle = p|B_s^0\rangle + q|\overline{B}_s^0\rangle \tag{2.25}$$

$$|B_L\rangle = p|B_s^0\rangle - q|\overline{B}_s^0\rangle, \tag{2.26}$$

where p and q complex constants. The value of p and q expressed in terms of the Hamiltonian elements is,

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{2.27}$$

The time-evolution of B_H and B_L is given by the solutions to Schrödinger equation:

$$|B_H(t)\rangle = e^{-iM_H t - \frac{\Gamma_H}{2}t}|B_H(0)\rangle \tag{2.28}$$

$$|B_L(t)\rangle = e^{-iM_L t - \frac{\Gamma_L}{2}t}|B_L(0)\rangle, \tag{2.29}$$

where the $M_{L,H}$ and $\Gamma_{L,H}$ are masses and decay widths[3] of B_L and B_H. The B_s^0 mass $m_{B_s^0}$, decay width Γ_s are defined as,

$$m_{B_s^0} = \frac{M_H + M_L}{2} \tag{2.30}$$

$$\Gamma_s = \frac{\Gamma_H + \Gamma_L}{2}. \tag{2.31}$$

The ratio $\left|\frac{M_{12}}{\Gamma_{12}}\right| \approx 5 \cdot 10^{-3}$, therefore the decay width difference $\Delta\Gamma_s$ and oscillation frequency Δm_s of B_s^0 to $\overline{B}_s^0$ are approximated as [24],

$$\Delta m_s = M_H - M_L = 2|M_{12}| + \mathcal{O}(\left|\frac{M_{12}}{\Gamma_{12}}\right|^2) \tag{2.32}$$

$$\Delta\Gamma_s = \Gamma_L - \Gamma_H = 2|\Gamma_{12}|\cos\left(\arg\left[-\frac{M_{12}}{\Gamma_{12}}\right]\right) + \mathcal{O}(\left|\frac{M_{12}}{\Gamma_{12}}\right|^2) \tag{2.33}$$

where M_{12}, Γ_{12} are from Eq. 2.24.

[3]Note that decay width Γ is inversely proportional to the lifetime τ.

From Eq. 2.25 and Eq. 2.28, the time propagation of B_s^0 and $\overline{B}_s^0$ is,

$$|B_s^0(t)\rangle = g_+(t)|B_s^0\rangle + \frac{q}{p}g_-(t)|\overline{B}_s^0\rangle \tag{2.34}$$

$$|\overline{B}_s^0(t)\rangle = \frac{p}{q}g_-(t)|B_s^0\rangle + g_+(t)|\overline{B}_s^0\rangle, \tag{2.35}$$

where $g_+(t)$ and $g_-(t)$ are complex time-dependent functions,

$$g_+(t) = \frac{1}{2}\Big(e^{-iM_Ht-\frac{1}{2}\Gamma_Ht} + e^{-iM_Lt-\frac{1}{2}\Gamma_Lt}\Big) \tag{2.36}$$

$$g_-(t) = \frac{1}{2}\Big(e^{-iM_Ht-\frac{1}{2}\Gamma_Ht} - e^{-iM_Lt-\frac{1}{2}\Gamma_Lt}\Big). \tag{2.37}$$

The probability of observing the $\overline{B}_s^0$, given that at $t = 0$ the B_s^0 was produced is then,

$$|\langle\overline{B}_s^0|B_s^0(t)\rangle|^2 = |g_-(t)|^2\left|\frac{q}{p}\right|^2. \tag{2.38}$$

For a $\overline{B}_s^0$ at $t = 0$, the probability of observing a B_s^0 at t is,

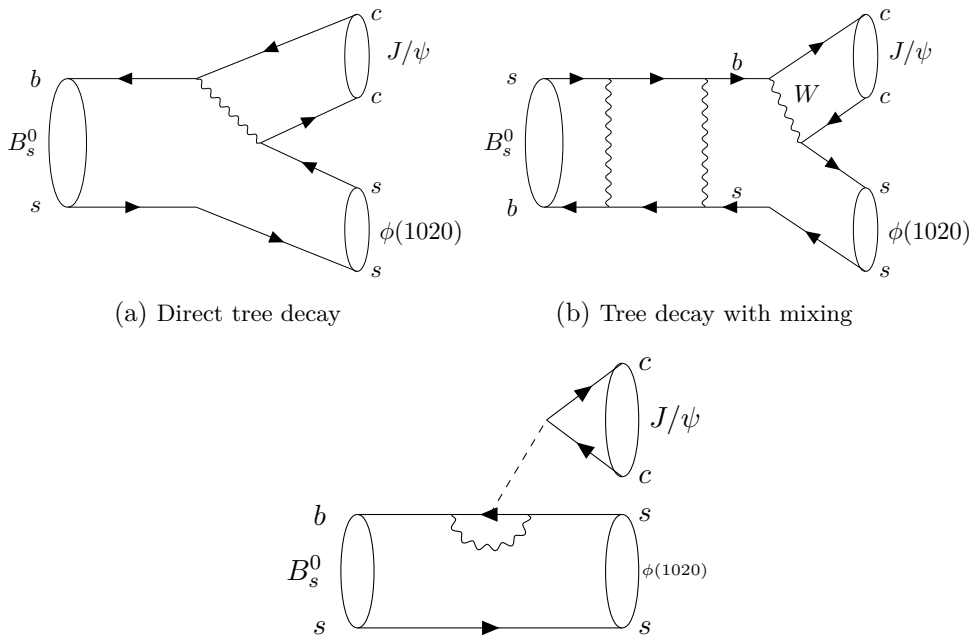
$$|\langle B_s^0|\overline{B}_s^0(t)\rangle|^2 = |g_-(t)|^2\left|\frac{p}{q}\right|^2. \tag{2.39}$$

The $\mathcal{CP}$-violation in mixing appears when the rate of $B_s^0 \to \overline{B}_s^0$ transitions is different from $\overline{B}_s^0 \to B_s^0$ transitions. This $\mathcal{CP}$-violation is observed when $\left|\frac{q}{p}\right| \neq 1$ [13].

The magnitude of complex time-dependent functions $g_-(t)$ and $g_+(t)$ is,

$$|g_\pm(t)|^2 = \frac{1}{2}e^{-\Gamma_st}\Big(\cosh(\Delta\Gamma_st/2) \pm \cos(\Delta m_st)\Big). \tag{2.40}$$

As mentioned before, the $\mathcal{CP}$-violating phase ϕ_s is generated in interference between the $b \to c\bar{c}s$ transitions with neutral meson mixing and without neutral meson mixing. To define ϕ_s we also have to consider B_s^0 decay. The decay chosen to measure the $\mathcal{CP}$-violating phase ϕ_s is $B_s^0 \to J/\psi\,\phi(1020)$, see Fig. 2.7. The first Feynman diagram in Fig. 2.7(a) is the tree-level direct decay of B_s^0 to J/ψ $\phi(1020)$. Its amplitude is proportional to the V_{cb}^* and V_{cs}. But as was discussed in Sec. 2.3 B_s^0 can oscillate to $\overline{B}_s^0$. The $\overline{B}_s^0$ can also decay into J/ψ $\phi(1020)$, see Fig. 2.7(b). This amplitude is then proportional to the V_{cb} and V_{cs}^*. Both amplitudes interfere and their interference terms generate the ϕ_s phase. This type of observable $\mathcal{CP}$-violation is known as *mixing-induced* $\mathcal{CP}$-violation.



(a) Direct tree decay

(b) Tree decay with mixing

(c) Direct penguin decay

Figure 2.7: The Feynman diagrams for the direct tree-level $B_s^0 \to J/\psi\,\phi(1020)$ decay, tree-level $B_s^0 \to J/\psi\,\phi(1020)$ decay with mixing, and direct penguin-level $B_s^0 \to J/\psi\,\phi(1020)$ decay are shown. Diagrams are created with [10].

The higher order processes, like the penguin amplitude shown in Fig. 2.7(c), also contribute to the $B_s^0 \to J/\psi\,\phi(1020)$ decay. However, these contributions are suppressed by a factor of 0.005 [6], and the $B_s^0 \to J/\psi\,\phi(1020)$ decay is dominated by the tree level amplitude. Nevertheless, the tiny interference between the direct tree decay in Fig. 2.7(a) and direct penguin decay in Fig. 2.7(c) potentially generates direct $\mathcal{CP}$-violation in the decay.

The J/ψ in $B_s^0 \to J/\psi\,\phi(1020)$ is reconstructed from the decay $J/\psi \to \mu^+\mu^-$, which has a high branching fraction [20] and a high experimental reconstruction efficiency for muons [25]. The $\phi(1020)$ is reconstructed from the decay $\phi(1020) \to K^+K^-$, which is the highest branching fraction decay mode for a $\phi(1020)$ [20].

2.4 Decay to $J/\psi\,\phi(1020)$

Two decay amplitudes A contribute to $B_s^0 \to J/\psi\,\phi(1020)$ decay, where ϕ stands for $\phi(1020)$ resonance:

$$\begin{aligned} A_{J/\psi\,\phi} &= \langle J/\psi\,\phi|B_s^0\rangle \\ &= g_+(t)\langle J/\psi\,\phi|B_s^0\rangle + \frac{q}{p}g_-(t)\langle J/\psi\,\phi|\overline{B}_s^0\rangle, \qquad (2.41) \\ \bar{A}_{J/\psi\,\phi} &= \langle J/\psi\,\phi|\overline{B}_s^0\rangle \\ &= \frac{p}{q}g_-(t)\langle J/\psi\,\phi|B_s^0\rangle + g_+(t)\langle J/\psi\,\phi|\overline{B}_s^0\rangle. \qquad (2.42) \end{aligned}$$

The $\mathcal{CP}$-violation in $B_s^0 \to J/\psi\,\phi(1020)$ is described by a complex parameters $\lambda_{f=J/\psi\,\phi}$, which is defined as,

$$\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}. \qquad (2.43)$$

The decay time-dependent decay rates are equal to [24],

$$\begin{aligned} \frac{d\Gamma_{B_s^0\to J/\psi\,\phi}}{dt} &= |\langle J/\psi\,\phi|B_s^0\rangle|^2 \\ &= |A_f|^2\Big(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + 2\Re[\lambda_f g_+(t)^* g_-(t)]\Big) \\ &= |A_f|^2\frac{1}{1+C_f}e^{-\Gamma_s t}\Bigg[\cosh\Big(\frac{\Delta\Gamma_s t}{2}\Big) + D_f\sinh\Big(\frac{\Delta\Gamma_s t}{2}\Big) \\ &\quad + C_f\cos(\Delta m_s t) + S_f\sin(\Delta m_s t)\Bigg] \qquad (2.44) \end{aligned}$$

$$\begin{aligned} \frac{d\Gamma_{\overline{B}_s^0\to J/\psi\,\phi}}{dt} &= |\langle J/\psi\,\phi|\overline{B}_s^0\rangle|^2 \\ &= |A_f|^2\Big|\frac{p}{q}\Big|^2\Big(|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 + 2\Re[\lambda_f g_+(t) g_-(t)^*]\Big) \\ &= |A_f|^2\Big|\frac{p}{q}\Big|^2\frac{1}{1+C_f}e^{-\Gamma_s t}\Bigg[\cosh\Big(\frac{\Delta\Gamma_s t}{2}\Big) + D_f\sinh\Big(\frac{\Delta\Gamma_s t}{2}\Big) \\ &\quad - C_f\cos(\Delta m_s t) - S_f\sin(\Delta m_s t)\Bigg] \qquad (2.45) \end{aligned}$$

with coefficients C_f, D_f and S_f defined as,

$$C_f \equiv \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \tag{2.46}$$

$$D_f \equiv -\frac{2\Re(\lambda_f)}{1+|\lambda_f|^2}, \tag{2.47}$$

$$S_f \equiv -\frac{2\Im(\lambda_f)}{1+|\lambda_f|^2}. \tag{2.48}$$

C_f quantifies the amount of direct $\mathcal{CP}$-violation in the $B_s^0 \to J/\psi\,\phi(1020)$ decay; S_f quantifies the amount of mixing-induced $\mathcal{CP}$-violation, which is proportional to $\sin(\phi_s)$. If $|\lambda_f|$ is one, then no direct $\mathcal{CP}$-violation is present. If ϕ_s is zero, then no mixing-induces $\mathcal{CP}$-violation is present.

If the final state is a $\mathcal{CP}$-eigenstate, then

$$\mathcal{CP}|f\rangle = \eta_f|f\rangle = |\bar{f}\rangle, \tag{2.49}$$

where η_f is the $\mathcal{CP}$-eigenvalue of state f. For a final $\mathcal{CP}$-eigenstate, the following relations hold:

$$A_f = \eta_f A_{\bar{f}} \tag{2.50}$$

$$\bar{A}_f = \eta_f \bar{A}_{\bar{f}}. \tag{2.51}$$

The parameter λ_f is then given by,

$$\lambda_f = \eta_f \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_f} = \eta_f|\lambda|e^{-i\phi_s} = \eta_f\Big(\frac{V_{\mathrm{tb}}^* V_{\mathrm{ts}}}{V_{\mathrm{tb}} V_{\mathrm{ts}}^*}\Big)\Big(\frac{V_{\mathrm{cb}} V_{\mathrm{cs}}^*}{V_{\mathrm{cb}}^* V_{\mathrm{cs}}}\Big). \tag{2.52}$$

Finally, the $\mathcal{CP}$-violating phase ϕ_s is defined as [24],

$$\phi_s \equiv -\arg(\lambda_f) = -\arg\Big(\eta_f \frac{V_{\mathrm{tb}}^* V_{\mathrm{ts}}}{V_{\mathrm{tb}} V_{\mathrm{ts}}^*}\frac{V_{\mathrm{cb}} V_{\mathrm{cs}}^*}{V_{\mathrm{cb}}^* V_{\mathrm{cs}}}\Big) = -\pi + \phi_M - 2\phi_D, \tag{2.53}$$

where ϕ_M is a weak mixing phase and ϕ_D is a weak decay phase in the $B_s^0 \to J/\psi\,\phi(1020)$ transition. Assuming $\eta_f = 1$ and ignoring penguin contributions, $\phi_s \approx -2\beta_s$, where β_s is defined in Eq. 2.15.

The value of η_f depends on the orbital angular momentum L of the final state,

$$\eta_f = (-1)^L. \tag{2.54}$$

To determine the value of L, consider the total angular momentum J, which is the sum of the orbital angular momentum and the internal angular momentum (spin). As the B_s^0 meson is a spinless particle, it has zero total

angular momentum. This has to be conserved in the decay. In the case of a $B_s^0 \to J/\psi\,\phi(1020)$ decay, the decay products are spin-one particles, and as a result, the orbital angular momentum L has to balance out the spin orientation of J/ψ and $\phi(1020)$. This gives rise to three possible configurations, depending on the relative value of the orbital angular momentum, which are known as polarization amplitudes. Figure 2.8 shows all three polarization amplitudes. The first one, A_0, has polarization vectors parallel to the J/ψ and $\phi(1020)$ momenta and has a zero relative orbital angular momentum. The second one, $A_\parallel$, has aligned polarization vectors perpendicular to the J/ψ and $\phi(1020)$ momenta and has a relative orbital momentum of two. The last one, $A_\perp$, has perpendicular polarization vectors and has a relative orbital momentum of one.

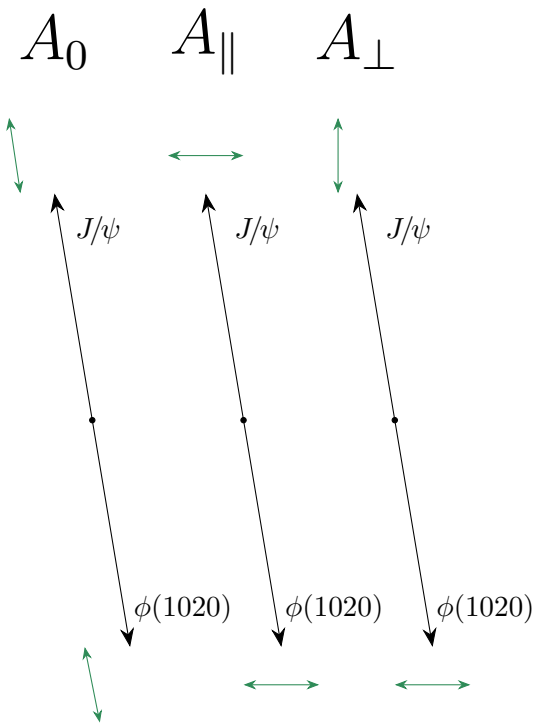


Figure 2.8: Three polarization of final state amplitudes in $B_s^0 \to J/\psi\,\phi(1020)$. Black arrows are momenta directions and green are polarization vectors.

These three polarization amplitudes have different $\mathcal{CP}$-eigenvalues. Both A_0 and $A_\parallel$ are $\mathcal{CP}$-even with $\eta_f = 1$, while $A_\perp$ is $\mathcal{CP}$-odd with $\eta_f = -1$. The $\mathcal{CP}$-even part then contributes to $\mathcal{CP}$-violation of the decay with positive

λ_f, see Eq. 2.52, and $\mathcal{CP}$-odd with negative λ_f. This introduces a dilution to the overall measured λ_f. Therefore, to increase the statistical power of the measurement, the three polarization amplitudes should be disentangled from each other. Luckily, each polarization has a unique angular distribution, and by measuring the decay angles, the relative contributions of the different polarization amplitudes can be determined. How many angles are necessary to describe the $B_s^0 \to J/\psi\,\phi(1020)$ decay and disentangle the three polarization amplitudes? As both J/ψ and $\phi(1020)$ decay into two tracks, $B_s^0 \to J/\psi\,\phi(1020)$ is a four-body decay. Three angles are needed to describe such a four-body decay. They are defined within the "helicity" basis and are shown in Fig. 2.9. The angle θ_K is the angle between the direction of B_s^0 in

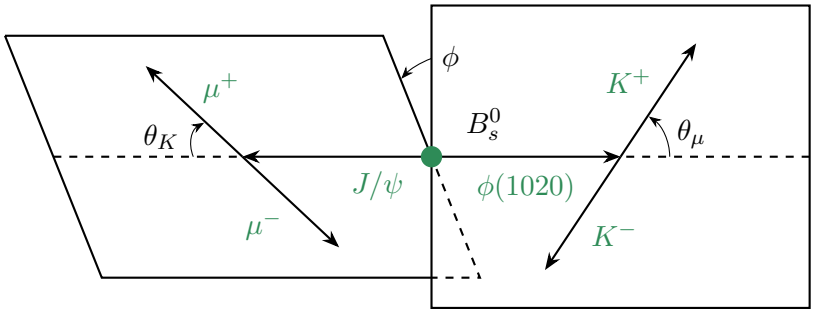


Figure 2.9: The three helicity angles of $B_s^0 \to J/\psi\,\phi(1020)$ decay, where $J/\psi \to \mu\mu^+$ and $\phi(1020) \to K^-K^+$.

the $\phi(1020)$ center-of-mass frame and the K^-. The angle θ_μ is defined as the angle between the direction of B_s^0 in the J/ψ center-of-mass frame and the μ^+. The angle ϕ is the angle between the J/ψ and the $\phi(1020)$ decay planes in the B_s^0 center-of-mass frame. The decay time-dependent decay rate in Eq. 2.44 must include the angular dependence to disentangle the polarization amplitudes. The three amplitudes and their interferences make up six real terms, which contribute to the decay time- and angular-dependent decay rate.

However, this is not the full story. The probability of a B_s^0 meson to decay into J/ψ and a kaon pair via the $\phi(1020)$ resonance is $(1.04 \pm 0.04) \times 10^{-3}$ [20]. However, the same final state can be reached either without an intermediate resonance with probability $(7.9 \pm 0.7) \times 10^{-4}$ or via the spin-zero $f_0(980)$ resonance with probability $(1.24 \pm 0.15) \times 10^{-4}$ [20]. The $B_s^0 \to J/\psi\,K^+K^-$ decay that went through the $\phi(1020)$ resonance is known as P-wave. The P-wave includes all three polarizations: A_0, $A_\parallel$, and $A_\perp$. Whereas the non-resonant and decays via the spin-zero $f_0(980)$ are known as S-wave. In a

similar manner to the P-waves, the S-wave has a unique angular dependence. It also interferes with the P-wave amplitudes, which adds four terms to the decay time- and angular-dependent decay rate, including the S-wave term itself. The inclusion of S-wave into the discussion means that more generically $B_s^0 \to J/\psi K^+K^-$ decays are measured, rather than exclusively $B_s^0 \to J/\psi\phi(1020)$ decays.

Including the contributions from the polarization amplitudes, the decay time- and angular-dependent decay rates can be expressed as [26],

$$\frac{d^4\Gamma_{B_s^0\to J/\psi K^+K^-}}{dtd\Omega} = \sum_{k=1}^{10} N_k f_k(\Omega) h_{k,B_s^0}(t) \tag{2.55}$$

$$\frac{d^4\Gamma_{\overline{B}_s^0\to J/\psi K^+K^-}}{dtd\Omega} = \sum_{k=1}^{10} N_k f_k(\Omega) h_{k,\overline{B}_s^0}(t) \tag{2.56}$$

where $\Omega \equiv (\cos(\theta_K), \cos(\theta_\mu), \phi)$; N_k ($[N_k] = E^2$) are defined in Tab. 2.1; $f_k(\Omega)$ are the angular distributions, defined in Tab. 2.1; $h_{k,B_s^0}(t)$ and $h_{k,\overline{B}_s^0}(t)$ are decay rates for B_s^0-mesons with initial B_s^0 and $\overline{B}_s^0$ flavours respectively. The decay rates $h_{k,B_s^0}(t)$ and $h_{k,\overline{B}_s^0}(t)$ are given by,

$$\begin{aligned} h_{k,B_s^0}(t) = \frac{3}{4\pi}\frac{1}{1+C} e^{-\Gamma_s t}\Big(a_k \cosh\Big(\frac{\Delta\Gamma_s t}{2}\Big) + b_k \sinh\Big(\frac{\Delta\Gamma_s t}{2}\Big) \\ + c_k \cos(\Delta m_s t) + d_k \sin(\Delta m_s t)\Big) \end{aligned} \tag{2.57}$$

$$\begin{aligned} h_{k,\overline{B}_s^0}(t) = \frac{3}{4\pi}\frac{1}{1+C} e^{-\Gamma_s t}\Big(a_k \cosh\Big(\frac{\Delta\Gamma_s t}{2}\Big) + b_k \sinh\Big(\frac{\Delta\Gamma_s t}{2}\Big) \\ - c_k \cos(\Delta m_s t) - d_k \sin(\Delta m_s t)\Big). \end{aligned} \tag{2.58}$$

The a_k, b_k, c_k, d_k factors can be found in Tab. 2.1.

2.4.1 $\mathcal{CP}$-asymmetry

Similar to Eq. 2.21, the time-dependent $\mathcal{CP}$-asymmetry $A_{CP}(t)$ is defined as,

$$A_{CP}(t) \equiv \frac{d\Gamma(\overline{B}_s^0 \to J/\psi\,\phi)/dt - d\Gamma(B_s^0 \to J/\psi\,\phi)/dt}{d\Gamma(\overline{B}_s^0 \to J/\psi\,\phi)/dt + d\Gamma(B_s^0 \to J/\psi\,\phi)/dt}. \tag{2.59}$$

From inserting Eq. 2.44 and Eq. 2.45, it follows, that,

$$A_{CP}(t) = \frac{-C_f \cos(\Delta m_s t) - S_f \sin(\Delta m_s)}{\cosh(\Delta\Gamma_s t/2) + D_f \sinh(\Delta\Gamma_s t/2)}. \tag{2.60}$$

Ignoring the direct $\mathcal{CP}$-asymmetry in $B_s^0 \to J/\psi\,\phi(1020)$, $C_f \approx 0$ and $A_{CP}(t) \approx \sin\phi_s \sin(\Delta m_s t)$.

k	N_k	f_k	a_k	b_k	c_k	d_k
1	$\|A_0\|^2$	$2\cos^2(\theta_K)\sin^2(\theta_\mu)$	1	D	C	S
2	$\|A_{\|\|}\|^2$	$\sin^2(\theta_K)(1-\sin^2(\theta_\mu)\cos^2(\phi))$	1	D	C	S
3	$\|A_\perp\|^2$	$\sin^2(\theta_K)(1-\sin^2(\theta_\mu)\sin^2(\phi))$	1	$-D$	C	$-S$
4	$\|A_\perp A_{\|\|}\|$	$\sin^2(\theta_K)\sin^2(\theta_\mu)\sin(2\phi)$	$C\sin(\delta_\perp-\delta_{\|\|})$	$-S\cos(\delta_\perp-\delta_{\|\|})$	$\sin(\delta_\perp-\delta_{\|\|})$	$D\cos(\delta_\perp-\delta_{\|\|})$
5	$\|A_0 A_{\|\|}\|$	$\frac{1}{\sqrt{2}}\sin(2\theta_K)\sin(2\theta_\mu)\cos(\phi)$	$\cos(\delta_{\|\|}-\delta_0)$	$D\cos(\delta_{\|\|}-\delta_0)$	$C\cos(\delta_{\|\|}-\delta_0)$	$S\cos(\delta_{\|\|}-\delta_0)$
6	$\|A_0 A_\perp\|$	$-\frac{1}{\sqrt{2}}\sin(2\theta_K)\sin(2\theta_\mu)\sin(\phi)$	$C\sin(\delta_\perp-\delta_0)$	$-S\cos(\delta_\perp-\delta_0)$	$-\sin(\delta_\perp-\delta_0)$	$D\cos(\delta_\perp-\delta_0)$
7	$\|A_S\|^2$	$\frac{2}{3}\sin^2(\theta_\mu)$	1	$-D$	C	$-S$
8	$\|A_S A_{\|\|}\|$	$\frac{\sqrt{6}}{3}\sin(\theta_K)\sin(2\theta_\mu)\cos(\phi)$	$C\cos(\delta_{\|\|}-\delta_S)$	$-S\sin(\delta_{\|\|}-\delta_S)$	$\cos(\delta_{\|\|}-\delta_S)$	$D\sin(\delta_{\|\|}-\delta_S)$
9	$\|A_S A_\perp\|$	$-\frac{\sqrt{6}}{3}\sin(\theta_K)\sin(2\theta_\mu)\sin(\phi)$	$\sin(\delta_\perp-\delta_S)$	$-D\sin(\delta_\perp-\delta_S)$	$-C\sin(\delta_\perp-\delta_S)$	$-S\sin(\delta_\perp-\delta_S)$
10	$\|A_S A_0\|$	$\frac{4}{\sqrt{3}}\cos(\theta_K)\sin^2(\theta_\mu)$	$C\cos(\delta_S-\delta_0)$	$-S\sin(\delta_S-\delta_0)$	$\cos(\delta_0-\delta_S)$	$D\sin(\delta_0-\delta_S)$

Table 2.1: The $N_k, f_k, a_k, b_k, c_k, d_k$ in Eq. 2.55-2.58 with $C=\frac{1-|\lambda|^2}{1+|\lambda|^2}$; $S=\frac{2|\lambda|\sin(\phi_s)}{1+|\lambda|^2}$; $D=-\frac{2|\lambda|\cos(\phi_s)}{1+|\lambda|^2)}$.

2.5 Looking for new phenomena with $\mathcal{CP}$-violating phase ϕ_s

The effective $\mathcal{CP}$-violating phase ϕ_s, measured in the experiment, is,

$$\phi_s{}^{\text{eff}} = -2\beta_s + \Delta\phi_s{}^{\text{SM}} + \phi_s{}^{\text{BSM}} \tag{2.61}$$

where β_s describes the tree-level contribution, see Fig. 2.7(a); $\Delta\phi_s{}^{\text{SM}}$ is the penguin contributions to ϕ_s, see for example Fig. 2.7(c); $\phi_s{}^{\text{BSM}}$ is the possible contribution to ϕ_s from beyond the SM processes.

The state-of-the-art SM model prediction of the tree-level contribution is obtained from a global fit of the CKM parameters [14],

$$\phi_s{}^{SM} = -0.03686^{+0.00093}_{-0.00067} \text{ rad}, \tag{2.62}$$

where uncertainties on both experimental and theoretical inputs are included.

The $\Delta\phi_s{}^{\text{SM}}$ within Standard Model is predicted to be [6],

$$\Delta\phi_s{}^{\text{SM}} = 0.003^{+0.0010}_{-0.0012} \text{ rad}, \tag{2.63}$$

from the combined analysis of $B^0 \to J/\psi\pi^0$, $B^0 \to J/\psi K^0$, $B^0 \to J/\psi\rho^0$, $B_s^0 \to J/\psi K_S^0$ and $B_s^0 \to J/\psi\phi(1020)$. Due to its smallness, it is common to ignore $\Delta\phi_s{}^{\text{SM}}$ contribution when discussing experimental results. Assuming one can ignore $\Delta\phi_s{}^{\text{SM}}$, if a deviation from value 2.62 is observed experimentally, it must be attributed to the beyond the SM process.

2.6 Polarization-dependent $\mathcal{CP}$-violation

The penguin contributions introduce polarization dependence in Eq. 2.52 such that both the magnitude and phase of λ_f are polarization dependent and $\phi_s{}^f = -2\beta_s + \Delta\phi_s{}^{SM,f} + \phi_s{}^{BSM}$. The only difference between the polarization-independent decay rate in Eq. 2.55 and polarization-dependent decays rate is introduction of $|\lambda_0|, |\lambda_\parallel|, |\lambda_\perp|, |\lambda_S|$ and $\phi_s{}^0, \phi_s{}^\parallel, \phi_s{}^\perp, \phi_s{}^S$. The polarization-dependent $\mathcal{CP}$-violation parameters are also measured in this thesis. This indirectly probes the validity of the assumption that the penguin contributions can be ignored.

3 Detecting $B_s^0 \to J/\psi\, K^+K^-$ decays with the LHCb detector

Measuring the $\mathcal{CP}$-violating phase ϕ_s is the main objective of this thesis. As mentioned in Sec. 2.5, the best decay to access ϕ_s is $B_s^0 \to J/\psi\, K^+K^-$, where the J/ψ is reconstructed from its decay to muons: $J/\psi \to \mu^+\mu^-$.

The sample of $B_s^0 \to J/\psi\, K^+K^-$ candidates for this thesis was collected with the LHCb detector between 2015 and 2018, a period denoted as Run 2. The LHCb detector is one of the four main experiments at the Large Hadron Collider, located at CERN. The Large Hadron Collider is a circular collider with 27 km circumference, which collides protons and heavier ions [27]. The data sample for this thesis was collected from p- p collisions at 13 TeV center-of-mass energy. The amount of data collected is quantified by the *integrated (over time) luminosity*. Luminosity is the ratio of the number of measured interactions over a period of time to the scattering cross-section of two beams. The total integrated luminosity presented in this thesis is $5.7\,\mathrm{fb}^{-1}$.

In this section, first, the requirements for measuring ϕ_s with $B_s^0 \to J/\psi\, K^+K^-$ decays are presented. Second, a comprehensive description of the LHCb detector and its trigger system. Lastly, the previous measurements of ϕ_s are described.

3.1 Requirements for the $B_s^0 \to J/\psi\, K^+K^-$ measurement

The B_s^0 and $\overline{B}_s^0$ mesons are produced in the p-p collisions at LHC. The number of collected B_s^0 mesons depends on the instantaneous luminosity, the duration of data collection and the production cross-section. The luminosity

is defined by the beam characteristics of LHC [27]. The production cross-section is energy-dependent [28]. The LHCb detector was designed to measure about one to two *p-p* interactions simultaneously. The number of visible *p-p* interactions, μ, in Run 2 is 1.3 per one crossing of p bunches. An increase in μ leads to an increase in the occupancy of the detector. When operating at a large occupancy, the LHCb data collection rate drops. To avoid inefficiency in the detector performance, the value of μ is tuned.

The momenta that B_s^0 mesons gain from the excess of energy above the necessary energy to produce the B_s^0 meson, increases the distance that B_s^0 mesons travel. The increased distance improves the experimental separation between the interaction point of the protons, which is also the origin vertex of B_s^0, also known as *primary vertex*, and the decay vertex of B_s^0, also known as *secondary vertex*. This benefits decay time measurements, which require a measurement of the distance between the primary and secondary vertices of the B_s^0 meson.

Now consider what affects the sensitivity to ϕ_s phase. Naturally, the detector should be able to measure the relevant observables: decay time, and helicity angles, as discussed Chapter 2. Because the $B_s^0 \to J/\psi K^+K^-$ final state contains both kaons and muons (from the J/ψ decay), the detector should be able to distinguish muons and kaons from other particles.

Most sensitivity for ϕ_s in Eq. 2.44 comes from the coefficient which multiplies the B_s^0-$\overline{B}_s^0$ oscillation $\sin(\Delta m_s t)$ term. The $\mathcal{CP}$-violating phase ϕ_s enters this coefficient, and, therefore, it is necessary to resolve the rapid B_s^0-$\overline{B}_s^0$ oscillations to measure ϕ_s. Because of the finite decay time resolution of the detector, the reconstructed decay time of B_s^0 is smeared, which dilutes the observed amplitude. Assuming the decay time resolution is Gaussian, the resulting dilution D is [29],

$$D = e^{-\frac{1}{2}\sigma^2 {\Delta m_s}^2}, \tag{3.1}$$

where σ is the width of the Gaussian decay time resolution. Aiming for the effective loss in the size of the data sample to be less than two, the dilution has to be above $\sqrt{0.5}$. Using the current world average on the Δm_s[1], this condition is satisfied if the decay time resolution σ is at least as good as 47 ps. The measured decay time uncertainty is derived by propagating the uncertainty on the decay length σ_L and the momentum uncertainty σ_p [25],

$$\sigma_t = \Big(\frac{m_{B_s^0}}{p_{B_s^0}}\Big)^2 \sigma_L^2 + \Big(\frac{t_{B_s^0}}{p_{B_s^0}}\Big)^2 \sigma_p^2 + \text{correlation}, \tag{3.2}$$

[1]The current world average is $\Delta m_s = 17.756 \pm 0.005\,\text{ps}^{-1}$ [20].

where $m_{B_s^0}$ is the B_s^0 mass; $p_{B_s^0}$ is the B_s^0 momentum and $t_{B_s^0}$ is the B_s^0 decay time. The measured decay time uncertainty of the LHCb detector is defined mainly by the decay length uncertainty for decay times up to a few B_s^0 lifetimes [25]. The decay length L of B_s^0 is computed as the distance between the primary and the decay vertices projected over the B_s^0 momentum,

$$\vec{L} = \frac{(\vec{x}_{\text{decay}} - \vec{x}_{\text{PV}}) \cdot \vec{p}_{B_s^0}}{|\vec{p}_{B_s^0}|}, \tag{3.3}$$

where $\vec{x}_{\text{PV}}$ is the primary vertex position; $\vec{x}_{\text{decay}}$ is the decay vertex position and $\vec{p}_{B_s^0}$ is the B_s^0 meson momentum. In practice, however, the decay length is taken from the full decay chain fit. Equation 3.3 is used for the qualitative discussion on the decay length. The decay length uncertainty is limited by the primary and decay vertices resolutions, where the decay vertex resolution usually contributes the most to the decay time uncertainty. The secondary vertex resolution depends on the multiple scattering in the detector material; the detector hits spatial resolution and the opening angles of outgoing particle trajectories. In Sec. 4.5.1 it is shown that, on average, the decay time uncertainty is $42 - 46$ fs for $B_s^0 \to J/\psi K^+K^-$ candidates collected in 2015 - 2018.

Before being able to determine the properties of $B_s^0 \to J/\psi K^+K^-$ decays, one must first recognize them and separate them from other background processes coming from the p-p collisions. Every particle is identified with two key properties: their mass and average lifetime. The B_s^0 meson average lifetime is $\tau = 1.521 \pm 0.005$ ps [20]. This means that on average, a boosted B_s^0 meson with 80 GeV/c momentum will travel for about 7 mm from the p-p collision point before decaying. This travel path is used to separate the final state particles produced in the B_s^0 decay from the particles that originate directly from the collision point, called *prompt particles*. The compatibility of the particle with the collision point is quantified with the impact parameter IP. The impact parameter is defined as the distance of the closest approach between the B_s^0 daughter tracks and the primary vertex. A simplified two-dimensional IP is shown in Fig. 3.1. It is used to separate the B_s^0 mesons (and other relatively long-lived particles) from the prompt particles. Hence, the resolution on IP should be as good as possible. In the simplified case of the two-dimensional IP, see Fig. 3.1, the covariance matrix V_{IP_x} is derived as [25],

$$V_{\text{IP}_x} = V_{xx} + V_{xx}^{\text{PV}} + t_x^2 V_{zz}^{\text{PV}} - 2t_x V_{xz}^{\text{PV}} + (z - z_{\text{PV}})^2 V_{t_x t_x} - 2(z - z_{\text{PV}}) V_{x t_x}, \tag{3.4}$$

where IP_x is the (x, z) plane impact parameter; V is the track parameters covariance matrix; V^{PV} is the primary vertex covariance matrix; t_x is the slope of the track; z_{PV} is the primary vertex z-position and z is the z-position

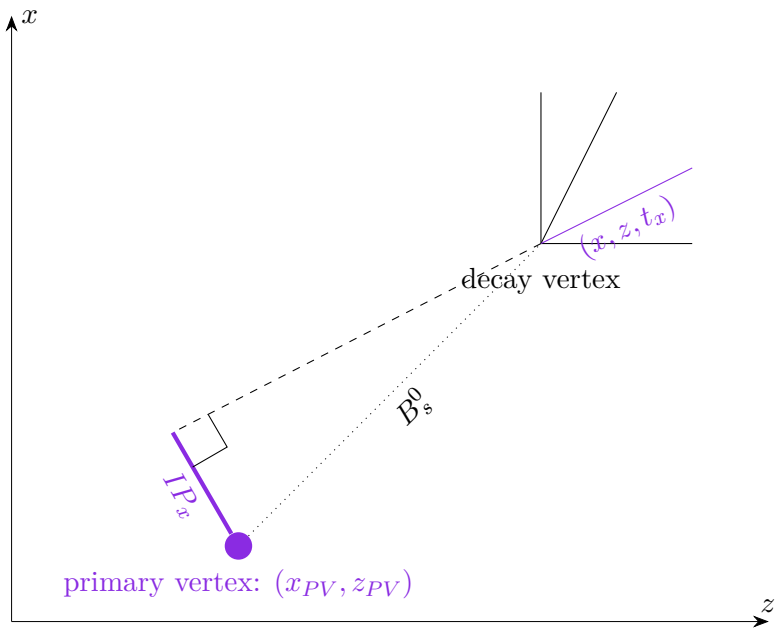


Figure 3.1: The two-dimensional impact parameter in (x, z) plane, where z is aligned with beam axis and x is perpendicular to it.

of the tracks first state. Consider an approximation of a two-body decay for B_s^0, that travels 7 mm at 0.1 radangle with respect to the beam axis. The typical impact parameter in the lab frame for J/ψ that decays under angle $\pi/3$ in the rest frame B_s^0, is $\mathrm{IP}_x \approx 0.8$ mm. Assuming that one needs to know the IP_x with the 10% relative precision, the resolution on IP_x should be as good as 80 µm.

The $B^0 \to J/\psi K^+K^-$ decay is topologically almost identical to $B_s^0 \to J/\psi K^+K^-$, which makes it hard to distinguish it from $B_s^0 \to J/\psi K^+K^-$. The average lifetime and mass of the B^0 meson are close to the B_s^0 meson: $\tau_{B^0} = 1.519 \pm 0.004$ ps and $m_{B^0} = m_{B_s^0} - (87.42 \pm 0.14)$ [20]. Good mass resolution is needed to separate the B^0 and B_s^0 mass peaks. Defining as "sufficient", the separation required such that the mean of the mass peaks is at least five times the mass resolution apart, which yields a maximum resolution of 17 MeV/c^2. As before, the derived mass uncertainty is used as a proxy for the true mass resolution. Using a two-body decay approximation,

the mass uncertainty σ_m of B_s^0 mass is derived as [30],

$$\sigma_m(B_s^0) = \sqrt{\left(\frac{\sigma_{pJ/\psi}}{p_{J/\psi}}\right)^2 + \left(\frac{\sigma_{p\phi}}{p_\phi}\right)^2 + \left(\frac{\sigma_{\theta_{J/\psi\,\phi}}}{1-\cos\theta_{J/\psi\,\phi}}\right)^2} \frac{m^2_{B_s^0} - m^2_{J/\psi} - m^2_{\phi(1020)}}{2m_{B_s^0}}, \tag{3.5}$$

where pJ/ψ and $\sigma_{pJ/\psi}$ are the J/ψ momentum and uncertainty on the momentum; $p\phi$ and $\sigma_{p\phi}$ are the $\phi(1020)$ momentum and uncertainty on the momentum; $\theta_{J/\psi\,\phi}$ and $\sigma_{\theta_{J/\psi\,\phi}}$ are the angle between J/ψ and $\phi(1020)$ and uncertainty on the angle; $m_{B_s^0} = 5366.92 \pm 0.10$ MeV/c^2 is the B_s^0 mass; $m_{J/\psi} = 3096.900 \pm 0.006$ MeV/c^2 is the J/ψ mass and $m_{\phi(1020)} = 1019.461 \pm 0.016\,\text{MeV}/c^2$ is the $\phi(1020)$ mass [20]. Assuming that $\frac{\sigma_{pJ/\psi}}{p_{J/\psi}} \approx \frac{\sigma_{p\phi}}{p_\phi} \equiv \frac{\sigma_p}{p}$ and angular resolution can be ignored, the necessary momentum uncertainty to achieve $\sigma_m = 17$ MeV/c^2 is $\frac{\sigma_p}{p} \approx 0.7\%$.

To measure ϕ_s it is necessary to know the initial flavour of B_s^0-meson produced in the p-p collision. Both $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ decay have the same final state $J/\psi K^+K^-$, which does not allow to identify experimentally whether $J/\psi K^+K^-$ came from B_s^0 or $\overline{B}_s^0$ decaying and if the said $B_s^0/\overline{B}_s^0$ are in the flavor state at the production (initial) or a result of the oscillation. This creates an issue because the ϕ_s-sensitive terms enter the time-dependent decay rates of B_s^0 and $\overline{B}_s^0$ in Eq. 2.44 with opposite signs. Because the $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ are not recognizable by their final state, their decay time distribution would add and the most ϕ_s-sensitive terms would cancel out. Luckily, a set of flavour tagging algorithms is developed in LHCb, which allows the separation between two distributions [31–33]. These algorithms are described in detail in the context of measuring ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays in Sec. 4.4. Here, only the detector requirements are discussed. The flavour tagging requires good particle identification and vertex reconstruction. The flavour tagging algorithm performance is expected to degrade with an increase in the number of primary vertices and multiplicity of charged particles [31].

To sum up, these are the requirements to measure ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays:

1. The B_s^0 yield per unit of time must be sufficiently large to get a sample of $\mathcal{O}(100)$ flavour tagged events;
2. The decay times and angles can be measured, and the final state particles can be identified.
3. It is possible to resolve the rapid oscillations of B_s^0-$\overline{B}_s^0$;
4. It is possible to distinguish initial B_s^0 and $\overline{B}_s^0$ flavours;

3.2 LHCb detector in 2015-2018

The LHCb detector is a forward-arm spectrometer located at the LHC [8]. It was designed for forward-boosted b-quark physics and covers the pseudorapidity range of $2 < \eta < 5$. Despite being optimized for b-quark physics, the modern LHCb physics program covers multiple topics, including charm physics, spectroscopy, exotic particle searches, electroweak physics, Higgs physics, and QCD with heavy-ion collisions [34]. Figure 3.2 shows a schematic view of the LHCb detector.

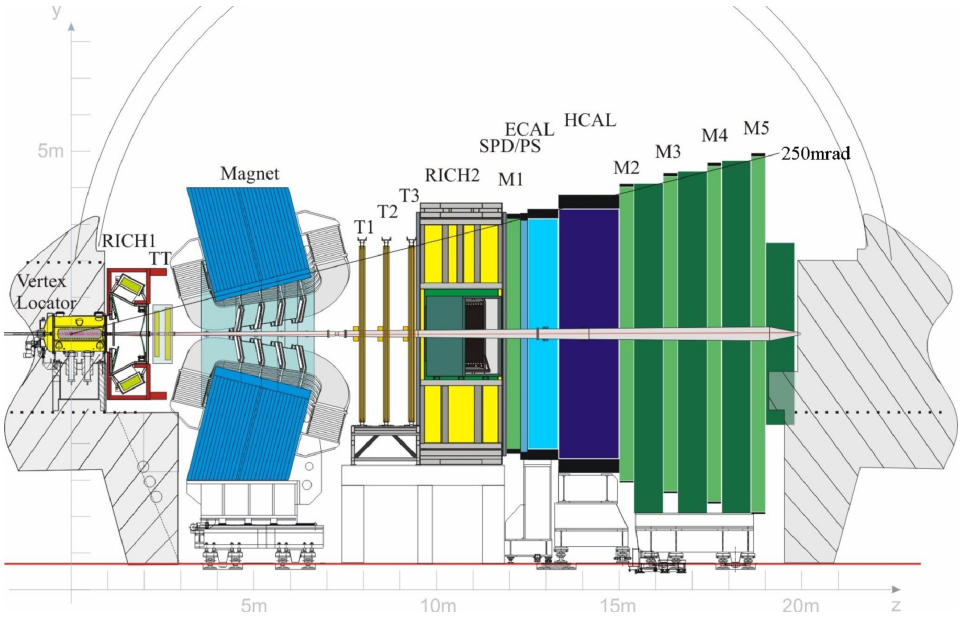


Figure 3.2: The schematic view of the LHCb detector. The collision point is at $z = 0$. From left to right: the Vertex Locator (VELO, tracking system); the first Ring-Imaging Cherenkov detector (RICH1, particle identification system); the Tracker Turicensis station (TT, tracking system); the Magnet; Tracking stations (T1-T3, tracking system); the second Ring-Imaging Cherenkov detector (RICH2, particle identification system); the first Muon station (M1, muon system); the Scintillating Pad detector and the pre-shower (SPD/PD, calorimeter system); the electromagnetic calorimeter (ECAL, calorimeter system); the hadron calorimeter (HCAL, calorimeter system); muon stations (M2-M5, muon system). Taken from [8].

3.2.1 Tracking system: Vertex Locator

The tracking system comprises three subdetectors: Vertex Locator, Tracker Turicensis, and T-stations[2]. The tracking stations around the collision point play a crucial role in the measurement of ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays because they define the quality of the vertex reconstruction. The tracking stations that surround the interaction points are called the Vertex Locator or simply VELO; see Fig. 3.3 [35].

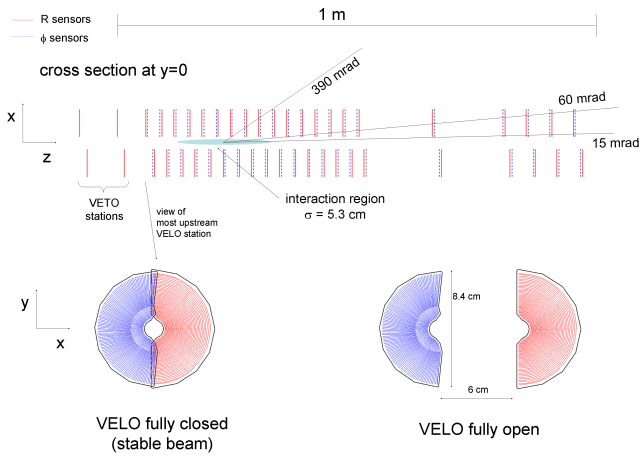


Figure 3.3: The schematic view of the VELO detector. The 21 VELO stations are divided into two modules, each containing r- and ϕ- sensors. The VELO "open" and "close" positions are also shown. Taken from [35].

The inner strips of the VELO modules are about 8.2 mm from the beam axis, and the entire detector is about 1 m long. As a safety measure, the VELO stations are moved out from the beam axis when the beam is unstable and moved in only after the LHC accelerator declares the stable beam operation conditions, see Fig. 3.3. Therefore, the VELO stations are split into two halves, left and right, to allow this movement. To protect the VELO from the beam-induced electrical fields and protect the LHC vacuum from the VELO outgassing, the VELO is enclosed in a so-called RF-box - a thin aluminum box. Each station hosts two silicon-strip sensors, r and ϕ, named after the strip orientation. The r-sensor strips have a radial

[2]Note, that Muon stations are also tracking detectors. However, they are described separately due to their importance for the hardware trigger and the muon particle identification

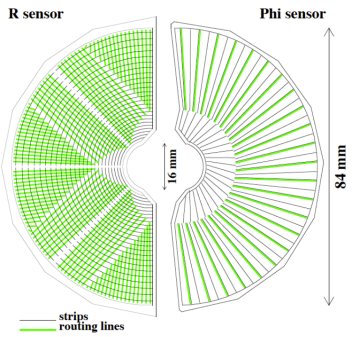


Figure 3.4: The schematic view of the VELO module, where the silicon strip r- and ϕ-sensors are shown. Taken from [36].

orientation, and ϕ-sensor strips have an azimuth orientation; see Fig. 3.4.

The Vertex Locator measures the first hits of the charged particle trajectories, called *tracks*, which are used to reconstruct the primary vertex. Figure 3.5 shows spatial primary vertex resolution for the plane transverse to the beam axis. The primary vertex resolution depends on the number of

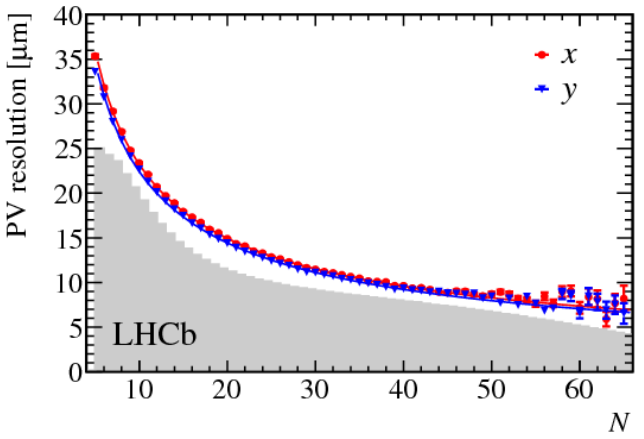


Figure 3.5: Resolution on x- and y-coordinates of reconstructed primary vertex (PV) with respect to the number of tracks (N) coming out of the primary vertex. (x, y) is a plane perpendicular to the beam axis. Taken from [25].

tracks in the primary vertex. The more tracks the primary vertex is reconstructed with, the better the resolution on its position is. Because the primary vertex is required to be reconstructed from at least five tracks, its resolution

is usually better than the secondary (decay) vertex of $B_s^0 \to J/\psi K^+K^-$ decay, which only has four tracks originating from it. The primary vertex resolution is about 70 μm along the beam axis (z-coordinate) and on average 13 μm in the transverse plane (x- and y-coordinate) for a representative primary vertex with 25 outgoing tracks [25].

The impact parameter resolution heavily depends on the p_{T} of the particle, as shown in Fig. 3.6. For the particles with $p_{\mathrm{T}} > 1$ GeV/c the resolution

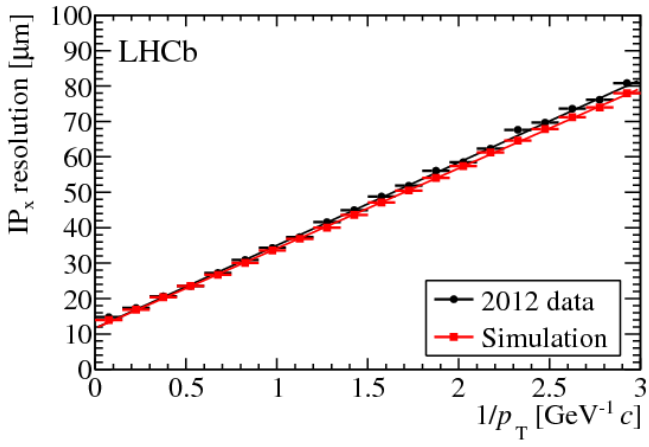


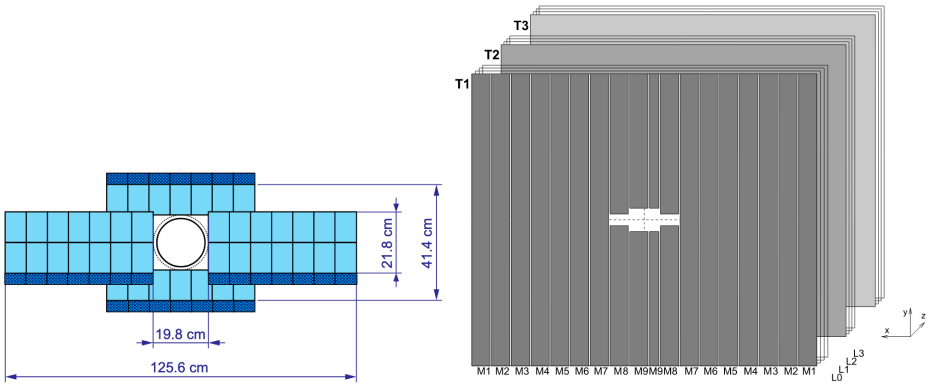
Figure 3.6: The IP_x resolution versus the inverse of p_{T}. Taken from [25].

on IP_x becomes better than 35 μm, reaching asymptotically 13 μm at the very high-transverse momentum [25].

3.2.2 Tracking system: TTa-b and T1-3

The Tracker Turicensis, or TTa-b stations, are located before the magnet [8, 37], see Fig. 3.2. It comprises four silicon strip modules with 200 μm strip width. The main task of the TT-stations is reconstructing the tracks of long-lived particles that decay at the end or outside of the VELO detector, like K_{S}^0 and Λ.

The T1-3 stations are located behind the magnet and reconstruct tracks after they were curved in the magnet [37, 38]. The area closest to the beam is covered by 200 μm silicon strips *the Inner tracker* [37]. All other area is covered by the gaseous detector - *the Outer Tracker* with 5 mm straws [38]. Figure 3.7 shows schematics of both Inner and Outer trackers.



(a) Inner tracker (b) Outer tracker

Figure 3.7: Inner and outer trackers. Taken from [8] and [39].

3.2.3 Magnet

The LHCb dipole magnet is located between the TT-stations and T-stations and has an integrated magnetic field of $4Tm$ for tracks that transverse detector from the vertex locator to calorimeter system [8]. The tracks are deflected when passing through the magnetic field. From the observed curvature, the momentum is estimated. Figure 3.8 shows the main component of the LHCb magnetic field. To reduce the systematic uncertainty on detector

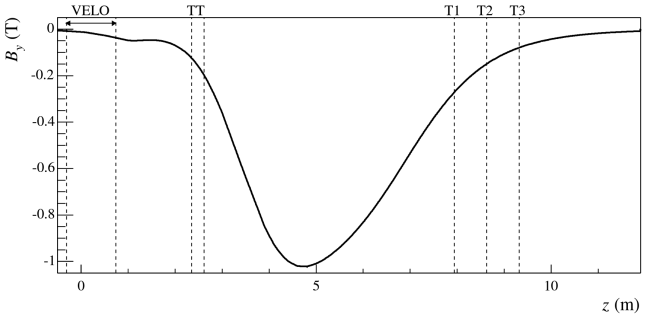


Figure 3.8: The y-component of the magnetic field strength B_y along the z-axis. Taken from [40].

left-right asymmetry and detector alignment the data at LHCb is collected in two regimes by flipping the current in the magnet: “Magnet Up” or MU when the magnetic field is aligned with the y-axis, and “Magnet Down” or

MD when the magnetic field points opposite the y-axis.

Figure 3.9 shows the relative momentum resolution achieved by the LHCb tracking system. Depending on the momentum value, the relative momentum resolution σ_p/p varies between 0.5 % and 1 %.

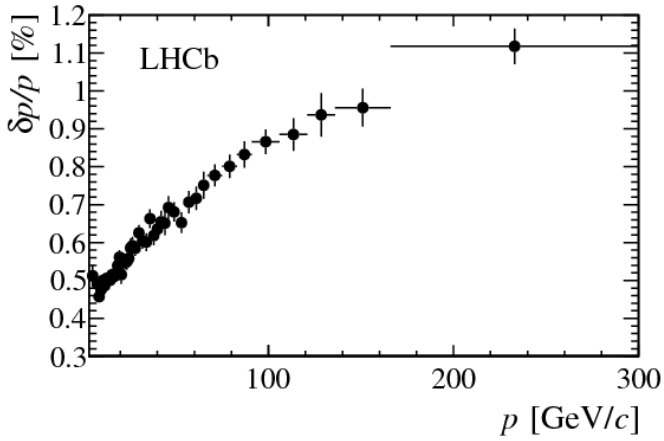


Figure 3.9: Momentum resolution versus momentum. Obtained using $J/\psi \to \mu^-\mu^+$ decays collected at LHCb. Taken from [25].

3.2.4 Particle identification

The $B_s^0 \to J/\psi K^+K^-$ candidates are reconstructed from four charged tracks: two muons that make the J/ψ candidate, and two kaons. The kaons and muons are identified using information from the LHCb particle identification subdetectors, calorimeter system, and muon stations. The identification of kaons helps to separate $B_s^0 \to J/\psi K^+K^-$ decays from the background decays, especially those with pions in the final state. By convention, each charged track is a-priori assumed to be a pion, the most abundant charged particle produced in the p-p collisions at LHC. The likelihood for an alternative particle hypothesis, like a kaon hypothesis $\mathcal{L}_K$, combines likelihoods from particle identification subdetectors $\mathcal{L}_{RICH}$, calorimeter systems $\mathcal{L}_{CAL}$, and muon stations $\mathcal{L}_{MUON}$:

$$\mathcal{L}_K = \mathcal{L}_{RICH} \cdot \mathcal{L}_{CAL} \cdot \mathcal{L}_{MUON}. \tag{3.6}$$

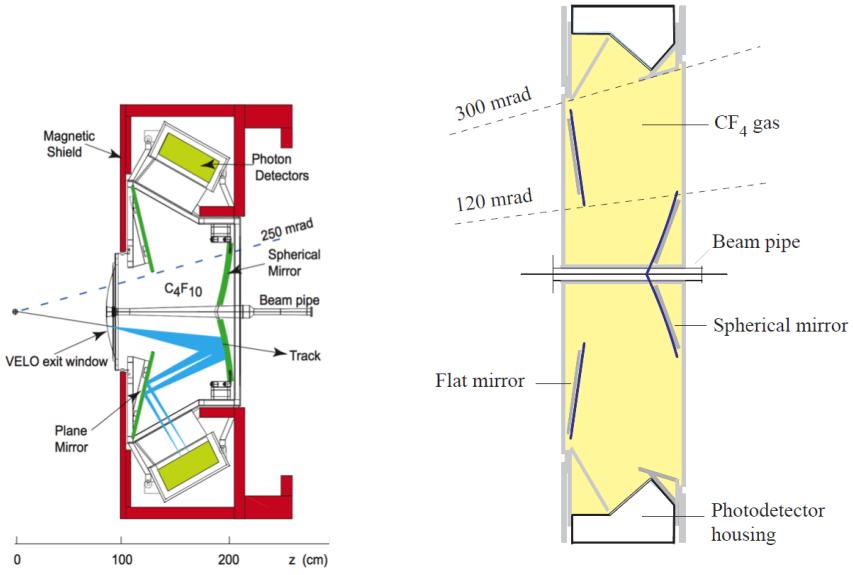
The logarithm of total likelihood to be a kaon $\mathcal{L}_K$ is compared to the logarithm likelihood (log-likelihood) of the pion hypothesis. The log-likelihood

difference of the kaon hypothesis with respect to the pion hypothesis is called $\mathrm{DLL}_{K\pi}$ and is used to identify kaons.

Another class of particle identification algorithms uses a neural-net machine learning algorithm, which also combines information from the particle identification detectors, calorimeters, and muon stations, but also adds tracking system information. The particle identification neural-net output for a particle under the hypothesis that is of type X is called ProbNN_X. In particular, ProbNN_K is a neural-net particle discriminator for charged kaons, and ProbNN_p is a neural-net particle discriminator for protons.

The muons from the J/ψ decays are identified by the hits in the muon stations M2-5. Because muons are minimally interacting particles, they survive beyond the calorimeter, where the M2-5 muon stations are located. Therefore, the presence of hits in the muon stations associated to a charged track is a strong indication of a muon. This information is combined with the calorimeter and particle identification subdetectors into the $\mathrm{DLL}_{\mu\pi}$ hypothesis. Using additionally the tracking system information a neural-net particle discriminator ProbNN_μ is obtained.

Figure 3.2 shows two Ring Image Cherenkov detectors, used for particle identification. The first Ring Image Cherenkov detector, or RICH1, is located after the VELO and before the TT station. It is used to identify charged particles, mainly pions, kaons, and protons in the momentum range below $2 - 60$ GeV/c. The second Ring Image Cherenkov detector or RICH2 is located after the T-stations and identifies the charged particles in the higher momentum range below $15 - 150$ GeV/c. A charged particle emits Cherenkov radiation (light) when it passes through a medium at a speed larger than the speed of light in the same medium. Each of the RICH detectors is filled with specific gas mixtures to create conditions for Cherenkov radiation for particles in different momentum ranges. The emitted light is collected in both detectors with mirror systems onto an array of photodetectors. Figure 3.10 shows the optical systems for both RICH 1 and RICH 2. From the collected light, the opening angle of the Cherenkov light cone is reconstructed. The opening angle depends on the velocity of the particle, which combined with the momentum measurement from the track's curvature gives an estimate of the particle's mass. The estimated mass is used to build a particle identification hypothesis. Figure 3.11 shows the opening angle dependence on momentum for different particle types.



(a) RICH 1 (b) RICH 2

Figure 3.10: Schematics of two Ring Image Cherenkov detectors optical systems. An elaborate system of mirrors is used to collect the emitted Cherenkov light on the photon detectors for measurement. Taken from [41].

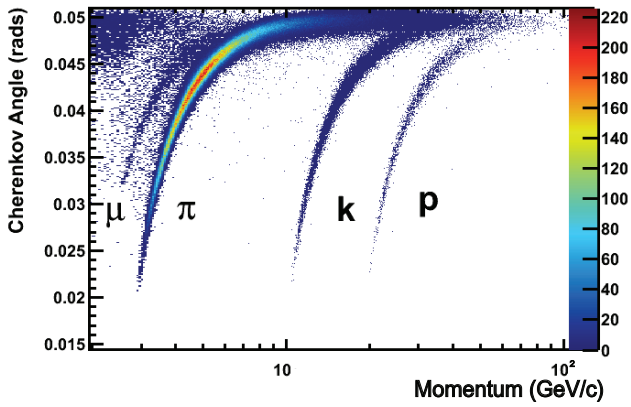


Figure 3.11: The distribution of the momentum of the particle and the opening angle of the Cherenkov light cone in RICH 1. Different particle types are shown. Taken from [42].

3.2.5 Calorimeter system

The primary purpose of the calorimeter system is energy measurement and providing information for the hardware trigger [43]. It also provides the input to the particle identification algorithms and is especially important for photons, π^0, and electrons. The calorimeter system consists of four parts: a pre-shower detector (PS), a scintillating pad detector (SPD), an electromagnetic calorimeter (ECAL), and a hadron calorimeter (HCAL), see Fig. 3.2. Most particles, except for the minimum interacting muons, decay or are stopped within the calorimeter system volume, leaving showers of light behind. These showers are used to reconstruct particle's energy E with $1\% + \frac{10\%}{\sqrt{E}}$ resolution in the electromagnetic calorimeter and $10\% + 80\sqrt{E}\%$ resolution in the hadron calorimeter [44]. The electromagnetic calorimeter comprises lead and scintillating plates arranged in an alternating order, known as "shashlik" [45]. The hadron calorimeter is made of iron-scintillator pads, which also make a "shashlik" structure [45]. The neutral particles are identified by having a response in the calorimeter system without an associated track. In contrast, charged particle showers have an associated track that points toward the shower. The shower shape provides additional information for particle identification. For example, the electron shower would be observed in the scintillating pad detector, pre-shower, and ECAL. The photon shower would only have energy deposited pre-shower and ECAL [43].

3.2.6 Muon system

The muon system is made from the five muon stations M1-5. As was mentioned before, the majority of muons in LHCb are minimally ionizing particles, which means that they propagate through most of the detector with minimal energy loss. The M2-5 stations are used for muon identification and are located at the end of the detector; see Fig. 3.2. The M1 station is only used for the track slope estimation in the muon L0 trigger and is located in front of the calorimeter system [43]. The muon stations are multi-wire proportional chamber tracking detectors [46]. The M2-5 stations have additional iron plates installed to stop hadrons that escape the calorimeter system.

3.2.7 Trigger system in 2015-2018

The raw data from one p-p collision collected within 25 ns period requires $\sim$ 70 kB of storage and is referred to as an "event". The available computing

resources limit the output rate for long-term storage to approximately 10 kHz rate. The p-p bunch crossing rate at LHC, however, is 30 MHz, which is 3×10^3 times larger than this. This puts a strong constraint on the number of collision events that can be saved. The trigger system makes a decision on whether an event is interesting enough to be preserved. In Run 2 the trigger system consisted of three stages: the hardware trigger `L0`; the first stage of the software trigger, denoted by high-level trigger `HLT1`; the second stage of the software trigger, denoted by high-level trigger `HLT2`. Figure 3.12 shows the trigger system of LHCb in Run 2.

The hardware trigger `L0` uses a fixed reconstruction implemented with FPGAs that use information from calorimeter and muon systems to make a decision [43]. It reduces the overall rate from 30 MHz to 1 MHz. If an event passes any of the `L0` triggers, the data from the individual detectors is transmitted to the event builder which accumulates the data from all subdetectors, and then passes the event to the first stage of the software Hlt1 trigger, which performs a partial reconstruction. The `HLT1` reduces the rate to about 150 kHz. Next, the events are collected in a disk buffer. A subset of these events is used to perform calibration. The alignment and calibration are then used in the second stage of the software trigger `HLT2` to fully reconstruct the events. By default, all tracks, vertices, calorimeter showers, etc., are reconstructed for each event, which is referred to as "full event reconstruction". However, if a specific decay is searched for, there is an option to save a subset of reconstructed objects that correspond to this decay [47]. To save disk space, for a subset of events only the reconstructed objects related to this decay are saved, see *"candidate"* in Fig. 3.12. Additionally, some subset of data goes to the calibration samples. If the candidate is reconstructed already in `HLT2`, the dataset can be directly used for the data analysis. However, if collision events are saved fully, then additional processing, known as *"Stripping"*, is needed offline to reconstruct the candidates.

3.3 Previous measurements

The $\mathcal{CP}$-violating phase ϕ_s was measured with $B_s^0 \to J/\psi K^+K^-$ decays by LHCb with $B_s^0 \to J/\psi K^+K^-$ decays ($J/\psi \to \mu^-\mu^+$) in the mass region of $\phi(1020)$ with datasets from Run 1 [48] and partial Run 2 [49]; in the mass region above $\phi(1020)$ resonance [50]; with $B_s^0 \to J/\psi K^+K^-$ decays, where $J/\psi \to e^-e^+$ [51]; with $B_s^0 \to (J/\psi \to \mu^-\mu^+)\pi^-\pi^+$ decays [52]; with $B_s^0 \to (\psi(2S) \to \mu^-\mu^+)K^+K^-$ decays [53]; and with $B_s^0 \to D_s^+D_s^+$

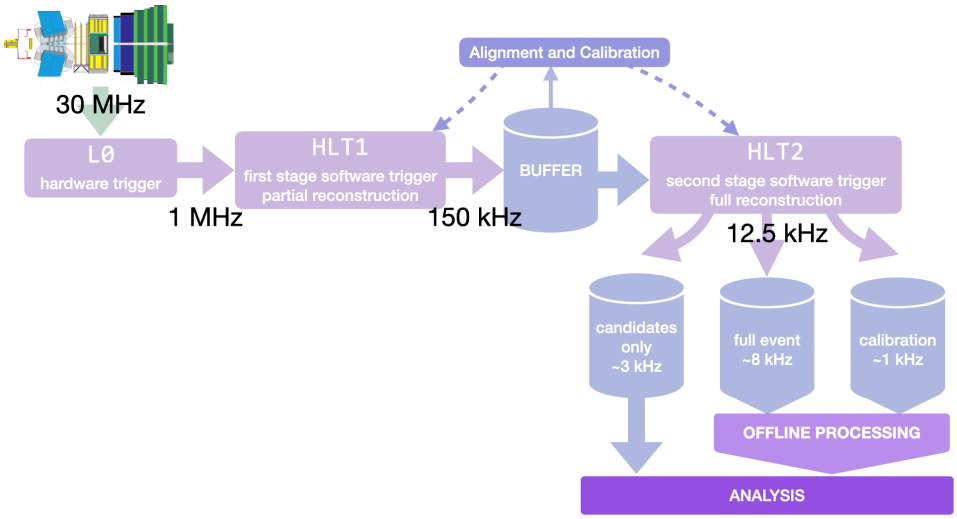


Figure 3.12: LHCb trigger system in 2015-2018 [47].

decays [54]. CDF, D0, ATLAS and CMS also performed measurements with $B_s^0 \to J/\psi K^+K^-$, where $J/\psi \to \mu^-\mu^+$: CDF [55], D0 [56], ATLAS [57], [58], [59] and CMS [60], [61]. They are all summarized in Fig. 3.13, which also shows the world average for ϕ_s.

For all measurements done so far, the precision is statistics-limited. The world average value of ϕ_s is 0.049 ± 0.019 rad [62]. More data is needed to observe a non-zero value of ϕ_s and discover $\mathcal{CP}$-violation in the $B_s^0 \to J/\psi K^+K^-$ decays.

An initial LHCb measurement utilizing $\sim 1.8 = 9$ fb^{-1} of Run 2 data was published in 2019 [49]. Since then, LHCb collected an additional ~ 3.8 fb^{-1}, doubling the sample size. This thesis uses the full Run 2 sample, which corresponds to the total luminosity of 5.7 fb^{-1}. The result presented in this thesis has been first published in [63].

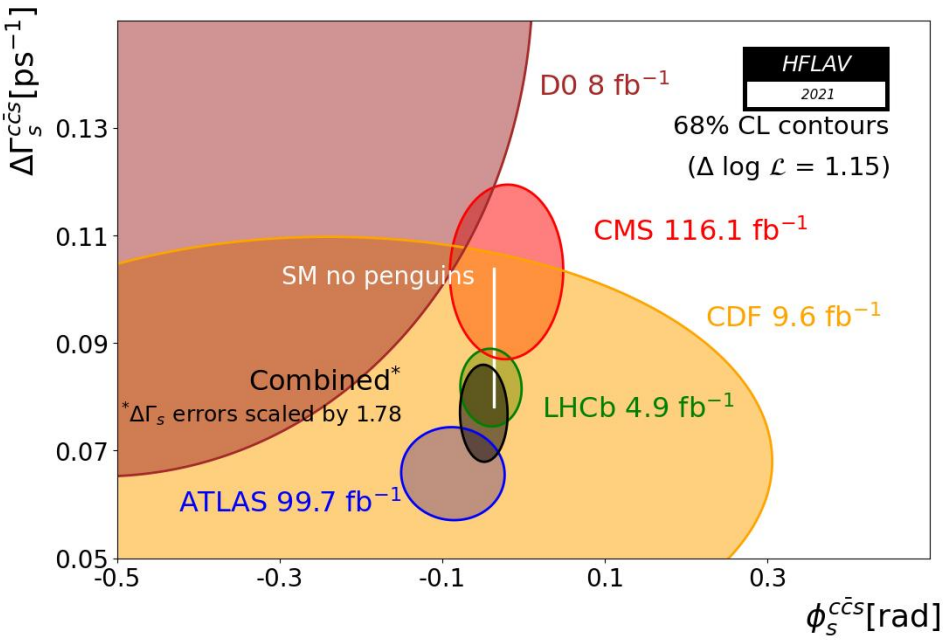


Figure 3.13: The HFLAV group [62] world average of ϕ_s and B_s^0 width split $\Delta\Gamma_s$.

4 Measurement of the $\mathcal{CP}$-violating phase ϕ_s

4.1 Analysis strategy

This analysis aims to measure the $\mathcal{CP}$-violating phase ϕ_s by measuring the decay time dependent difference between the $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ decay rates. The selection procedure and methodology are optimized to enhance sensitivity to ϕ_s. Other parameters of B_s^0 mesons, like the decay width Γ_s, the decay width difference $\Delta\Gamma_s$, and the B_s^0-$\overline{B}_s^0$ oscillation frequency Δm_s, are also accessible with $B_s^0 \to J/\psi K^+K^-$ decays. They are measured simultaneously with ϕ_s.

As it is the difference between $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ decays that is sensitive to ϕ_s, the measurement requires knowing the flavour-specific time-dependent decay rates. This entails measuring the decay time and distinguishing between initial B_s^0 and $\overline{B}_s^0$ mesons. The initial flavour (B_s^0 or $\overline{B}_s^0$) is determined using the flavour tagging tools of LHCb [31–33].

An additional complication arises because the $B_s^0 \to J/\psi K^+K^-$ is not a specific $\mathcal{CP}$-eigenstate decay but a combination of several. As mentioned in Sec. 2.4, because of total angular momentum conservation, different polarization amplitudes contribute to this final state. Due to the differences in the $\mathcal{CP}$-eigenvalues of these amplitudes, their contribution towards $\mathcal{CP}$-asymmetry and thus ϕ_s, enter with different signs, see Eq. 2.52. Consequentially, the sensitivity to ϕ_s is reduced. The contributions from these polarization amplitudes must be statistically resolved, which can be done by measuring the angular distributions. Therefore, the measured decay rate is determined as a function of three helicity angles, Ω. The helicity angles are defined in Fig. 2.9.

In addition, the $B_s^0 \to J/\psi K^+K^-$ decay rate is distorted by imperfect

detector resolution and limited acceptance. The finite decay time resolution leads to a damping factor in front of ϕ_s and, therefore, dilutes the observed value. Acceptances distort distributions of the decay time and helicity angles and can bias the values of the measured physics parameters. The detector effects are accounted for in the model for the observed decay rate.

The maximum likelihood estimation method (MLE) is used to determine ϕ_s and other parameters of interest. MLE estimates the parameters of the assumed model such that they maximize the likelihood of observing the given data. The likelihood corresponding to a probability distribution $P(x|\theta)$ (model) for a given dataset $\vec{x}$ is,

$$\mathcal{L}(\theta|\vec{x}) = \prod_{i=1}^{n} P(x_i|\theta) \tag{4.1}$$

The maximum likelihood estimator $\hat{\theta}$ is then defined as,

$$\hat{\theta} = \arg\max \mathcal{L}(\theta|\vec{x}) \tag{4.2}$$

In practice, $\ln \mathcal{L}(\theta|\vec{x})$ is used instead of $\mathcal{L}(\theta|\vec{x})$, as it is a better match to the floating point representation used by computers. Then the $\prod_{i=1}^{n}$ in Eq. 4.1 becomes $\sum_{i=1}^{n}$, and as ln is a monotonously rising function the maximum for $\ln \mathcal{L}(\theta|\vec{x})$ is the same as of $\mathcal{L}(\theta|\vec{x})$.

The variance V of the maximum likelihood estimator is [64],

$$V_{\hat{\theta}} = -\Big[\frac{\partial^2 \ln \mathcal{L}(\theta|\vec{x})}{\partial \theta^2}\Big|_{\theta=\hat{\theta}}\Big]^{-1}. \tag{4.3}$$

In the limit of the infinite dataset the maximum likelihood estimator is efficient, meaning that it has the lowest possible variance and is unbiased.

To construct the decay time and angular dependent PDF, the B_s^0 and $\overline{B}_s^0$ decay rates from Eq. 2.55 are combined and normalized. The combination is done with the help of the flavour tag q[1]. The flavour tag is not 100% efficient and has an associated probability of being wrong, known as *mistag probability* or η. The non-zero mistag probability reduces sensitivity to ϕ_s and thus effectively reduces the collected sample size. The normalized combination of the $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ decay rates is then,

$$\frac{d^4N}{dtd\Omega} = \frac{1}{\tilde{N}} \sum_{k=1}^{10} N_k \big[(1+q(1-2\eta)) \cdot h_{k,B_s^0}(t) + (1-q(1-2\eta)) \cdot h_{k,\overline{B}_s^0}(t)\big] f_k(\Omega) \tag{4.4}$$

[1] *+1* for tagged as B_s^0, *0* for untagged and *-1* for tagged as $\overline{B}_s^0$

where $q = \pm 1$; η is the mistag probability and $\tilde{N}$ is the normalization, which is defined further in Eq. 4.7.

Including the detector and reconstruction effects such as decay time acceptance, angular acceptances, and decay time resolution, Eq. 4.4 becomes,

$$\frac{d^4N}{dtd\Omega} = \frac{1}{\tilde{N}} \sum_{k=1}^{10} N_k \cdot \big[(1 + q(1-2\eta)) \cdot h_{k,B_s^0}(t) + (1 - q(1-2\eta)) \cdot h_{k,\overline{B}_s^0}(t)\big] \cdot f_k(\Omega) \cdot \varepsilon(t,\Omega) \otimes R(t|\sigma_t) \tag{4.5}$$

where $\varepsilon(t,\Omega)$ is the decay time and angular acceptance and $R(t|\sigma_t)$ is the decay time resolution, which depends on the estimated decay time uncertainty σ_t. The acceptance function $\varepsilon(t,\Omega)$ tells how likely for a candidate with decay time t and angles Ω to pass the selection criteria. The acceptance function takes only two values: zero if the candidate is not accepted and one if the candidate is accepted. The detailed introduction to the acceptances can be found in [65].

Parameters of interest are estimated using Eq. 4.1, which de-facto means differentiating the log-likelihood of Eq. 4.5 with respect to parameters of interest to find its maximum. Efficiencies are independent of the parameters of interest and, therefore, can be dropped from the numerator of the log-likelihood when minimizing. However, they still contribute to the denominator $\tilde{N}$, which contains the integral over decay time and angles. Assuming that the decay time acceptance and the angular acceptance factorize,

$$\varepsilon(t,\Omega) = \varepsilon(t) \cdot \varepsilon(\Omega), \tag{4.6}$$

the denominator can be written as,

$$\tilde{N} = \int dt \sum_{k=1}^{10} \omega_k N_k \cdot \big[(1 + q(1-2\eta)) \cdot h_{k,B_s^0}(t) + (1 - q(1-2\eta)) \cdot h_{k,\overline{B}_s^0}(t)\big] \cdot \varepsilon(t) \otimes R(t|\sigma_t), \tag{4.7}$$

where the angular weights ω_k, represent the angular acceptance $\varepsilon(\Omega)$ and are defined as:

$$\omega_k \equiv \int d\Omega f_k(\Omega)\varepsilon(\Omega) \tag{4.8}$$

As described in Sec. 2.4, the $B_s^0 \to J/\psi K^+K^-$ decay can not only go via the $\phi(1020)$ resonance but also without resonance or via the spin-zero $f_0(980)$ resonance [66]. The former is known as the P-wave, and the latter (non-resonant and $f_0(980)$ decays) is the S-wave. The (relative) contribution of

the S-wave to the observed decay rate naturally depends on the K^+K^- pair mass or, more precisely, the proximity to the $\phi(1020)$ resonance. Therefore, the A_S in Eq. 2.55 is measured in the bins of $m(K^+K^-)$, denoted here m_{KK}. The P-wave amplitudes $A_{P,j}$, where $j \in (0, \|, \perp)$ and S-wave amplitude A_s as functions of m_{KK} are,

$$A_{P,j}(m_{KK}) = A_j \rho(m_{KK}), \text{ for } j \in (0, \|, \perp) \tag{4.9}$$

$$A_S(m_{KK}) = A_S \sigma(m_{KK}), \tag{4.10}$$

where $\rho(m_{KK})$ is P-wave mass line shape and $\sigma(m_{KK})$ is S-wave mass line shape. The interference between line shapes of P-wave and S-wave is accounted for by a set of multiplicative constants C_{SP}, such that

$$N_k = C_{SP} N_k, \text{ for } k \in [s\,\|, s\,\perp, s0]. \tag{4.11}$$

The C_{SP} constants are computed in bins of $m_{KK} \in [990, 1008, 1016, 1020, 1024, 1032, 1050]\,\text{MeV}/c^2$ as,

$$C^i_{\text{SP}} e^{-i\theta^i_{\text{SP}}} = \frac{\int_{m_i}^{m_{i+1}} \rho(m_{KK}) \cdot \sigma^*(m_{KK}) dm_{KK}}{\sqrt{\int_{m_i}^{m_{i+1}} |\rho(m_{KK})|^2 dm_{KK} \int_{m_i}^{m_{i+1}} |\sigma(m_{KK})|^2 dm_{KK}}}, \tag{4.12}$$

where i is an i^{th} bin of m_{KK}, such that $m_{KK} \in [m_i, m_{i+1}]$; θ_{SP} is a phase that is absorbed by the strong phase of S-wave δ_s. Table 4.1 shows the computed C_{SP} constants. The corresponding interference terms ($k \in [s\,\|, s\,\perp, s0]$) in

$m_{KK} \in$, MeV/c^2	$[990, 1008)$	$[1008, 1016)$	$[1016, 1020)$	$[1020, 1024)$	$[1024, 1032)$	$[1032, 1050]$
C_{SP}	0.8458	0.8673	0.8127	0.8558	0.9359	0.9735

Table 4.1: The multiplicative constants in front of P-wave and S-wave interference, called C_{SP}, in different bins of $m(K^+K^-)$.

Eq. 4.5 are multiplied by an effective coupling, depending on the observed m_{KK} of a candidate. The C_{SP} factors are described in detail in [67].

With all the above, Eq. 4.5 becomes,

$$\frac{d^4N}{dtd\Omega} = \frac{1}{\tilde{N}} \sum_{k=1}^{10} N_k \cdot \big[(1 + q(1-2\eta)) \cdot h_{k,B^0_s}(t) + (1 - q(1-2\eta)) \cdot h_{k,\bar{B}^0_s}(t)\big] \cdot f_k(\Omega) \cdot \varepsilon(t) \cdot \varepsilon(\Omega) \otimes R(t|\sigma_t) \tag{4.13}$$

where the normalisation $\tilde{N}$ is defined as,

$$\tilde{N} = \int dt \frac{1}{\tilde{N}} \sum_{k=1}^{10} \omega_k N_k \cdot \left[(1 + q(1-2\eta)) \cdot h_{k,B_s^0}(t) + (1 - q(1-2\eta)) \cdot h_{k,\overline{B}_s^0}(t) \right] \cdot \varepsilon(t) \otimes R(t|\sigma_t) \tag{4.14}$$

where ω_k are weights that model angular acceptance, defined in Eq. 4.70. The parameters of interest are estimated by performing the maximum likelihood fit of Eq. 4.13. Inputs to the fit, like acceptance shapes or decay time resolution shapes, are derived from the control samples.

The analysis steps can be summarised in the following workflow:

1. $B_s^0 \to J/\psi K^+K^-$ decays are selected from the proton-proton LHCb data sample collected during the 2015 - 2018 data-taking years. The selection is the same for all four years, and only a multivariate cut is optimized for each year separately. The selection is described in Sec. 4.2.

2. After the selection is applied, some background contamination is still expected. Therefore, the background subtraction is performed with the *sPlot* method from [68] using the reconstructed mass of $J/\psi K^+K^-$ as a discriminating variable. In this method, each candidate gets a weight, called *sWeight*, that represents a "signal-like" score. The sum of the sWeights is equal to the estimated total number of $B_s^0 \to J/\psi K^+K^-$ decays in the sample. The sWeights are then applied to the maximum likelihood fit [69]. The method essentially corresponds to a statistically optimal sideband subtraction. The method is described in more detail in Sec. 4.3.

3. The initial flavour of B_s^0 and $\overline{B}_s^0$ is estimated using the flavour tagging tools developed by LHCb collaboration. Algorithms are trained on specific training samples. Their output has to be calibrated to be applicable to the selected $B_s^0 \to J/\psi K^+K^-$ sample. This calibration is described in Sec. 4.4.

4. Detector effects modify the measured decay rates. Incorrectly accounting for the decay time and angular acceptances can result in biased parameters of interest. The finite decay time resolution of the detector directly dilutes the value of ϕ_s. In Sec. 4.5, the detector effects are described.

5. The final maximum likelihood fit is performed on the selected sWeighted data. Systematic uncertainties on the parameters of interest are assigned to all the procedures of this analysis, as described in Sec. 4.6. The results are quoted in Sec. 4.8

This analysis was performed as a part of the extended research group and is published under the name "Improved measurement of CP violation parameters in $B_s^0 \to J/\psi K^+K^-$ decays in the vicinity of the $\phi(1020)$ resonance", [63]. The author of this thesis contributed to the statistical background subtraction, decay time resolution, decay time angular maximum likelihood fit, and corresponding systematical uncertainties. Other analysis components, such as selection, acceptance studies, and flavour tagging, are described for completeness. The decay time angular maximum likelihood fit was cross-checked for consistency with independent implementations from collaborators. The author reports results obtained with her version of the likelihood fit in this thesis. The results are consistent, but not identical, with those published in [63].

4.2 $B_s^0 \to J/\psi\, K^+K^-$ candidates selection

The bunch crossing rate of the LHC beams is 40 MHz, which corresponds to a 25 ns period. A set of reconstructed tracks, vertices, and clusters from a 25 ns period is called an *event*. On average, 1.3 *p*-*p* interactions are expected in each *p*-*p* collision event in LHCb. A reconstructed *p*-*p* interaction point is called a primary vertex. Each track in an event has a "best primary vertex" associated with it. The best primary vertex is the primary vertex for which the track has the lowest $\chi^2_{\rm IP}$, where $\chi^2_{\rm IP}$ is defined as follows: the $\chi^2_{\rm IP}$ of track T is the difference between the χ^2 of vertex V with track T removed and the χ^2 of vertex V with track T included. Reconstructed decays that match the decay of interest on which the measurement is performed (signal) are called *signal candidates*. Decays that do not originate from the decay of interest but are selected as signal candidates are called *background*. Two main types of backgrounds pollute the $B_s^0 \to J/\psi\, K^+K^-$ signal sample. The first type, *combinatorial background*, consists of random track combinations. The second type, *peaking background*, are background decays that peak in the B_s^0 invariant mass spectrum but do not originate from $B_s^0 \to J/\psi\, K^+K^-$ decays. The $B_s^0 \to J/\psi\, K^+K^-$ candidate selection consists of four stages applied sequentially:

1. *Online trigger selection* selects *events* that possibly contain signal

candidates during the data taking. The online trigger selection is based on event properties.

2. *Offline stripping selection* reconstructs $B_s^0 \to J/\psi K^+K^-$ candidates in trigger-selected events. Only events with at least one $B_s^0 \to J/\psi K^+K^-$ candidate are saved.

3. *Post-stripping offline selection* increases the purity of signal candidates based on candidate properties.

4. *Misidentified background veto* removes peaking background candidates, accidentally selected as signal candidates due to the wrongly assigned particle identification.

4.2.1 Online trigger selection

For this analysis, events that passed any of the `L0` triggers (Sec. 3.2.7) are considered. At `HLT1` (see Sec. 3.2.7), $B_s^0 \to J/\psi K^+K^-$ candidates are required to pass either the `Hlt1DiMuonHighMass`, `Hlt1TrackMuon` or `Hlt1TwoTrackMVA` line at the HLT1 stage. `Hlt1DiMuonHighMass` selects two oppositely charged muon tracks with high transverse momentum, with a small distance of closest approach and an invariant mass larger than 2900 MeV/c^2. The distance of closest approach between two tracks (`DOCA`) is the shortest distance between a point on the first track and a point on the second track. `Hlt1TrackMuon` selects a displaced muon track with high transverse momentum. `Hlt1TwoTrackMVA` selects two displaced tracks with high transverse momenta. Displaced tracks are selected by requiring high χ^2_{IP} with respect to all reconstructed primary vertices.

At HLT2, defined in Sec. 3.2.7, only events containing candidates that passed `Hlt2DiMuonDetachedJPsi` line are used. `Hlt2DiMuonDetachedJPsi` requires two displaced muon tracks with high transverse momentum to have a mass close to the nominal J/ψ mass. This thesis uses the nominal J/ψ mass of 3096.9 MeV/c^2, [70]. Displaced tracks are selected by high χ^2_{IP}, and significant decay length, where decay length is defined in Eq. 3.3. Decay length significance is decay length divided by its uncertainty.

Requirements on the transverse momenta of B_s^0 daughter tracks introduce signal inefficiency as a function of the angles Ω. Displacement requirements on the B_s^0 daughter tracks introduce signal inefficiency for the smaller values of the B_s^0 decay time. Data that passes through the trigger requirements are divided into two subsamples based on the amount of decay time inefficiency introduced. Candidates that passed

`Hlt1TwoTrackMVA` and `Hlt2DiMuonDetachedJPsi` or `Hlt1TrackMuon` and `Hlt2DiMuonDetachedJPsi` are affected the most and are labeled as the **"biased"** category. Candidates that passed `Hlt1DiMuonHighMass` and `Hlt2DiMuonDetachedJPsi` do not have a displacement requirement at the HLT1 stage but only at the HLT2 stage and fall into the **"almost unbiased"** category. In the later text, they are called **"unbiased"**. In Tab. 4.2, the fraction of $B_s^0 \to J/\psi K^+K^-$ candidates in the "biased" and "unbiased" trigger categories is shown.

	"biased"	"unbiased"	total number of candidates
2015	21.91 ± 0.32 %	78.07 ± 0.73 %	25879
2016	16.81 ± 0.11 %	83.19 ± 0.30 %	171862
2017	16.54 ± 0.11 %	83.46 ± 0.31 %	164140
2018	17.33 ± 0.10 %	82.67 ± 0.28 %	188368

Table 4.2: Fractions of candidates in the "biased" and "unbiased" trigger categories and the total number of selected candidates for all data taking years.

4.2.2 Offline stripping selection

$B_s^0 \to J/\psi K^+K^-$ candidates are built from the events that pass the trigger selection. Two oppositely charged muon tracks are combined into the J/ψ candidates. Muon tracks are long tracks with several hits in the LHCb `MUON` subsystem. Only those muon tracks with $\mathrm{DLL}_{\mu\pi}$ larger than zero and transverse momentum bigger than 500 MeV/c are considered. The dimuon invariant mass must not exceed 80 MeV/c difference from the J/ψ nominal mass. The J/ψ candidates two-track vertices must have χ^2_{vtx} less than 16. The J/ψ candidates are then combined with two oppositely charged kaon tracks. Kaon tracks are long tracks that have an associated `RICH` particle identification input and a loose $\mathrm{DLL}_{K\pi}$ larger than zero. The two kaon tracks are selected if the vertex they form has χ^2_{vtx} less than 25 and an invariant mass in the range $[980, 1060]$ MeV/c around the $\phi(1020)$ resonance. In this mass window, the K^+K^- pair originates from $\phi(1020)$ and other resonances, but for simplicity, the two kaons are collectively called ϕ candidates in the following text. J/ψ candidates and ϕ candidates are combined into a B_s^0 candidates, which are accepted if $\chi^2_{\mathrm{vtx}}/\mathrm{ndof}$ is less than 20, where ndof is 5, and if the B_s^0 candidate invariant mass is in the $[5150, 5550]$ MeV/c range. To reduce the short-lived combinatorial background, the B_s^0 candidates are required to have a reconstructed decay time of at least 0.2 ps. The above

selection is known as the `StrippingBetaSBs2JpsiPhiDetached` line.

4.2.3 Post-stripping offline selection

The post-stripping offline selection can be split into two stages: single-variable requirements and multi-variate requirements. Single-variable requirements tighten the offline stripping selection. The invariant mass of B_s^0 candidates is recomputed with an additional constraint to improve its resolution. Namely, the mass of the two muon tracks is forced to the nominal J/ψ mass. The allowed invariant mass range for B_s^0 candidates is then narrowed to $[5200, 5550]$ MeV/c^2 and for K^+K^- combinations to $[990, 1050]$ MeV/c^2.

In the vertex fit, the B_s^0 candidate is constrained to originate from the primary vertex. The constraint is applied in the kinematics χ^2 fit of $B_s^0 \to J/\psi K^+K^-$ decay chain, which is done by the *Decay Tree Fitter* algorithm [71]. Adding this primary vertex constraint on B_s^0 kinematics improves the angular resolution. Here and further, all B_s^0 kinematics is computed with primary vertex constraint. The primary vertices are reconstructed prior to selecting signal candidates. As a result, the candidate daughter tracks can mistakenly be added to the primary vertex. If this happens, the primary vertex position is biased towards the signal candidate. To remove the bias, tracks that are part of the selected B_s^0 candidates are removed from the primary vertices, and primary vertices are refitted. The above *primary vertex unbiasing* procedure is always applied.

Adding the primary vertex constraint to the decay time computation introduces an edge effect close to the 0.2 ps stripping cut-off. To avoid it, the lower boundary is shifted to 0.3 ps. The upper boundary is set to 15 ps. Candidates with a lifetime above 15 ps are prone to reconstruction errors, like wrong primary vertex association. Including them in the sample can bias the final fit. As the B_s^0 decay time has to be well estimated, only candidates with estimated decay time uncertainty, see Eq. 3.2, below 0.15 ps are considered for the analysis. The final single-variable selection is summarized in Tab. 4.3.

Multi-variate classifiers are powerful tools for signal-background separation. They exploit the discriminating power of multiple variables at the same time. A gradient-boosted decision tree (`BDT`) machine learning technique is used to define a multivariate classifier. `BDT` algorithm selects the important data features, making it a popular tool for data categorization problems. It combines multiple boosted decision trees, each representing a set of data classifiers. Two samples are used for the $B_s^0 \to J/\psi K^+K^-$ `BDT` training. First, a sample of simulated $B_s^0 \to J/\psi \phi(1020)$ events is used as a signal

Table 4.3: Final selection, single-variable cuts for $B_s^0 \to J/\psi K^+K^-$ candidates.

	Variable	Final selection
$J/\psi \to \mu^+\mu^-$	$\Delta\mathcal{L}_{\mu\pi}$ $(\mu^\pm)$	> 0
	p_{T} $(\mu^\pm)$	> 500 MeV/c
	χ^2_{vtx}	< 16
	$m(\mu^+\mu^-)$	$\in [3016.9, 3176.9]\,\mathrm{MeV}/c^2$
$K^+K^- \equiv \phi$	$\chi^2_{\mathrm{track}}/\mathrm{nDoF}$ $(K^\pm)$	< 5
	$\Delta\mathcal{L}_{K\pi}(K^\pm)$	> 0
	p_{T} (K^+K^-)	> 500 MeV/c
	χ^2_{vtx}	< 25
	$m(K^+K^-)$	$\in [990, 1050]\,\mathrm{MeV}/c^2$
$B_s^0 \to J/\psi\phi$	χ^2_{vtx}	< 20
	$m(J/\psi K^+K^-)$	$\in [5200, 5550]\mathrm{MeV/c^2}$
	t	$\in [0.3, 15]$ ps
	σ_t	< 0.15 ps

proxy. As a proxy for combinatorial background, real data candidates are selected from the upper $J/\psi K^+K^-$ invariant mass side band ($\in [5450, 5550]$ MeV/c^2). This is done because the lower mass contains $B^0 \to J/\psi K^+\pi^-$ misidentified background, see Sec. 4.2.4. This background can mimic the signal behavior and, therefore, bias the `BDT` training, which is undesirable. Based on the discriminating power, the following data features are chosen for training: final state particles track χ^2_{trk}/ndof, neural net-based particle identification of final state particles ProbNN_K and ProbNN_μ, vertex χ^2_{vtx} of J/ψ and B_s^0, B_s^0 χ^2_{IP}, the *decay chain fitter* χ^2_{DTF} and transverse momentum of B_s^0 and ϕ, see Fig. 4.1.

The cut on the output of the `BDT` is chosen to maximize the signal significance in the signal mass region $\mathrm{m}_{J/\psi K^+K^-} \in (5320., 5420)$ MeV/c^2. Separate `BDTs` are trained for each year, except 2015. For 2015, the 2016 `BDT` is used.

The event generation and the detector description are imperfect in the LHCb simulation. The imperfections are corrected in the simulation used for the `BDT` training to get the optimal classifier. Corrections are derived from a comparison between simulation and background subtracted real data samples or calibration real data samples.

Neural net particle identification variables are corrected using the `PIDCalib` package of the URANIA framework [72]. For muons, ProbNN_μ

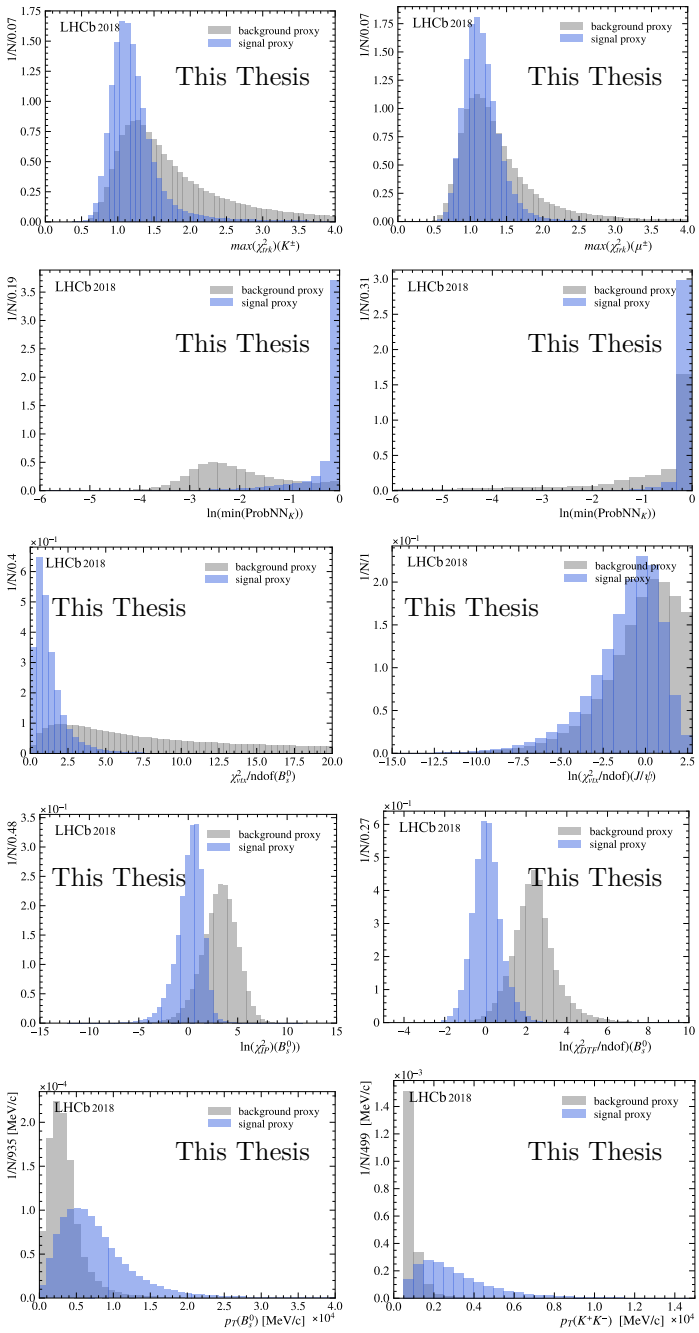
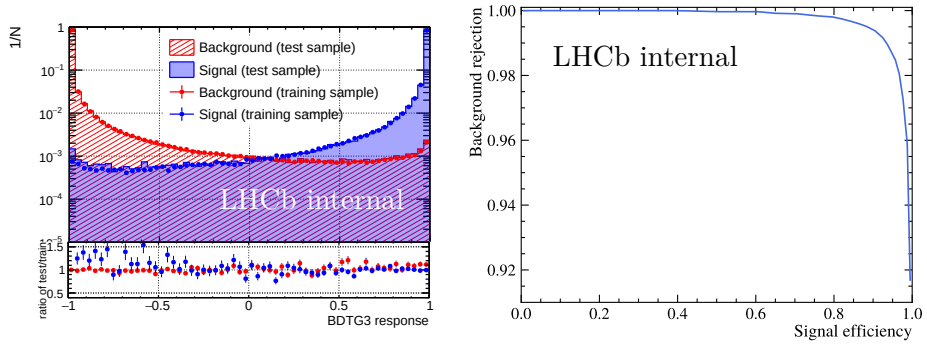


Figure 4.1: *B*oosted *D*ecistion *T*rees input variables for 2018. From left to right: kaon tracks quality ($\chi^2_{trk}(K^\pm)$), muon track quality ($\chi^2_{trk}(\mu^\pm)$); neural net-based particle identification ProbNN_K for kaons and ProbNN_μ for muons; vertex quality of B_s^0 ($\chi^2_{\mathrm{vtx}}(B_s^0)$) and J/ψ ($\chi^2_{\mathrm{vtx}}(J/\psi)$); B_s^0 χ^2_{IP}; quality of decay chain fitter χ^2_{DTF}; transverse momentum of B_s^0 ($p_{\mathrm{T}}(B_s^0)$) and K^+K^- combination ($p_{\mathrm{T}}(K^+K^-)$).



(a) BDT response

(b) ROC curve

Figure 4.2: 2018 BDT output. Taken from [67].

is resampled from particle identification distributions from $J/\psi \to \mu^+\mu^-$ calibration samples. For kaons, it is not enough to correct only ProbNN_K, as in later section ProbNN_π and ProbNN_p of kaon tracks are used for vetoing peaking background and these particles. Particle identification variables are correlated; therefore, the ProbNN_K, ProbNN_π, and ProbNN_p correction have to account for this correlation. However, the resampling procedure described for muons assumes no correlation between the particle identification responses. A different form of correction, called particle identification transformation, is applied for kaons. In this method, ProbNN_K, ProbNN_π, and ProbNN_p are simultaneously transformed so that the cumulative distribution function for the particle identification observable in simulation matches the cumulative distribution function in calibration data [73]. The calibration sample used is $D^{*+} \to D^0(\to K\pi)\pi^+$. The simulated sample, used for the particle identification transformation, already includes a correlation between variables. This correlation is preserved during the transformation. Particle identification transformation is the most applicable technique for multi-variable kaons particle identification correction. The method relies on the parametrization of the particle identification distributions in the simulated sample.
Kinematic distributions in the simulation are corrected using multi-variate weights to match the background subtracted real data distributions.

4.2.4 Misidentified backgrounds veto

The $\mathrm{DLL}_{K\pi}$ observable compares the kaon hypothesis to the pion hypothesis. Requiring $\mathrm{DLL}_{K\pi}(K^\pm) > 0$, about 6% of the selected kaon tracks are expected to originate from pions, see Fig.4.3.

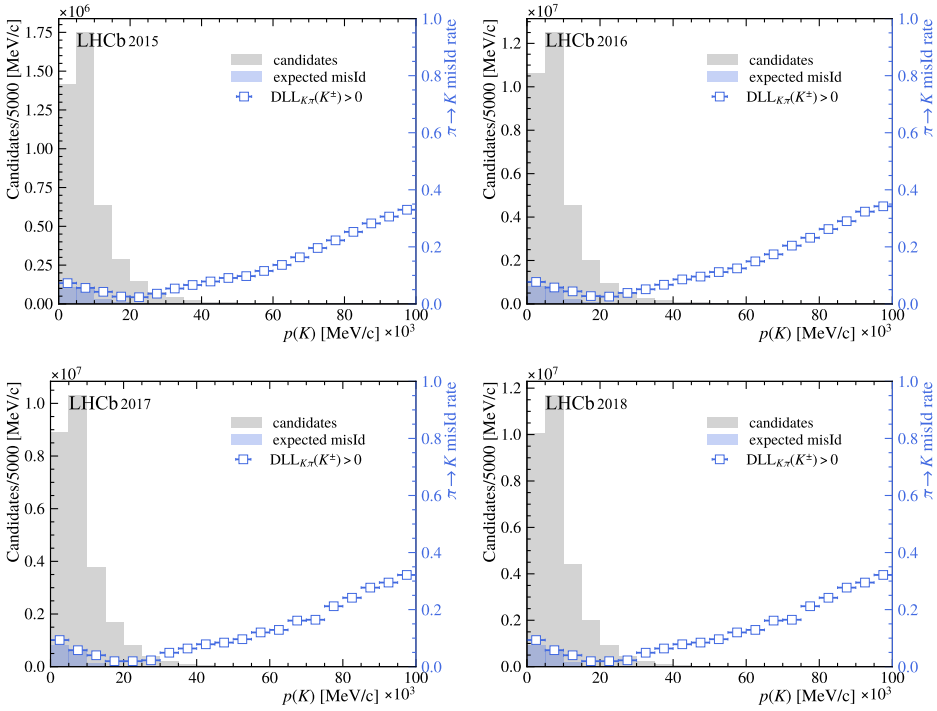


Figure 4.3: π to K misidentification rates for the $\mathrm{DLL}_{K\pi} > 0$ cut from [74](blue boxes), momentum distribution of kaon tracks in $B_s^0 \to J/\psi K^+K^-$ samples before offline selection (grey) and expected distribution of pion tracks misidentified as kaons(blue).

When a misidentified track is given the correct particle identification hypothesis, the $J/\psi\ h^+h^-$ mass shows a Λ_b^0 or B^0 mass peak, depending on whether one of the kaons is given a proton or a pion mass. If a pion is misidentified as a kaon, $B^0 \to J/\psi K^{*0}(\to K^+\pi^-)$ candidates, denoted as $B^0 \to J/\psi K^+\pi^-$, can be falsely assigned the $B_s^0 \to J/\psi K^+K^-$ hypothesis. After offline single and multi-variate cut selections, about 3.4% $B_s^0 \to J/\psi K^+K^-$ candidates come from a $\Lambda_b^0 \to J/\psi pK^-$ decay and about 0.5% from $B^0 \to J/\psi K^+\pi^-$ decay. Figure 4.4 shows the $J/\psi K^+K^-$ invariant mass distributions of simulated $\Lambda_b^0 \to J/\psi pK^-$, $B^0 \to J/\psi K^+\pi^-$

and $B_s^0 \to J/\psi K^+K^-$ candidates. The $B^0 \to J/\psi K^+\pi^-$ distribution drops down at 5200 MeV due to the cut in $m(J/\psi K^+\pi^-)$[2], which unfortunately is applied in the stripping-level selection of the $B^0 \to J/\psi K^+\pi^-$ sample. Both $\Lambda_b^0 \to J/\psi pK^-$ and $B^0 \to J/\psi K^+\pi^-$ can be found under the $B_s^0 \to J/\psi K^+K^-$ signal peak and thus, should be removed with care.

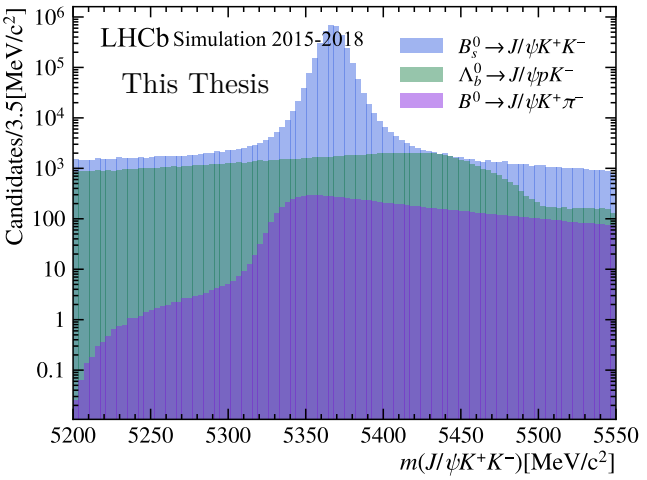


Figure 4.4: $J/\psi K^+K^-$ invariant mass from $\Lambda_b^0 \to J/\psi pK^-$, $B^0 \to J/\psi K^+\pi^-$ and $B_s^0 \to J/\psi K^+K^-$ simulated samples.

In both cases, vetoing the mass ranges around Λ_b^0 or B^0 peaks in the correct mass spectra and applying tighter particle identification requirements removes the majority of the misidentified $\Lambda_b^0 \to J/\psi pK^-$ and $B^0 \to J/\psi K^+\pi^-$ decays. A candidate is excluded if one of the kaon tracks has $\text{ProbNN}_\pi > 0.7$ or $\text{ProbNN}_K < 0.35$ and $|m(J/\psi K^+\pi^-) - m(B^0)| < 15\,\text{MeV}/c^2$ ($B^0 \to J/\psi K^+\pi^-$ *veto*); or if one of the kaon tracks has $\text{ProbNN}_p > 0.7$ or $\text{ProbNN}_K < 0.35$ and $|m(J/\psi pK^+) - m(\Lambda_b^0)| < 15\,\text{MeV}/c^2$ ($\Lambda_b^0 \to J/\psi pK^-$ *veto*). The nominal masses of B^0 and Λ_b^0 are taken from [70].

The vetoes remove about 74% of all Λ_b^0 background candidates and 76% of B^0 background candidates. The number of remaining background candidates in the tight signal region $m(J/\psi K^+K^-) \in (5347, 5387)$ MeV/c^2 is estimated by extrapolating the number of background candidates in the B_s^0 mass sidebands[3]. The background yields are estimated with the mass

[2]The cut is $m(J/\psi K^+\pi^-) \in (5150, 5450)$ MeV.

[3]Sidebands: $m(J/\psi K^+K^-) \in [5387, 5550)$ MeV/c^2 for $B^0 \to J/\psi K^+\pi^-$ and $m(J/\psi K^+K^-) \in [5200, 5347) \vee [5387, 5550)$ MeV/c^2 for $\Lambda_b^0 \to J/\psi pK^-$

fit of $m(J/\psi pK^+)$ for Λ_b^0 (or $m(J/\psi K^+\pi^-)$ for B^0) in two categories: one that passed the PID part of veto, but failed the mass criteria in the veto, $N_{PIDpass}$ and one that failed the PID part, $N_{PIDfail}$. The number of the remaining candidates is then,

$$N = \frac{N_{PIDpass}\varepsilon_f + N_{PIDfail}}{\varepsilon_{sig}}, \tag{4.15}$$

where ε_{sig} is the ratio of background candidates in the B_s^0 signal mass range and sidebands estimated with simulated background samples; ε_f is the ratio between $N_{PIDpass}$ and $N_{PIDfail}$. It is estimated with integrals of the $m(J/\psi pK^+)(m(J/\psi K^+\pi^-))$ shapes over the full mass range and the vetoed background mass window.

The number of estimated surviving Λ_b^0 candidates in the signal region is shown in Tab. 4.4. The remaining Λ_b^0 background is statistically subtracted

	$N_{\Lambda_b^0}$	N_{B^0}
2015	269.2 ± 38.0	8.7 ± 5.8
2016	1488.3 ± 89.0	174.7 ± 102.6
2017	1330.2 ± 79.1	250.2 ± 65.9
2018	1652.1 ± 84.8	119.6 ± 82.1

Table 4.4: The number of surviving Λ_b^0 and B^0 candidates in the B_s^0 signal mass region after the Λ_b^0 veto is applied. The statistical and systematics (from mass fit model) uncertainty added in quadrature are shown.

by injecting simulated Λ_b^0 events with negative weights, computed to match the resonant structure from the pentaquark analysis [75]. The absolute sum of weights equals the number of estimated remaining Λ_b^0 candidates.

The remaining $B^0 \to J/\psi K^+\pi^-$ candidates, see Tab. 4.4, make up only $0.100 \pm 0.004\%$ of the total selected signal candidates and are neglected in further analysis.

4.3 Statistical background subtraction with *sPlot*

As seen from Fig. 4.5, a predominately combinatorial background remains in the sample after the peaking background is vetoed. To subtract the remaining background from the sample, the *sPlot* method is used, [68]. It removes surviving peaking background and combinatorial background by giving a positive weight to the events that are more signal-like and negative

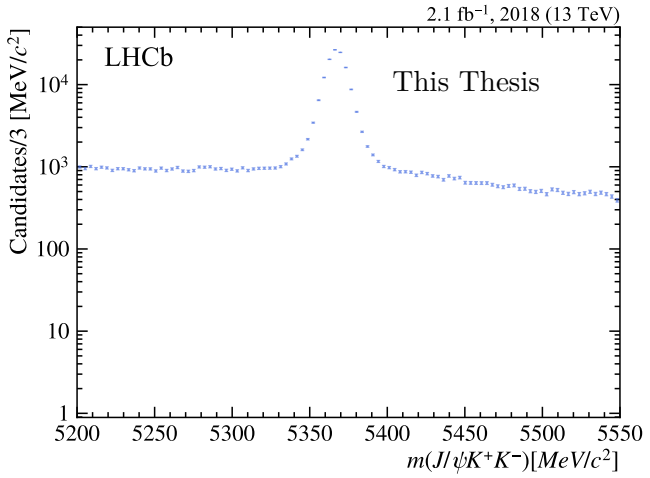


Figure 4.5: $J/\psi K^+K^-$ invariant mass, after offline selection and peaking background vetoes are applied.

to those that are more background-like. The *sPlot* method disentangles signal and background components in the observable of interest x using a discriminating variable y, assuming that signal and background distributions of y are known [68]. An extended likelihood fit is performed to the sample y to obtain the yields n_k for each component k and the corresponding yields covariance matrix V_{jk}. In this fit, the shapes of the signal and background components are fixed, and only the yields are varied. A sWeight $\mathcal{W}$ for component j is then assigned for a candidate i based on the value of the signal and background probability distribution functions at the discriminating variable y_i as,

$$\mathcal{W}_{j,i} = \frac{\sum_{k=1}^{n} V_{jk} f_k(y_i)}{\sum_{k=1}^{n} n_k f_k(y_i)} \tag{4.16}$$

where i is candidate's index; y_i is candidate's value of y; $f_k(y)$ is the PDF of k^{th} component; n is the total number of components. For each candidate, the sum of the sWeights for all components is 1,

$$\sum_{k=1}^{n} \mathcal{W}_{k,i} = 1. \tag{4.17}$$

The sWeights for component k are normalized such that,

$$\sum_{i=1}^{n} \mathcal{W}_{k,i} = n_k, \tag{4.18}$$

where n is the total number of candidates and n_k is the yield of k^{th} component. The covariance V_{ij} of the yields is,

$$V_{jk} = \sum_{i=1}^{n} \mathcal{W}_{j,i}\mathcal{W}_{k,i}. \tag{4.19}$$

The distribution of the observable of interest x with the applied sWeight $\mathcal{W}_k$ for k-th components reproduces, on average, the true distribution of k-th components in variable x. With the true distributions for the signal component "revealed", it is sufficient to only describe the signal distributions in the time and angular-dependent model as in Eq. 4.13. However, compared with the model that described the background, the signal-only model with applied sWeights will have a bigger variance in the parameters of interest. This is because the model with backgrounds exploits additional constraints from the background distributions.

On top of that, applying the sWeights comes at a price of underestimated statistical uncertainties in the time and angular-dependent fit. The covariance matrix V_{jk} in Eq. 4.16 is computed with the MLE fit with only yields floating. Therefore, the uncertainties on the parameters of the PDF $f_k(y)$ are not included. To overcome this problem, a systematic uncertainty is assigned: the parameters of $f_k(y)$ are varied within their uncertainties, and new sWeights are computed. The difference between using the baseline sWeights and the sWeights with varied PDF parameters is assigned as systematical uncertainty in Sec. 4.7.

The reconstructed invariant mass of $J/\psi K^+K^-$ is chosen as the discriminating variable because it provides a clear separation between the signal and the background. The mass model used to compute sWeights is described below. The yields and probability distribution parameters estimated from the fit are used as inputs to Eq. 4.16.

Baseline mass model

Without background and mass resolution effects and neglecting the natural width of B_s^{04}, $m(\mu^+\mu^-K^+K^-)$ would be a simple delta-peak at the invariant mass of B_s^0. The simplified measured mass uncertainty, see Eq. 3.5, depends on the momentum and angular uncertainties. The measured mass uncertainty is the sum of many non-Gaussian contributions: multiple scattering, single hit error, etc. Following the Central limit theorem, their sum is Gaussian. Therefore, the measured invariant mass $m(\mu^+\mu^-K^+K^-)$ distribution has

[4] about 0.0004 eV/c^2

a Gaussian-like core with a mean μ at the invariant mass of B_s^0 and width σ, which is defined by the invariant mass resolution of the LHCb detector. To improve the mass resolution on $\mu^+\mu^- K^+K^-$ combinations, $\mu^+\mu^-$ mass is constrained to the nominal J/ψ mass. Even though the core of the mass distribution is Gaussian, the tails are defined by non-Gaussian contributions: resolution, radiative energy loss, and multiple scattering. The non-Gaussian resolution and multiple scattering contribute to both the left and right side tails. The left-side tail has an additional contribution from the radiative energy loss of the final state particles. Therefore, a double-sided Crystal Ball distribution DCB is chosen to model $B_s^0 \to J/\psi K^+K^-$ invariant mass distribution,

$$f(m;\mu,\sigma,\alpha_L,n_L,\alpha_R,n_R) = N \cdot \begin{cases} A_L \cdot (B_L - \frac{m-\mu}{\sigma})^{-n_L}, & \text{for } \frac{m-\mu}{\sigma} < -\alpha_L \\ A_R \cdot (B_R + \frac{m-\mu}{\sigma})^{-n_R}, & \text{for } \frac{m-\mu}{\sigma_R} > \alpha_R \\ \exp\left(-\frac{1}{2} \cdot \left[\frac{m-\mu}{\sigma}\right]^2\right), & \text{otherwise,} \end{cases} \tag{4.20}$$

where

$$\begin{aligned} A_i &= \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \cdot \exp\left(-\frac{|\alpha_i|^2}{2}\right) \\ B_i &= \frac{n_i}{|\alpha_i|} - |\alpha_i| \\ N &= \frac{1}{\sigma(\sum_i (C_i + D_i))} \\ C_i &= \frac{n_i}{|\alpha_i|}\frac{1}{n_i - 1}\exp\left(-\frac{|\alpha_i|^2}{2}\right) \\ D_i &= \sqrt{\frac{\pi}{2}}\left(1 - \mathrm{erf}\left(\frac{|\alpha_i|}{\sqrt{2}}\right)\right) \end{aligned} \tag{4.21}$$

where the left side exponential tail is described by α_L and n_L; and the right side exponential tail is described by α_R and n_R; and $i \in (L, R)$. Together α_L, n_L, α_R and n_R are refered to as *"tail"* parameters in the text. The right tail in the DCB shape is the result of the multiple scattering. Multiple scattering is a process of particles interacting with the material they pass through, which slightly modifies their momenta and leads to larger uncertainty on tracks.

The width of the DCB σ is set to be a function of the measured per-event mass uncertainty σ_m. Figure 4.6 shows distributions of σ_m for the $B_s^0 \to J/\psi K^+K^-$ samples of each of the four data taking years.

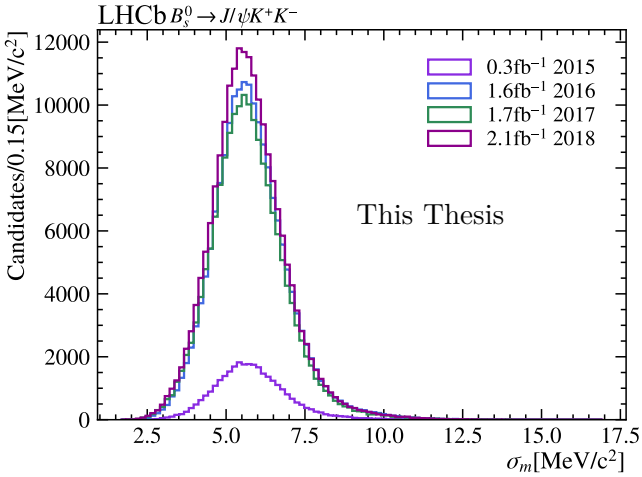


Figure 4.6: Distribution of per event mass uncertainty σ_m in collected samples.

The $\Lambda_b^0 \to J/\psi pK^-$ background is added with negative weights to the $B_s^0 \to J/\psi K^+K^-$ sample as described in Sec. 4.2.4. The remaining two components of background decays are $B^0 \to J/\psi K^+K^-$ and a combinatorial background. The same model describes the invariant mass distribution of $B^0 \to J/\psi K^+K^-$ as $B_s^0 \to J/\psi K^+K^-$ 4.20, once it is shifted by the difference $\Delta\mu_{B^0}$ between the B_s^0 and B^0 nominal masses. The invariant mass distribution of combinatorial background is modeled by an exponential function [76].

The probability of observing n events is given by,

$$p(n, m, \sigma_m) = \frac{(\sum_k \mu_k)^n e^{-\sum_k \mu_k}}{n!} \prod_{i=1}^{n} \frac{\sum_k \mu_k f_k(m_i, \sigma_{m,i})}{\sum_k \mu_k}, \tag{4.22}$$

where n is observed number of candidates; k is the component index and is defined as $k \in (\text{sig}, \text{comb}, B^0)$ for the signal, combinatorial background and $B^0 \to J/\psi K^+K^-$ background, respectively; μ_k is the expected number of candidates of k^{th} component; $f_k(m_i, \sigma_{m,i})$ is the PDF of the k^{th} component; i is the candidates index. The components of the PDF are,

$$f_{sig}(m, \sigma_m) = f(m; \mu, s_1, s_2, \alpha_L, n_L, \alpha_R, n_R|\sigma_m)p(\sigma_m) \tag{4.23}$$
$$f_{comb}(m, \sigma_m) = \gamma e^{-\gamma m} p(\sigma_m) \tag{4.24}$$
$$f_{B^0}(m, \sigma_m) = f(m; \mu - \Delta\mu_{B^0}, s_1, s_2, \alpha_L, n_L, \alpha_R, n_R|\sigma_m)p(\sigma_m) \tag{4.25}$$

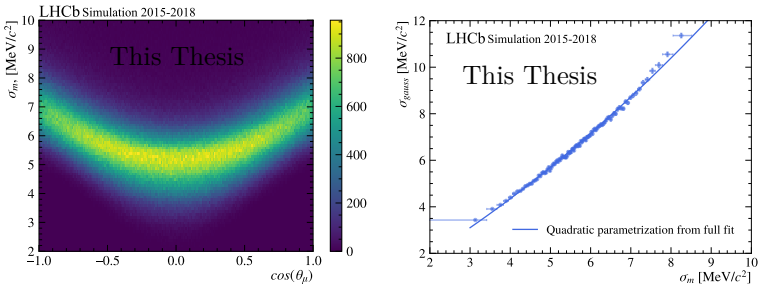
where f is defined in 4.20; γ is a combinatorial background exponent; $p(\sigma_m)$ is the PDF for the per-event estimated uncertainty on the invariant mass of the candidate. The $p(\sigma_m)$ is assumed to be the same for signal and background. Since $p(\sigma_m)$ is a multiplicative term in the total likelihood, it does not affect the likelihood minimization and, therefore, can be ignored.

Mass shape dependence on $\cos\theta_\mu$ described by σ_m as proxy

The invariant mass resolution of reconstructed resonances in LHCb depends on the daughter tracks momenta and opening angles. For the decay $B_s^0 \to J/\psi K^+K^-$, this leads to a correlation between the shape of the invariant mass distribution and the decay angle $\cos\theta_\mu$. This correlation was observed in [48], [30] and was found to introduce a significant bias on the parameters of interests. The per-event estimated uncertainty for the mass σ_m, see Eq. 3.5, can be used as an approximation for the invariant mass resolution. In Figure 4.7(a) the two-dimensional distribution of σ_m and $\cos\theta_\mu$ is shown. If the underlying dependence is ignored, background subtraction with *sPlot* does not yield the correct $B_s^0 \to J/\psi K^+K^-$ signal distribution of the observables of interest, which results in a non-negligible bias in the parameters of interests.

To include these correlation effects, the width of the double Crystal Ball is parametrized in terms of σ_m. The parameterization model is taken from the $B_s^0 \to J/\psi K^+K^-$ simulation, but the parameters are left free. Figure 4.7(b) shows the results of the simulation fit with the double Crystal Ball function, which width is parameterized by a second-order polynomial in the form $s_1 \cdot \sigma_m + s_2 \cdot \sigma_m^2$ (a baseline). The projected data points are computed by fitting the simulated sample in bins of σ_t with a double Crystal Ball. In each σ_m bin, the width of Crystal Ball is a single parameter, which is shown on the y-axis. The tail parameters $\alpha_{L,R}$ and $n_{L,R}$ of the double Crystal Ball in the binned fit are fixed to the result of the full fit. For the higher σ_m bins, the quadratic parameterization starts to deviate from the width of the double Crystal Ball obtained in bins of σ_m. Therefore for a subset of candidates with large σ_m the second-order polynomial parametrization is sub-optimal. Therefore, alternative parametrizations and their influence on the parameters of interest are explored in Sec. 4.6.

Figure 4.8 shows distributions of parameters that describe DCB width for two mass models fitted to the $B_s^0 \to J/\psi K^+K^-$ sample collected in 2018. The first model is defined in Eq. 4.22-4.23 with the baseline second-order polynomial DCB width parameterization defined by the coefficients s_1 and s_2. The second model does not depend on σ_m and directly fits the σ of DCB.



(a) (b)

Figure 4.7: Candidate's mass uncertainty σ_m per muons helicity angle $\cos\theta_\mu$ (a); The baseline double Crystal Ball width parametrization model projected over the Crystal Ball width in bins of σ_t (b)

The fit was performed in bins of K^+K^- mass and three bins of $\cos\theta_\mu$. As can be seen, in Fig. 4.8 the *DCB* width σ varies depending on the $\cos\theta_\mu$ bins. However, the polynomial coefficients s_1 and s_2 only fluctuate within uncertainties, remaining almost constant in the three $\cos\theta_\mu$ bins. Therefore, the second-order polynomial correctly describes the mass shape correlation with $\cos\theta_\mu$ within uncertainties.

The parameters s_1 and s_2 are referred to as *"width"* parameters in the text.

Fitting procedure and results

The tails of the mass distribution differ per the K^+K^- mass bin, which motivated a split of the sample into K^+K^- mass bins. Due to the changes in the momentum trigger selection criteria and detector response with time, the mass shape can differ slightly from year to year. Therefore, the sample is split into four data-taking years: 2015, 2016, 2017, and 2018. *Biased* and *unbiased* trigger categories have different muon tracks transverse momentum requirements and are treated separately in the mass fit. The mass fit is done in 48 bins: six K^+K^- mass bins, *biased* and *unbiased* trigger categories, and four years 2015-2018.

The signal *DCB* shape is taken from the signal simulation for each bin. Ideally, the simulation should contain only the real signal candidates but also the combinatorial background candidates. The combinatorial background candidates are built from a mix of tracks generated from signal and non-signal events. Such candidates are explicitly excluded from the simulation

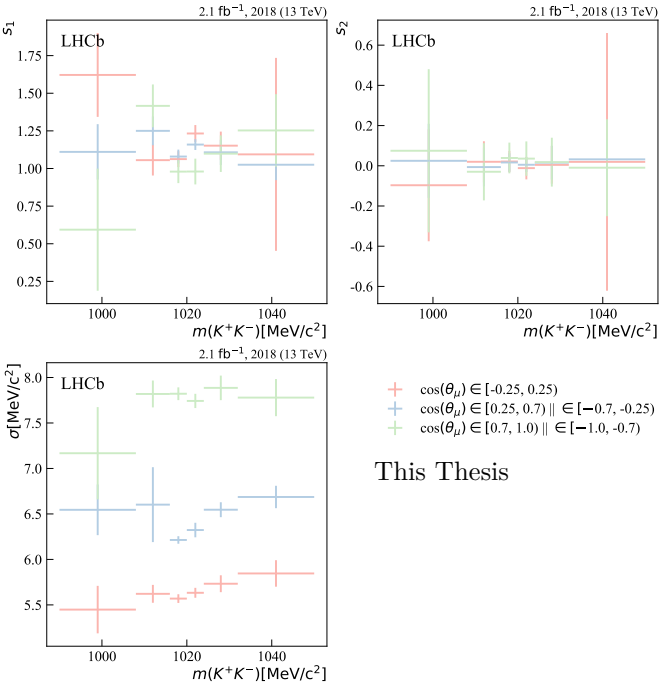


Figure 4.8: Dependence of the parameters that describe the width of Double Crystal Ball function on the K^+K^- mass and $\cos\theta_\mu$. Top row: plots show s_1, s_2, which are coefficients of the second-order polynomial in 4.22. They describe the dependence of the DCB width on the per event mass uncertainty σ_m. Bottom row: plot shows the fitted σ parameters of the DCB if the per-event estimated mass uncertainty is not used.

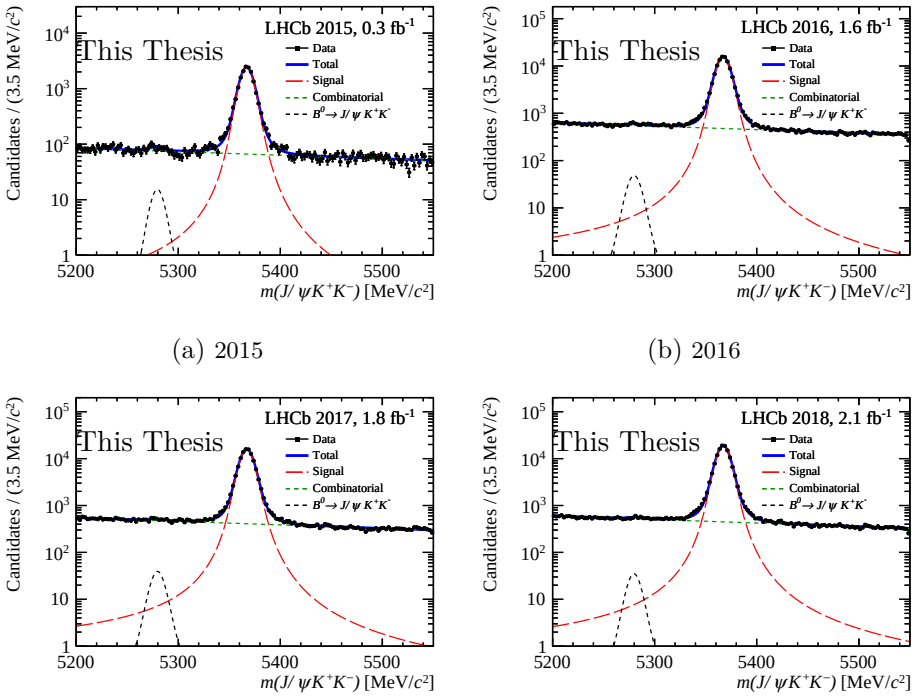
sample when determining the signal shape.

The mass resolution description in the simulation differs from the real detector mass resolution, so width parameters are refitted in the real data sample. The real data sample also refits the mean μ of DCB. The tail parameters are taken directly from the signal simulation mass shape. Figure 4.9 shows the total mass fit distributions for four years. The total number of signal and background events is shown in Tab. 4.5.

The sWeights for all three components are shown in Fig. 4.10 for 2018. The spread of the sWeights is a consequence of dependence on the measured per-event mass uncertainty; see Fig. 4.12.

year	signal	background
2015	16208 ± 134	9382 ± 114
2016	103234 ± 342	67156 ± 305
2017	105177 ± 355	57662 ± 288
2018	123703 ± 414	63030 ± 394

Table 4.5: The total number of signal and background events per year from the background subtraction procedure. Background combines both combinatorial and $B^0 \to J/\psi K^+K^-$ background.



(a) 2015 (b) 2016

(c) 2017 (d) 2018

Figure 4.9: Fits to invariant mass of $J/\psi K^+K^-$ for 2015-2018 years. The signal double-sided Crystal Ball is shown in red, B^0 peak is shown in black, the combinatorial background exponent is in green, and the total shape is in blue.

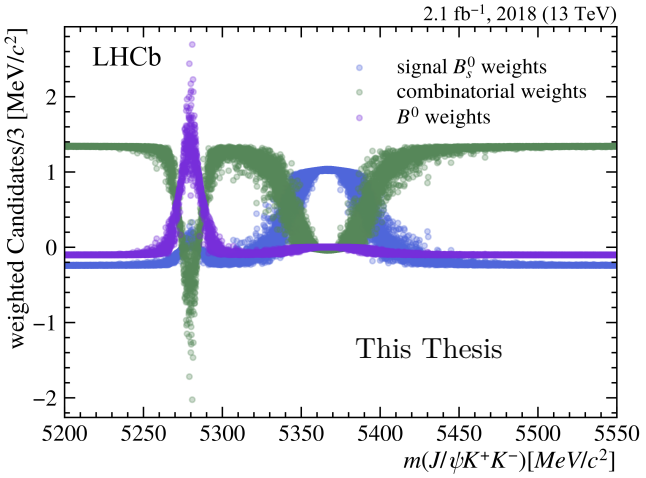


Figure 4.10: sWeights for $B_s^0 \to J/\psi K^+K^-$ (blue), combinatorial background (green) and $B^0 \to J/\psi K^+K^-$ (violet). Shown for the *unbiased* trigger category and $m(K^+K^-) \in [1016, 1020)\,\mathrm{MeV}/c^2$.

4.3.1 Baseline mass model limitations

The baseline mass model, described by Eq. 4.22, assumes the shape of the per-event mass uncertainty distribution is the same for the background and the signal. Fig. 4.11 depicts σ_m distribution in signal region ($m(J/\psi K^+K^-) \in [5300, 5440)\,\mathrm{MeV}/c^2$) and background sidebands ($m(J/\psi K^+K^-) \in [5200, 5300) \vee [5440, 5550)\,\mathrm{MeV}/c^2$).

The mass resolution depends on momentum and, therefore, σ_m depends on the measured $m(J/\psi K^+K^-)$. In Fig. 4.12, the average σ_m in $B_s^0 \to J/\psi K^+K^-$ simulation sample is plotted in bins of reconstructed $m(J/\psi K^+K^-)$. In the sidebands, the mass uncertainty grows with mass value. However, in the signal region, the average mass uncertainty is smaller.

In conclusion, the signal and background σ_m distributions differ. Thus, the assumption on which Eq. 4.22 is based is not fully satisfied.

To check the effect of this assumption, the mass fit is performed in bins of σ_m. In each bin, the width of signal DCB is not conditional but just one parameter σ. This parameter describes the width of both signal and background mass uncertainty distributions.

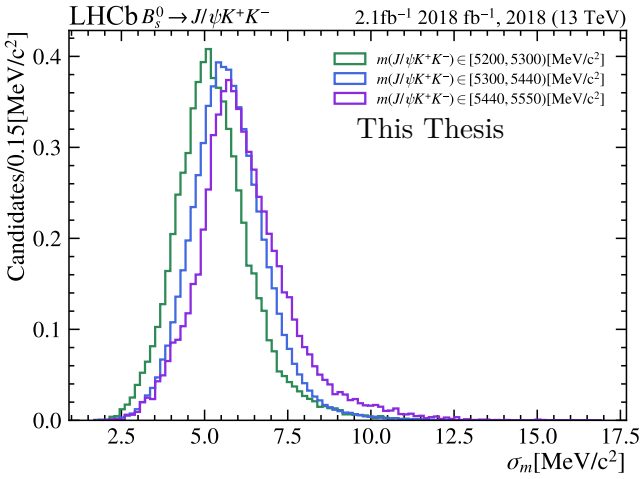


Figure 4.11: Normalised distributions of per event mass uncertainty σ_m in three different mass regions: signal - $(m(J/\psi K^+K^-) \in [5300, 5440)\,\mathrm{MeV}/c^2)$, lower sideband - $(m(J/\psi K^+K^-) \in [5200, 5300)\,\mathrm{MeV}/c^2)$ and higher sideband - $(m(J/\psi K^+K^-) \in [5440, 5550)\,\mathrm{MeV}/c^2)$. Distributions are shown for the sample collected in 2018.

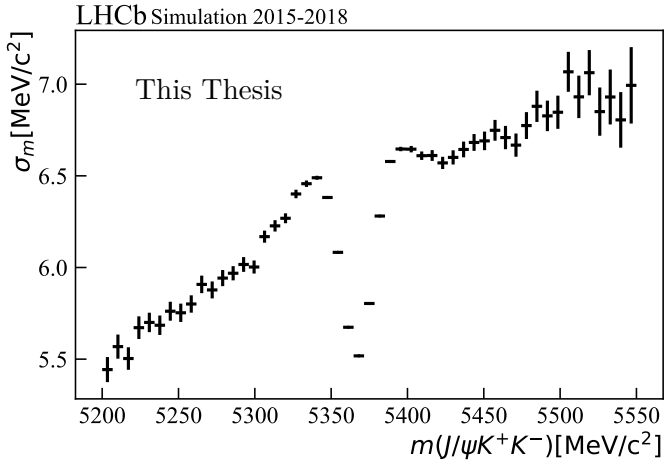


Figure 4.12: Average estimated mass uncertainty in bins of reconstructed $m(J/\psi K^+K^-)$ in $B_s^0 \to J/\psi K^+K^-$ simulation for 2015-2018.

Mass fit in bins of σ_m

For the mass fit in bins of σ_m, the fit is performed simultaneously in six bins of K^+K^- mass and eight bins of σ_m. The bins of σ_m are chosen such that the sample of simulated $B_s^0 \to J/\psi K^+K^-$ candidates populates each bin equally. The resulting bin boundaries are found to be $[0, 4.84, 5.41, 5.90, 6.52, 24]\,\mathrm{MeV}/c^2$.

The mass model in σ_m and K^+K^- mass bins is taken Eq. 4.22-4.23, but the width of signal DCB is defined as one parameter σ and not as a polynomial of σ_m. The mass model in each bin j is,

$$p^j(N, m) = \frac{(\sum_k \mu_k)^{n^j} e^{-\sum_k \mu_k}}{n^j!} \prod_{i=1}^{N^j} \frac{\sum_k \mu_k f_k(m_i)}{\sum_k \mu_k}, \tag{4.26}$$

$$f^j_{sig}(m) = f(m; \mu^j, \sigma^j, \alpha^j_L, n^j_L, \alpha^j_R, n^j_R) \tag{4.27}$$

$$f^j_{comb}(m) = \gamma^j e^{-\gamma^j m} \tag{4.28}$$

$$f^j_{B^0}(m) = f(m; \mu^j - \Delta\mu_{B^0}, \sigma^j, \alpha^j_L, n^j_L, \alpha^j_R, n^j_R), \tag{4.29}$$

where n^j is number of candidates in bin j; i is candidate index; k is component index and $k \in (\mathrm{sig}, \mathrm{comb}, B^0)$; f is defined in Eq. 4.20.

The fit is done in two steps:

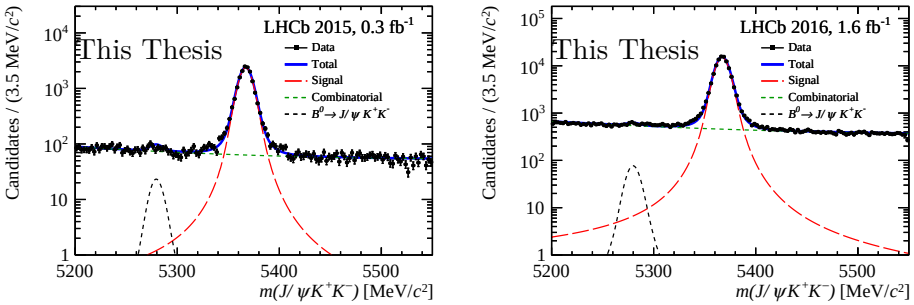
1. The signal shape is fitted to the $B_s^0 \to J/\psi K^+K^-$ simulation sample in bins of σ_m and $m(K^+K^-)$.

2. The model from Eq. 4.26 is fitted to the $B_s^0 \to J/\psi K^+K^-$ data sample. The tail parameters of the DCB shape in Eq. 4.26 are taken from the simulation results.

In Fig. 4.13, the combined mass distributions are shown for the four data-taking years.

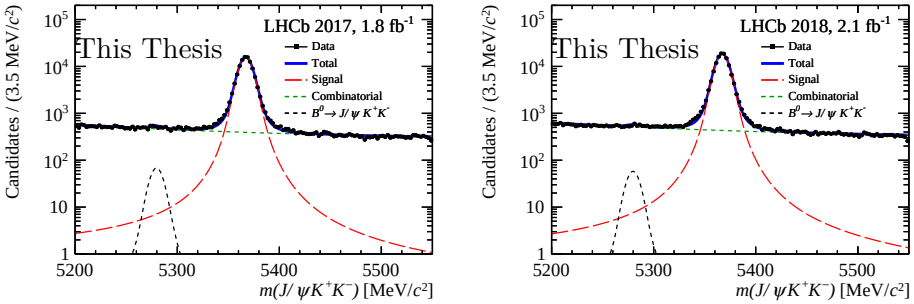
The total number of signal events in 2015, 2016, 2017, and 2018 is shown in Tab. 4.6. The statistics of the sample effectively change, as can be seen when comparing Tab. 4.6 and Tab. 4.5. The correlations between the two sets of the sWeights is computed with Eq. 4.19 and is shown in Tab. 4.6.

It is expected that the estimated parameter of interest will change if the sWeights change. The effect on the parameters of interest, defined in Sec. 4.1, of using a mass fit in bins of σ_m, rather than conditional, is within the systematic uncertainties of the baseline method, Sec. 4.7. The most significant difference is in $\Gamma_s - \Gamma_d$ and $|A_0|$, about the size of statistical uncertainty. For $\Gamma_s - \Gamma_d$, the systematic uncertainty is the size of statistical



(a) 2015

(b) 2016



(c) 2017

(d) 2018

Figure 4.13: The invariant mass fits of $J/\psi K^+K^-$ for 2015-2018 years, that was performed in 8 bins of σ_m. The model is described in Eq. 4.26.

year	signal	background	ρ_{sig}
2015	16270 ± 135	9320 ± 113	0.29
2016	103651 ± 371	66736 ± 335	0.38
2017	105579 ± 344	57257 ± 279	0.37
2018	124052 ± 573	62679 ± 540	0.37

Table 4.6: The total number of signal and background events per year from the background subtraction procedure in bins of measured mass uncertainty. Background combines both combinatorial and $B^0 \to J/\psi K^+K^-$ background. The last column shows the correlation ρ_{sig} (Eq. 4.19) between signal sWeights from the background subtraction procedure in bins of measured mass uncertainty and baseline signal sWeights.

uncertainty, and for $|A_0|$, the systematic uncertainty is twice as large as the statistical one. This fit was repeated in five and ten bins of σ_m, and the above conclusion does not change.

This analysis was done with other collaborators, and it was decided to keep the baseline mass fit in "Improved measurement of CP violation parameters in $B_s^0 \to J/\psi K^+K^-$ decays in the vicinity of the $\phi(1020)$ resonance", [63], as the difference is covered by the systematic. For future reference, this thesis's author believes this approach is more justified than the current baseline mass fit. In the following chapters, the baseline mass fit is assumed by default.

4.3.2 Better than a fit in σ_m bins?

A logical continuation of the mass fit in bins of σ_m is a full two-dimensional mass and mass-uncertainty fit. This fit would provide a smooth solution compared to the two-dimensional fit. However, the complication of such fit is understanding the shape for combinatorial background σ_m distribution. In Appendix A.1, a study of the two-dimensional fit is presented. Due to the unresolved issues with the combinatorial background σ_m distribution, it is only presented as a "first step" study to what can become a full description in the further iterations of this analysis.

4.3.3 *sPlot* procedure limitations

The *sPlot* procedure assumes that discriminating variables and the observables of interest are uncorrelated; see Sec. 4.3. However, a small correlation of mass and mass uncertainty with the final observables exists, as can be seen from Fig. 4.14 and Fig. 4.15. Figure 4.14 shows the correlation matrix, which assumes linear dependence between variables. Figure 4.15 shows the coefficient of determination (or r^2 score) for the quadratic model. The coefficient of determination measures the predictability of variable y from variable x,

$$r^2_{yx} = 1 - \frac{\sum\limits_i (y_i - f(x_i))^2}{\sum\limits_i (y_i - \langle y \rangle)^2}, \tag{4.30}$$

where i is the index of the data entry; f is the predicted dependence; $\langle y \rangle$ is average value of y. It ranges from zero to one, where zero means no predictability power, and one means the variable is perfectly predictable.

Properly treating these correlations between mass and final observables requires computing sWeights in bins of the final observables. However, in

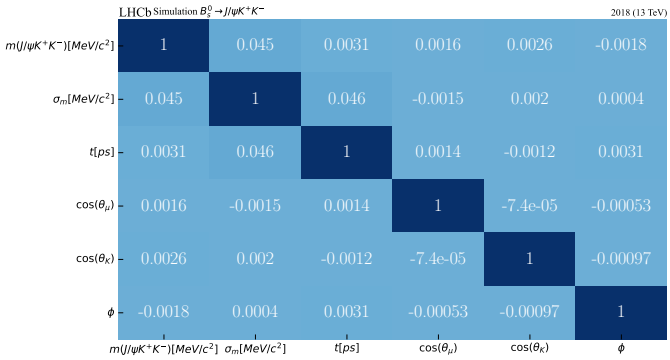


Figure 4.14: Correlation matrix between $m(J/\psi\, K^+K^-)$,σ_m and final observables t, $\cos\theta_\mu$, $\cos\theta_K$ and ϕ. Computed for the simulated $B_s^0 \to J/\psi\, K^+K^-$ sample. The smallness of correlations between σ_m and $\cos(\theta_\mu)$ is due to the quadratic dependence between them, which can be seen from Fig. 4.12 and Fig. 4.15.

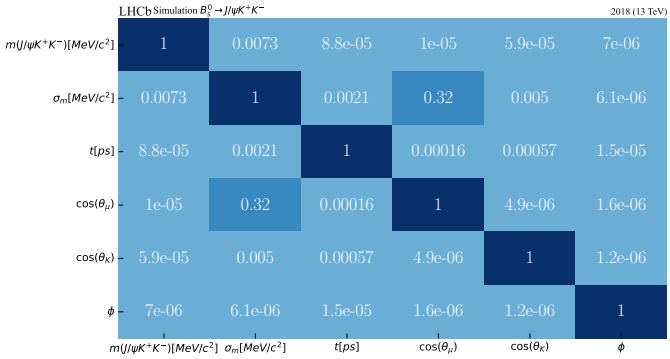


Figure 4.15: The coefficient of determination matrix coefficient of determination (r^2 score) for the quadratic model. Computed for the simulated $B_s^0 \to J/\psi\, K^+K^-$ sample.

comparison to the baseline method, this approach would increase the final statistical uncertainty.

Another way is including the mass distribution as an additional variable to the time-dependent angular maximum likelihood fit. This requires having a defined multidimensional model in mass, time, and angles, which is non-trivial but possible.

A novel approach is a more generalized *sPlot* procedure, called Custom Orthogonal Weights, proposed by Dembinski et al. [77]. According to this paper, this technique could correctly treat the correlations mentioned above

within the *sPlot* procedure. This statement was not checked as a part of this thesis but should be examined in future measurements.

For the baseline mass fit, the effect of ignoring correlations is estimated as a part of systematics studies, described in Sec. 4.7.

4.4 Flavour tagging

The $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ decay rates, where B_s^0 and $\overline{B}_s^0$ are the initial flavours of the B-meson, differ by a sign in front of $\cos(\Delta m_s t)$ and $\sin(\Delta m_s t)$. The amplitude of the $\sin(\Delta m_s t)$ term is proportional either to $\sin(\phi_s)$ or $\cos(\phi_s)$, depending on the polarization. Thus, if the initial flavour of B-meson is unknown, these terms cancel out, and the sensitivity to ϕ_s drops drastically[5].

In $p-p$ collisions, b-quarks are mostly produced as b-$\bar{b}$ pairs. The flavour of a signal quark b_{sig} from the $b-\bar{b}$ pair can be deduced from the flavour of the second b-quark in the pair, the tag b_{tag} [31]. Flavour tagging algorithms that use the other b-quark from the $b-\bar{b}$ pair to tag the signal b-quark are collectively known as *opposite-side* (`OS`) taggers [32], [33]. The flavour of b_{tag} is determined from secondary vertex charge and the charge of a lepton (μ, e) from a semileptonic B-decay or a kaon from the $b \to c \to s$ transition. This gives a name to the tagging algorithms used in this analysis: `OS vertex charge`, `OSCharm`, `OSMuon`, `OSElectron`, `OSKaon` [32], [33]. Their output is a continuous variable between -1 and 1, known as "signed dilution" qD,

$$qD = q(1-2\eta) \ \in [-1,1], \tag{4.31}$$

where η is the probability of the tagging algorithm to be wrong, known as *"mistag"* probability. The mistag probability has a range $[0, 0.5]$, where untagged candidates will have $\eta = 0.5$. The signed dilution is later separated into two variables: a tag q and the mistag probability. The tag q can take three values, $+1$ for B_s^0, -1 for $\overline{B}_s^0$ or 0 if the tagging algorithm did not tag the candidate. In Fig. 4.16, the `OS` taggers are shown in the bottom half of the schematics. The `OS vertex charge`, `OSCharm`, `OSMuon`, `OSElectron`, `OSKaon` are combined into one `OS` tag, following combination procedure described in [32].

Another type of flavour tagging algorithm is known as *same-side* algorithms. The same-side algorithms exploit the particles, generated by the

[5]The mixing phase ϕ_s also enters in front of some of the cosh terms in the decay rate, see Tab. 2.1. Therefore, some sensitivity ϕ_s to remains even without knowing the initial flavour.

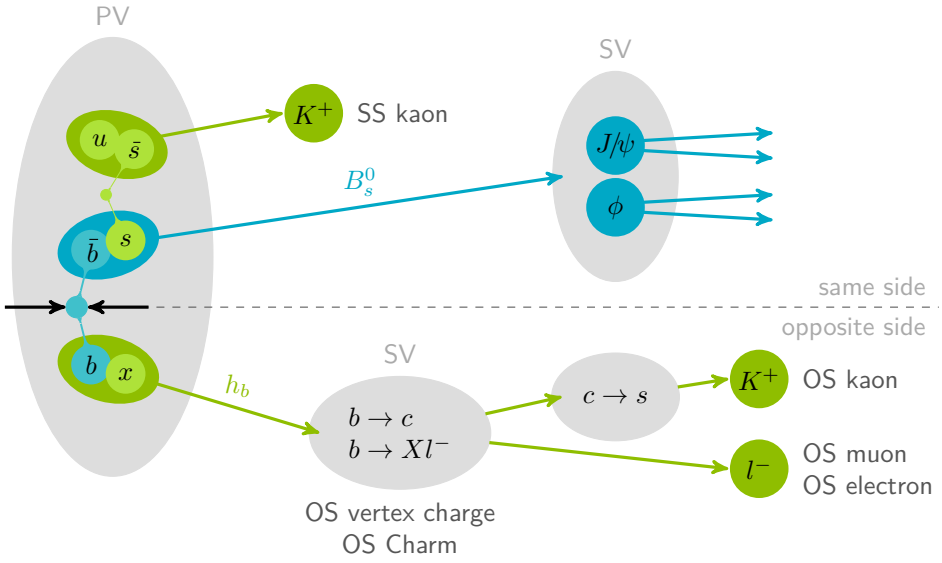


Figure 4.16: A scheme of flavour tagging algorithms from [78] used in the $B_s^0 \to J/\psi K^+K^-$ analysis. The signal b is shown in the top part. The *same-side* tagging algorithms are shown on the top too. The bottom half shows the *opposite-side* flavour tagging algorithms. The `tikz` version of this plot is taken from [79] and was modified to represent B_s^0 decays.

hadronization of the partner quark of the light quark in B_s^0, which was produced during with b_{sig} hadronization. For B_s^0, the light quark is an s quark, and its partner quark $\bar{s}$, has about 50% chance to hadronize as a charged kaon [80]. The charge of this kaon is then used to deduce the flavour of B_s^0. The corresponding tag is called `SSKaon` [80]. It is shown schematically in Fig. 4.16.

Flavour tagging algorithms are based on a multivariate classifier that takes the kinematics and topological properties of the candidates as inputs. The algorithms are trained on specially selected samples. The `OS` taggers are trained on the $B^+ \to J/\psi K^+$ data sample and `SS` taggers on simulated samples. If both tags are present for a candidate, their values are combined, considering the associated mistag probabilities.

4.4.1 Calibration of mistag probability ω

The efficiency of the flavour tagging algorithm depends on the signal decay and its selection. Therefore, the mistag probability is recalibrated to the true mistag probability ω to reflect the $B_s^0 \to J/\psi K^+K^-$ decays selected for this analysis. The calibration is done separately for SS tagger on $B_s^0 \to D_s^-\pi^+$ decays and for OS tagger on $B^+ \to J/\psi K^+$ decays. The true mistag probability ω for number of rightly tagged candidates R and wrongly tagged candidates W is,

$$\omega = \frac{W}{W+R}. \tag{4.32}$$

For both taggers the linear dependence between measured mistag η and true mistag ω is assumed,

$$\omega = \left(p_0 + \frac{\Delta p_0}{2}\right) + \left(p_1 + \frac{\Delta p_1}{2}\right)(\eta - \langle\eta\rangle) \text{ , for } B_s^0 \tag{4.33}$$

$$\bar{\omega} = \left(p_0 - \frac{\Delta p_0}{2}\right) + \left(p_1 - \frac{\Delta p_1}{2}\right)(\eta - \langle\eta\rangle) \text{ , for } \overline{B}_s^0 \tag{4.34}$$

where ω is true mistag for B_s^0 and $\bar{\omega}$ is true mistag for $\overline{B}_s^0$; $\langle\eta\rangle$ is an average measured η in the calibration sample; $p_{0,1}$ are calibration parameters and $\Delta p_{0,1}$ are mistag asymmetries. The mistag asymmetries $\Delta p_{0,1}$ account for a difference between ω for B_s^0 and $\overline{B}_s^0$.

The calibrated mistag probabilities then enter Eq. 4.13 instead of measured mistag probability η.

Calibration of SS tagger

The SS tagger has to be calibrated on a sample that contains information about B_s^0 hadronization. By construction, it thus has to be a B_s^0 decay. Therefore, the SS tagger calibrated on a $B_s^0 \to D_s^-\pi^+$ sample collected in Run 2. It is a flavour-specific decay, as the charges of π^+ and D_s^- identify the flavour of B_s^0. But, as B_s^0 oscillates to $\overline{B}_s^0$, only the flavour tag at the decay is known from the charge of the daughter particles. The decay time probability distribution for B_s^0, including decay time acceptance and resolution effects, is given by,

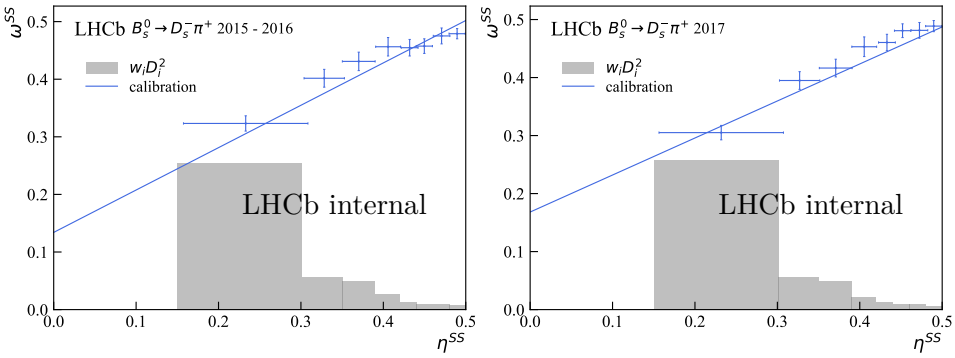
$$\mathcal{P}(t|\delta_t) = \varepsilon(t)\left[\Gamma(t') \otimes G(t - t'|\delta_t)\right] \text{ ,} \tag{4.35}$$

$$\Gamma(t) = e^{-\Gamma t}\left[\cosh \Delta\Gamma_s t/2. + q^{\text{mix}}(1 - 2\omega(\eta))\cos \Delta m_s t\right] \tag{4.36}$$

where $\varepsilon(t)$ is decay time acceptance; $G(t - t'|\delta_t)$ is Gaussian-like decay time resolution; t is the decay time of $B_s^0 \to D_s^-\pi^+$; δ_t is per event decay time

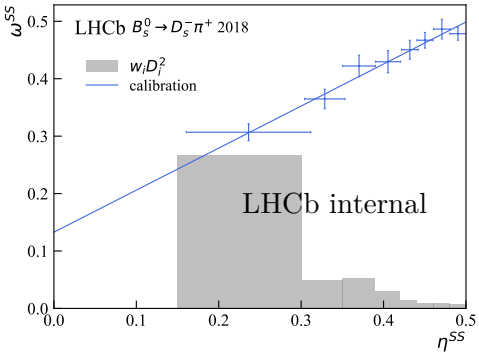
uncertainty; q^{mix} is 1 if the B_s^0 has not changed flavour from its initial flavour, and -1 if it did. The true mistag $\omega(\eta)$ can be estimated from the maximum likelihood fit of Eq. 4.35 to $B_s^0 \to D_s^- \pi^+$ decay time. The fit is described in detail in [67].

Figure 4.17 shows the SS combined tagger calibration for 2015-2016, 2017 and 2018 data taking years.



(a) SS 2015 - 2016

(b) SS 2017



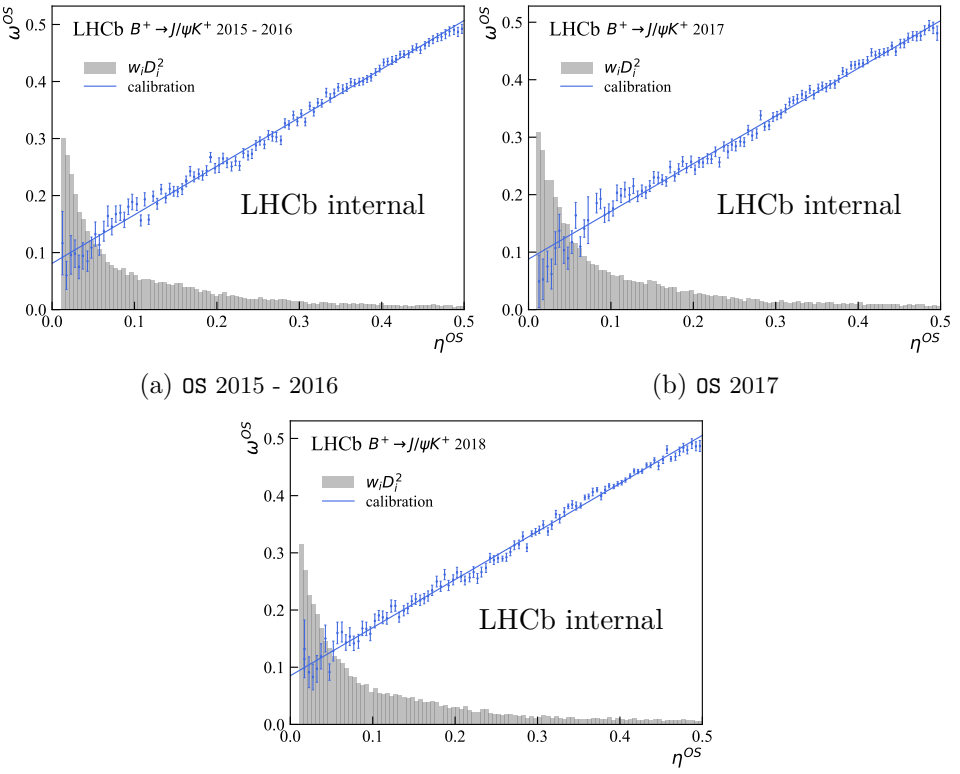
(c) SS 2018

Figure 4.17: SS combination tagger calibration for 2015-2016, 2017, 2018 data taking years. In grey, the effective statistical power $w_i D_i^2 = w_i(1 - 2\eta_i)$ is shown, where w is the sWeight and η is the mistag probability for $B_s^0 \to J/\psi K^+ K^-$ candidates. Calibrations taken from [67].

Calibration of OS combined tagger

The OS combined tagger is calibrated on a reweighted $B^+ \to J/\psi K^+$ sample collected in 2015-2018 data taking years. The $B^+ \to J/\psi K^+$ sample is

background subtracted with *sPlot* procedure and is kinematically reweighted to represent background-subtracted $B_s^0 \to J/\psi K^+K^-$ sample. The details are discussed in [63]. $B^+ \to J/\psi K^+$ is a self-tagged decay: the kaon charge defines the flavour of B^+. Therefore, the number of correctly and incorrectly tagged candidates can be counted directly. Figure 4.18 shows the OS combined tagger calibration for 2015-2016, 2017 and 2018 data taking years.



(a) OS 2015 - 2016

(b) OS 2017

(c) OS 2018

Figure 4.18: OS combination tagger calibration for 2015-2016, 2017, 2018 data taking years. In grey, the effective statistical power $w_i D_i^2 = w_i(1 - 2\eta_i)$ is shown, where w is the sWeight and η is the mistag probability for $B_s^0 \to J/\psi K^+K^-$ candidates. Calibrations taken from [67].

Calibration parameters for the OS and SS taggers are summarized in Tab. 4.7.

Parameter	2015 + 2016 OS	2015 + 2016 SS
$\langle\eta\rangle$	0.3546	0.4239
p_0	$0.3831 \pm 0.0007 \pm 0.0009$	$0.4459 \pm 0.0051 \pm 0.0026$
p_1	$0.8518 \pm 0.0062 \pm 0.0163$	$0.7357 \pm 0.0608 \pm 0.0235$
$\rho(p_0, p_1)$	0.120	-0.064
Δp_0	0.0092 ± 0.0013	-0.0164 ± 0.0009
Δp_1	0.0141 ± 0.0119	0.0277 ± 0.0089
$\rho(\Delta p_0, \Delta p_1)$	0.119	0.117
Parameter	2017 OS	2017 SS
$\langle\eta\rangle$	0.3546	0.4239
p_0	$0.3821 \pm 0.0007 \pm 0.0006$	$0.4387 \pm 0.0048 \pm 0.0016$
p_1	$0.8297 \pm 0.0065 \pm 0.0122$	$0.6382 \pm 0.0574 \pm 0.0414$
$\rho(p_0, p_1)$	0.139	0.049
Δp_0	0.0123 ± 0.0014	-0.0158 ± 0.0008
Δp_1	0.0510 ± 0.0135	0.0143 ± 0.0083
$\rho(\Delta p_0, \Delta p_1)$	0.137	0.128
Parameter	2018 OS	2018 SS
$\langle\eta\rangle$	0.3546	0.4239
p_0	$0.3832 \pm 0.0006 \pm 0.0009$	$0.4431 \pm 0.0050 \pm 0.0021$
p_1	$0.8406 \pm 0.0059 \pm 0.0079$	$0.7321 \pm 0.0594 \pm 0.0361$
$\rho(p_0, p_1)$	0.136	0.057
Δp_0	0.0114 ± 0.0013	-0.0158 ± 0.0007
Δp_1	0.0133 ± 0.0118	0.0229 ± 0.0077
$\rho(\Delta p_0, \Delta p_1)$	0.135	0.111

Table 4.7: OS and SS calibration parameters for $B_s^0 \to J/\psi K^+K^-$. The first uncertainty is statistical, and the second uncertainty is systematic. Taken from [67].

4.4.2 Performance of flavour tagging algorithms

The tagging performance can be quantified with the *tagging power* $\varepsilon_{tag} D_{tag}^2$. The factor $D_{tag} \equiv (1 - 2\omega)$ is the dilution factor from the imperfect flavour tagging, and ε_{tag} is the fraction of tagged signal candidates in the sample. The tagging power encapsulates the effective reduction of the statistical power of measurement due to the imperfect tagging. In Tab. 4.8, tagging dilution and tagging power are reported for $B_s^0 \to J/\psi K^+K^-$.

2015 - 2016	$\varepsilon_{tag}(\%)$	D^2_{tag}	$\varepsilon_{tag}D^2_{tag}(\%)$
exclusive OS	11.34	0.0776	0.88 ± 0.01
exclusive SS	42.58	0.0278	0.97 ± 0.14
OS $\cap$ SS	23.93	0.0974	2.33 ± 0.06
total	77.85	0.0495	4.18 ± 0.15
2017	$\varepsilon_{tag}(\%)$	D^2_{tag}	$\varepsilon_{tag}D^2_{tag}(\%)$
exclusive OS	11.11	0.0792	0.88 ± 0.01
exclusive SS	43.16	0.0218	0.94 ± 0.14
OS $\cap$ SS	24.45	0.0982	2.40 ± 0.07
total	78.67	0.0536	4.22 ± 0.16
2018	$\varepsilon_{tag}(\%)$	D^2_{tag}	$\varepsilon_{tag}D^2_{tag}(\%)$
exclusive OS	11.17	0.0814	0.91 ± 0.01
exclusive SS	42.96	0.0235	1.01 ± 0.14
OS $\cap$ SS	24.80	0.0987	2.45 ± 0.07
total	78.93	0.0552	4.36 ± 0.16

Table 4.8: $B^0_s \to J/\psi K^+K^-$ tagging performance in 2015-2016, 2017 and 2018 data taking years.Taken from [67].

4.5 Acceptance and resolution

The finite detector resolution, detector acceptance, and selection criteria affect the physical time- and angular-dependent decay rate. Therefore, to correctly describe the measured decay rate, those effects have to be included.

In Fig. 4.19, the decay time projection of the decay rate for both initial B^0_s and $\overline{B}^0_s$ is shown in the range $t \in (-0.1, 5.)$ ps with perfect decay time resolution, no decay time acceptance effects, and no angular acceptance effects. To illustrate the influence of decay time resolution, decay time acceptance, and angular acceptance, all three are modeled and applied to the distributions shown in Fig. 4.19.

The finite decay time resolution of the LHCb detector originates from the limited precision of the detector on the measured observables, like momentum and vertices positions. It smears the measured values of decay time. As a result, the values of ϕ_s phase and B^0_s-$\overline{B}^0_s$ oscillation frequency, Δm_s, are

affected [6]. To the first order, this is a dilution effect. It can be described by a Gaussian resolution function $R(t)$ with width σ. Its dilution D_{tr} is then [29],

$$D_{tr} = e^{-\frac{1}{2}\sigma^2 {\Delta m_s}^2}. \tag{4.37}$$

The decay time resolution is modeled as a Gaussian function centered at zero with 45 fs width. It is then convoluted with distribution on Fig. 4.19. The result is shown in Fig. 4.20. Because of the finite resolution, negative decay times appear, and the distribution gets damped.

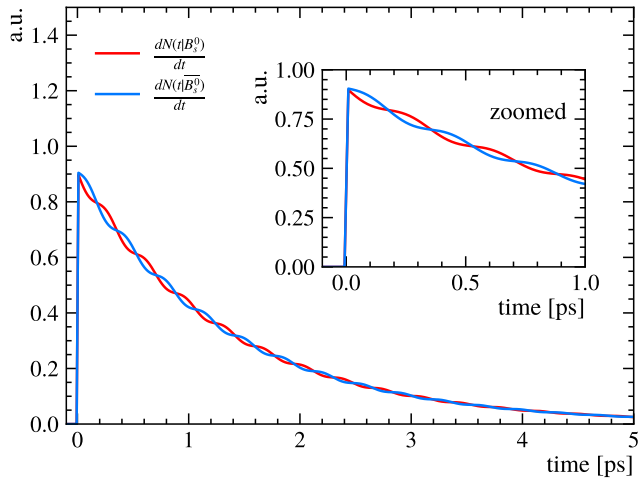


Figure 4.19: Decay time projection of decay rate PDF for initial B_s^0 flavour, $\frac{dN(t|B_s^0)}{dt}$ and $\overline{B}_s^0$, $\frac{dN(t|\overline{B}_s^0)}{dt}$ in the range $t \in (-0.1, 5.)$ ps. Parameter values used are described in Appendix A.2. No resolution or acceptance effects are applied. The normalization is done for $t \in (0.3, 15)$ ps as in the analysis. *Note that the steep drop near 0 is the consequence of plotting* `numpy` [82] *Heaviside step function.*

The angular resolution of the detector is about 10 mrad [83], which is small to the physical variations in the angular distributions. Therefore, the angular resolution is ignored in the analysis. The bias in the final parameters from ignoring the angular resolution is assigned as systematic uncertainty.

The trigger level selection requirements are split into two categories: *unbiased* and *biased*, as described in Sec. 4.2.1. The *biased* category has a displacement requirement that results in the measured decay time inefficiencies and, therefore, a non-trivial decay time acceptance shape. The *unbiased*

[6] In the dedicated measurements of Δm_s with $B_s^0 \to D_s^- \pi^+$ decays, the decay time resolution is one of the main sources of systematic uncertainties [81].

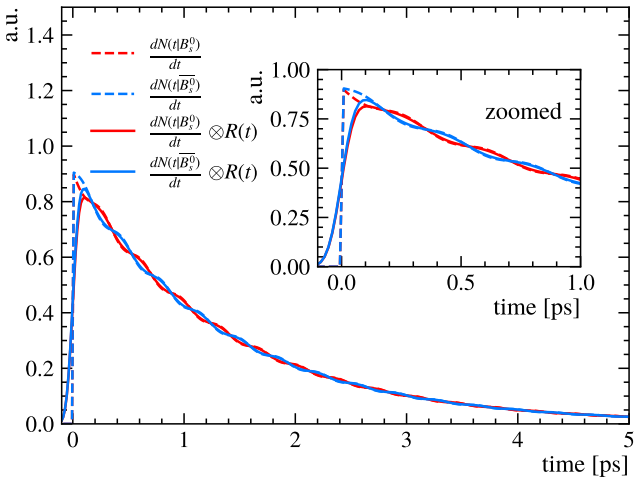


Figure 4.20: Decay time projection of decay rate PDF for initial B_s^0 flavour, $\frac{dN(t|B_s^0)}{dt}$ and $\overline{B}_s^0$, $\frac{dN(t|\overline{B}_s^0)}{dt}$ in the range $t \in (-0.1, 5.)$ ps. The dashed lines show the decay rate before the decay time resolution is applied. The solid lines show the decay rate after the decay time resolution is applied. The decay time resolution is shown in Fig. A.5.

category is almost free of any biases and has an almost uniform decay time acceptance. Here the decay time acceptance is modeled with a $\varepsilon(t) = 1 - e^{-\beta t}$ function with $\beta = 2$ ps^{-1}. It is then applied to the distributions in Fig. 4.19, which is shown in Fig. 4.21.

The geometrical acceptance of the LHCb detector, kinematic selection criteria, hit resolution, etc. lead to a non-uniform efficiency in the measured helicity angles. If uncovered, the angular acceptance biases the relative fractions of the polarised amplitudes. To illustrate this, the angular acceptance is modeled as a combination of Legendre polynomials $P_i(\cos\theta_K)$ and spherical harmonics $Y_{ml}(\theta_\mu, \phi)$ as

$$\varepsilon(\Omega) = \sum_{i,l,m} c_{lm}^i P_i(\cos\theta_K) Y_{ml}(\theta_\mu, \phi). \tag{4.38}$$

Appendix A.2.1 specifies the terms and parameters used for the illustration. In Fig. 4.22, the distortion from an example angular acceptance is shown with respect to the distribution with perfect angular acceptance.

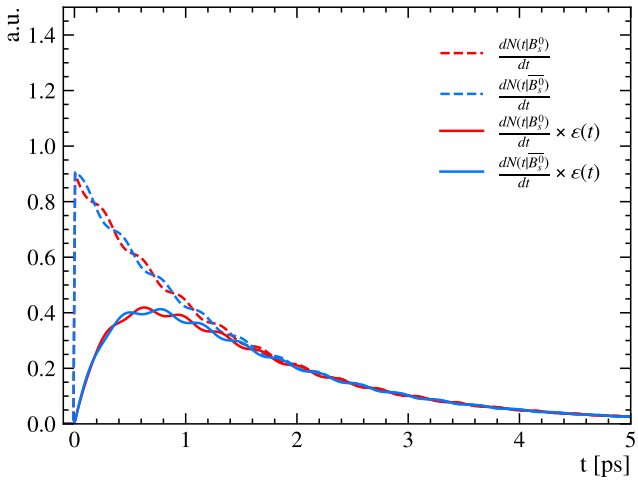


Figure 4.21: Decay time projection of decay rate PDF for initial B_s^0 flavour, $\frac{dN(t|B_s^0)}{dt}$ and $\overline{B}_s^0$, $\frac{dN(t|\overline{B}_s^0)}{dt}$ in the range $t \in (-0.2, 5.)$ ps. The dashed lines show the decay rate before the decay time acceptance is applied. The solid lines show the decay rate after the decay time acceptance is applied. The decay time acceptance is shown in Fig. A.6.

4.5.1 Decay time resolution

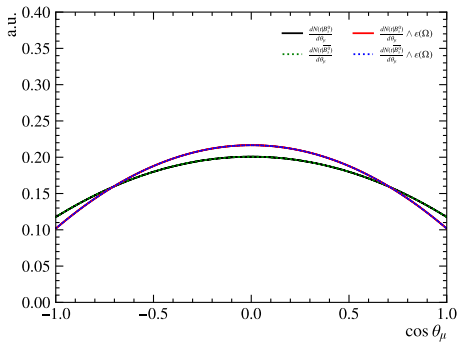
The decay time resolution of LHCb is small in comparison to $1/\Gamma_s$. Therefore, it mainly affects the oscillation terms of the physical PDF, and the influence on lifetimes is assumed to be negligible. For the small in comparison to $1/\Gamma_s$ the decay time resolution and under the assumption that $\Delta\Gamma_s = 0$ ps^{-1}, the observed B_s^0 PDF for one $\mathcal{CP}$-eigenstate can be approximated as,

$$\mathcal{P}^{\text{obs}}(t) \approx \Gamma_s e^{-\Gamma_s t} H(t) \left(1 + D_{tr}[-C\cos(\Delta m_s t) + S\sin(\Delta m_s t)]\right) \quad (4.39)$$

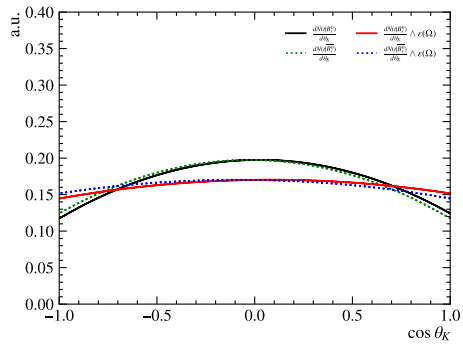
where $H(t)$ is the Heaviside function; D_{tr} is the damping factor, called *dilution*. The factor D_{tr} is a multiplicative factor in front of $S \approx \sin(\phi_s)$. Therefore, decoupling the decay time resolution in the amplitude is necessary to determine the value of ϕ_s.

In the full observed PDF, the decay time resolution is described by a convolution of the physical PDF with a Gaussian. The Gaussian width effectively describes the dilution of the sample and depends on the measured decay time uncertainty.

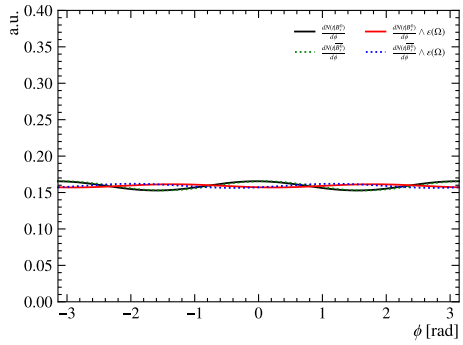
The decay time resolution procedure can be summarized in the following steps:



(a) $\cos\theta_\mu$



(b) $\cos\theta_K$



(c) ϕ

Figure 4.22: (a) The $\cos\theta_\mu$ projections of decay rate PDF for B_s^0 and $\overline{B}_s^0$. Shown in black-green with perfect angular acceptance effects and in blue-red with angular acceptance from Eq. 4.38. (b) The $\cos\theta_K$ projections of decay rate PDF for B_s^0 and $\overline{B}_s^0$. Shown in black-green with perfect angular acceptance effects and in blue-red with angular acceptance from Eq. 4.38. (c) The ϕ_s projections of decay rate PDF for B_s^0 and $\overline{B}_s^0$. Shown in black-green with perfect angular acceptance effects and in blue-red with angular acceptance from Eq. 4.38.

1. The true resolution is estimated in the bins of measured decay time uncertainty σ_t from the control sample.

2. Using the simulated samples, the true resolution is corrected to represent the signal sample.

3. The measured decay time uncertainty is calibrated with respect to the estimated true resolution.

4. The decay time resolution is described with a single Gaussian function in the final decay time angular PDF, where the Gaussian width depends on the calibrated measured decay time uncertainty.

The following sections discuss the details of these steps, starting with the selection of the control sample.

Selection for the decay time resolution control sample

The decay time resolution is estimated from the data. The measured decay time of *prompt* $J/\psi K^+K^-$ candidates from the inclusive $J/\psi K^+K^-$ sample is used as a proxy for decay time resolution shape. The prompt $J/\psi K^+K^-$ candidates are combinations of $J/\psi \to \mu^+\mu^-$ produced in the primary vertex with two random kaon tracks. The selection criteria for inclusive $J/\psi K^+K^-$ sample are identical to those applied to $B_s^0 \to J/\psi K^+K^-$ decays, as described in Tab. 4.3. The only exception is the absence of the displacement. This significantly increases the combinatoric background of the sample. As a result, selected inclusive $J/\psi K^+K^-$ candidates are dominated by prompt decays, which have zero true decay time ($t_{true} = 0\,\text{ps}$). However, about $5 - 10\%$ of the sample is contaminated by a long-lived background. Part of this background is real $B_s^0 \to J/\psi K^+K^-$ decays, and part is a random combination of tracks originating from the long(er)-lived particles with prompt ones. This long-lived background affects the extraction of the resolution function from the sample, as its distributions can be interpreted as resolution effects. Therefore, the long-lived background must be treated.

Origins of decay time resolution

In LHCb, the decay time is measured from the momentum p of the particle and decay length:

$$t = \frac{m\vec{L} \cdot \vec{p}}{|p|^2} \tag{4.40}$$

where the $\vec{L}$ is decay length defined in Eq. 3.3; $\vec{p}$ is momentum; m is mass of a particle.

The daughter tracks of the B_s^0 decays are known with finite precision. This then results in the finite precision on the vertex position and kinematics of a particle. Both contribute to the finite decay time resolution, where the vertex position precision is the most important effect, followed by the precision of the momentum of a particle. A priori, the true decay time resolution is not known, but the uncertainty on the measured decay time

can be estimated by propagating the estimated uncertainties on the distance between primary and decay vertices σ_L (decay length) and momentum σ_p,

$$\sigma_t^2 \approx \left(\frac{m}{p}\right)^2 \sigma_L^2 + \left(\frac{t}{p}\right)^2 \sigma_p^2 + corr. \tag{4.41}$$

The measured decay time uncertainty σ_t is not guaranteed to be the true decay time resolution. Therefore, the true decay time resolution has to be estimated separately. The measured decay time uncertainty is then calibrated to the true decay time resolution in the bins of σ_t.

The momentum resolution in LHCb is $\delta p/p \sim 0.005$ for particles with $p < 20$ GeV/c and 0.008 for particles with $p \sim 100$ GeV/c [25]. The contribution to the decay time uncertainty from the term proportional to σ_p is very small. Therefore, the decay time uncertainty is assumed to be time-independent and defined by the first term. This term is evaluated on the inclusive $J/\psi K^+K^-$ sample, which represents the time-independent, or "prompt" term. However, Fig. 4.23 shows that, in reality, a decay time dependence is present for a small number of $B_s^0 \to J/\psi K^+K^-$ candidates. This effect is not prominent in the simulated sample and is still linear within the statistical uncertainties. The real data sample has even bigger uncertainty, as its size is a magnitude smaller than the simulated signal sample. It is expected that for this analysis, the decay time dependence of the time resolution can be safely ignored.

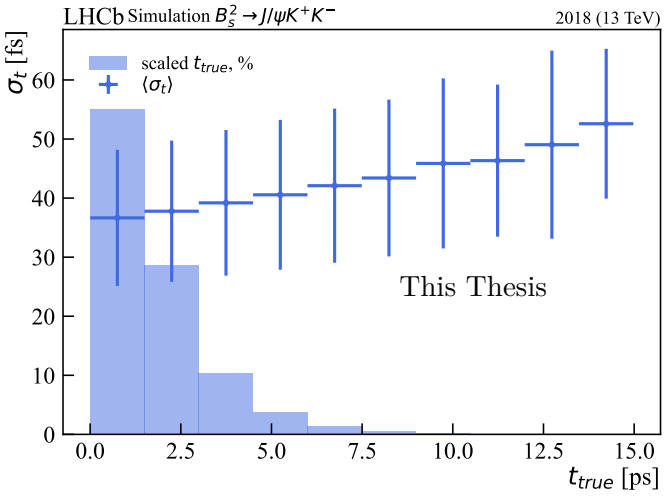


Figure 4.23: Decay time dependence of the measured per event decay time uncertainty in the 2018 simulated $B_s^0 \to J/\psi K^+K^-$ sample.

Decay time resolution estimation

The observed PDF $\mathcal{P}^{obs}(t)$ is a convolution of the physics PDF with the resolution function,

$$\mathcal{P}^{\text{obs}}(t) = \int_{-\infty}^{\infty} \mathrm{d}t' \; R(t;t') \; \mathcal{P}^{\text{phys}}(t') \; H(t') \tag{4.42}$$

where $R(t;t')$ is the resolution response function. The decay time resolution function $R(t;t')$ in the decay time angular fit is a Gaussian centered at zero with width dependent on the calibrated measured decay time uncertainty. If the decay time resolution is Gaussian, the resolution model can be extracted from a fit to the control sample. However, the non-Gaussian effects were observed in the decay time resolution control sample in the previous measurement [49]. The cause of these effects is the misassociation between the decay vertex and the primary vertex. Luckily, to measure ϕ_s, mainly the dilution factor the resolution generates is important; see Eq. 4.39. Therefore, the aim is to estimate the effective dilution from the control sample rather than the exact decay time resolution shape. The effective dilution is then used to compute the Gaussian width for the resolution model in the decay time angular fit; see Eq. 4.37. The resolution model in the decay time angular fit is then guaranteed to correctly represent the effective dilution.

The effective dilution D_{tr} is estimated using the control sample. Figure 4.24 shows the decay time distribution of the inclusive $J/\psi\, K^+K^-$ sample. The negative reconstructed decay time originates from the decay time resolution, whereas the positive reconstructed decay times are a mix of resolution effects and real decays. Under the assumptions that the decay time resolution function $R(t;t^{'})$ is time-invariant and symmetric around the mean, the left-tail of Fig. 4.24 can be used as a proxy for the decay time resolution. The distribution in Fig. 4.24 is centered roughly but not exactly at $t = 0$ ps. The tiny shift is known as *decay time bias* and is introduced in the next section. Here, the decay time distribution of the control sample is shifted to be centered at $t = 0$ ps and, in such a way, is separated from the decay time bias discussion.

Under the same assumptions as Eq. 4.39, the convolution integral in Eq. 4.42 becomes,

$$\begin{aligned}\mathcal{P}^{\text{obs}}(t) \;\approx\; \Gamma_s e^{-\Gamma_s(t)} H(t) \Big[1 - C \int \mathrm{d}t' R(t-t') \cos(\Delta m_s t') \\ + S \int \mathrm{d}t' R(t-t') \sin(\Delta m_s t') \Big].\end{aligned} \tag{4.43}$$

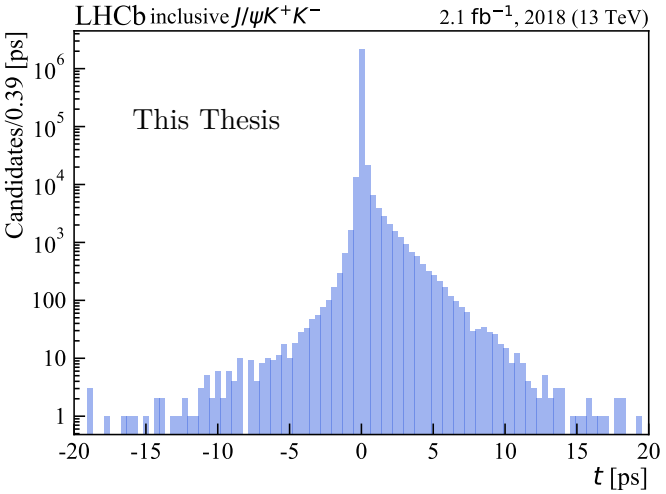


Figure 4.24: Decay time distribution of the inclusive $J/\psi K^+K^-$ control sample, collected in 2018.

Using the fact that $\sin(t)$ is asymmetric, while $R(t)$ and $\cos(t)$ are symmetric, one finds,

$$\int_{-\infty}^{\infty} \mathrm{d}t'\ R(t-t')\ \sin(\Delta m_s t') = \ \sin(\Delta m_s(t))\ \int_{-\infty}^{\infty} \mathrm{d}t'\ R(t')\ \cos(\Delta m_s t'). \tag{4.44}$$

The integral in front of the $\sin(\Delta m t)$ is denoted as dilution D_{tr},

$$D_{tr} \equiv \int_{-\infty}^{\infty} \mathrm{d}t'\ R(t')\ \cos(\Delta m_s t'). \tag{4.45}$$

The same expression for the dilution is obtained when repeating this exercise for a $\cos(\Delta m_s t')$ term. The dilution integral is evaluated numerically from the negative decay time distribution of the *inclusive* $J/\psi K^+K^-$ control sample, as

$$D_{tr} \ = \ \frac{1}{n}\sum_{i=1}^{n} \cos(\Delta m_s\,(t_i)),\ \text{for}\ t_i < 0, \tag{4.46}$$

where i is the index of a candidate in the control sample; n is the number of candidates in the control sample; t_i is the reconstructed decay time of a candidate in the control sample.

As was mentioned above, the benefit of using Eq. 4.46, rather than using the decay time resolution fit method from the previous analysis [49], is

$\sigma_t \in$	D_{tr}^{2015}	D_{tr}^{2016}	D_{tr}^{2017}	D_{tr}^{2018}
$(0.01, 0.021]$ ps	0.8487 ± 0.0035	0.8767 ± 0.0028	0.8920 ± 0.0029	0.8911 ± 0.0029
$(0.021, 0.026]$ ps		0.8423 ± 0.0014	0.8583 ± 0.0014	0.8660 ± 0.0013
$(0.026, 0.032]$ ps	0.8030 ± 0.0024	0.8019 ± 0.0009	0.8261 ± 0.0008	0.8307 ± 0.0007
$(0.032, 0.038]$ ps	0.7513 ± 0.0021	0.7542 ± 0.0008	0.7766 ± 0.0007	0.7815 ± 0.0006
$(0.038, 0.044]$ ps	0.7033 ± 0.0023	0.7042 ± 0.0008	0.7257 ± 0.0007	0.7307 ± 0.0007
$(0.044, 0.049]$ ps	0.6551 ± 0.0031	0.6590 ± 0.0011	0.6695 ± 0.0010	0.6816 ± 0.0010
$(0.049, 0.054]$ ps	0.6276 ± 0.0040	0.6136 ± 0.0014	0.6256 ± 0.0014	0.6254 ± 0.0014
$(0.054, 0.059]$ ps	0.5627 ± 0.0057	0.5663 ± 0.0020	0.5705 ± 0.0021	0.5738 ± 0.0020
$(0.059, 0.064]$ ps	0.4782 ± 0.0058	0.5233 ± 0.0027	0.5141 ± 0.0031	0.5202 ± 0.0030
$(0.064, 0.008)$ ps		0.4377 ± 0.0029	0.4374 ± 0.0034	0.4464 ± 0.0033

Table 4.9: Dilution of inclusive $J/\psi K^+K^-$ sample collected in 2015-2018.

that the estimated dilution incorporates the non-Gaussian resolution effects. However, this method only describes the dilution of the oscillation terms in the PDF, see Eq. 4.39. The exact impact of the non-Gaussian tails on the measured lifetimes is ignored. This impact is described in **Lifetime bias from decay time resolution** paragraph of Sec. 4.6.

Decay time resolution of the inclusive $J/\psi K^+K^-$ sample

The estimated dilutions from the inclusive $J/\psi K^+K^-$ sample can be found in Tab. 4.9 for different bins of per-event decay time uncertainty. They are computed with Eq. 4.46.

First, these dilutions are corrected for the long-lived background contamination of the inclusive $J/\psi K^+K^-$ sample.

Long-lived background correction in inclusive $J/\psi K^+K^-$

The equation 4.46 is valid under the assumption that the left tail of Fig. 4.24 contains only the *prompt* $J/\psi K^+K^-$ candidates. However, the long-lived candidates with short decay times can migrate to the left tail due to the resolution effects and contaminate it. The long-lived candidate contamination is accounted for by correcting the effective dilution to get the dilution of prompt candidates D_{prompt},

$$D_{\text{prompt}} = \frac{1}{f_{\text{prompt}}} \left(D_{\text{tr}} - f_{\text{ll}} D_{\text{ll}} \right), \tag{4.47}$$

where $D_{\rm tr}$ is the dilution computed with Eq. 4.46 from the control sample; $f_{\rm ll}$ is the fraction of long-lived candidates in the control sample and D^{ll} is their effective dilution; f_{prompt} is the fraction of the inclusive $J/\psi K^+K^-$ candidates.

The parameters $f_{\rm prompt}$, $f_{\rm ll}$ and $D_{\rm ll}$ are estimated from a fit to the decay time distribution of the control sample taken from [49],

$$\mathcal{P}(t) = \mathcal{G}(t)\otimes[f_{\rm prompt}\delta(t)+f_{\rm ll}(f_{\rm sl}e^{-t/\tau_1}+(1-f_{\rm sl})e^{-t/\tau_2})]+f_{\rm WPV}W(t) \quad (4.48)$$

where $f_{\rm WPV}$ is a fraction of non-Gaussian effects and $W(t)$ is an empirical model for candidates that have been assigned to the wrong primary vertex. The latter model is determined using the `PVMixer` tool. `PVMixer` tool adds a randomly chosen primary vertex from another event into the current event. The best primary vertex is researched for, and if the newly added primary vertex is chosen as the best, the candidate is flagged as the "wrong primary vertex". The wrongly assigned primary vertices shape is derived empirically from a sample filled with candidates flagged "wrong primary vertex". Figure 4.25 shows the inclusive $J/\psi K^+K^-$ candidates marked having the "wrong primary vertex", which were created with the `PVMixer` tool. The model shown in blue is the empirical shape, defined as,

$$W(t) = f_{{\rm WPV},\tau}e^{-|t|/\tau_1} + (1 - f_{{\rm WPV},\tau})e^{-|t|/\tau_2}. \quad (4.49)$$

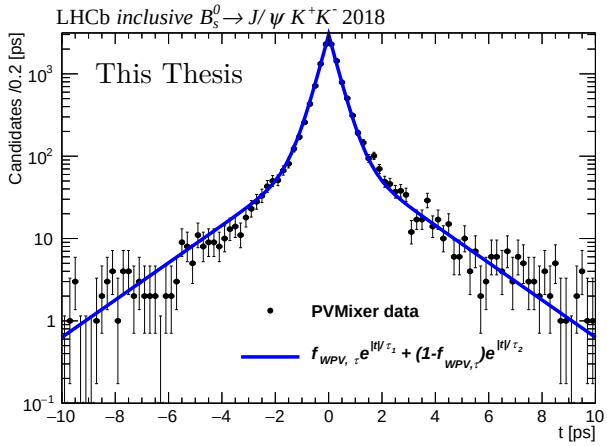


Figure 4.25: The inclusive $J/\psi K^+K^-$ candidates marked "wrong primary vertex" created with `PVMixer` tool. Shown for 2018 data-taking year.

The model in Eq. 4.48 has one important problem, it treats the non-Gaussian effects from the wrong primary vertex association as non-resolution effects. This is strictly speaking incorrect, but considering the small size of the long-lived correction to the dilution, the possible bias from treating the wrong primary vertex association as a non-resolution effect is assumed to be negligible and is ignored. In the next iteration of this analysis, one could try a convolution of exponential and Gaussian shapes to correctly describe the resolution in this fit.

From the fit of Eq. 4.48 to the control sample, it is estimated that only 2-5% of the long-lived background contributes to the negative decay time depending on the decay time uncertainty bin. The D_{ll} is computed using Eq. 4.46 from the sampled long-lived background shape, which is taken from the result of the decay time fit. The shape is shown in Fig. 4.26. Table 4.10 shows the long-lived background dilution D_{ll}, the long-lived background fraction f_{ll} in the negative decay times, and the corrected dilution D_{prompt}.

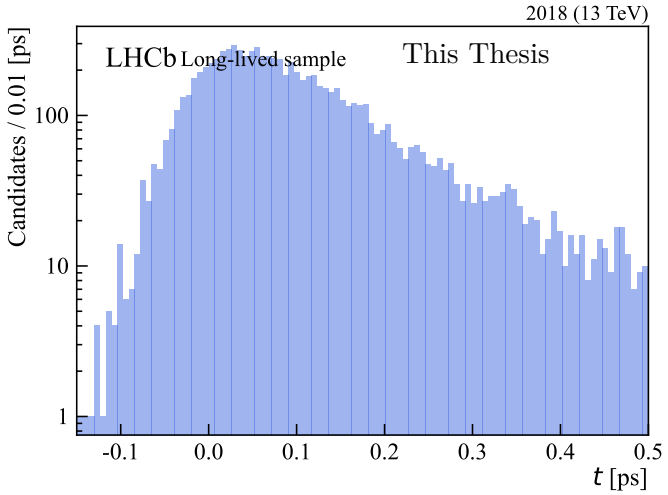


Figure 4.26: Long-lived component sampled from the 2018 decay time resolution shape. The shift towards positive decay time is the consequence of convolutions with Gaussian decay time resolution.

Compatibility between inclusive $J/\psi K^+K^-$ and $B_s^0 \to J/\psi K^+K^-$ time resolution and translation to $B_s^0 \to J/\psi K^+K^-$

The prompt $J/\psi K^+K^-$ candidates that are used for the resolution analysis are produced via a different mechanism than the $B_s^0 \to J/\psi K^+K^-$ decays.

$\sigma_t \in$	(0.01, 0.026) ps	(0.026, 0.032) ps	(0.032, 0.038) ps	(0.038, 0.044) ps	(0.044, 0.049) ps	(0.049, 0.054) ps	(0.054, 0.059) ps	(0.059, 0.08) ps
$D_{\rm li}$, 2015	0.9161 ± 0.0063	0.8852 ± 0.0077	0.8360 ± 0.0108	0.7887 ± 0.0095	0.7563 ± 0.0108	0.7061 ± 0.0090	0.6673 ± 0.0108	0.5936 ± 0.0143
$f_{\rm li}$, 2015	0.0203	0.0247	0.0297	0.0303	0.0336	0.0282	0.0349	0.0431
$D_{\rm prompt}$, 2015	0.8473 ± 0.0036	0.8009 ± 0.0024	0.7487 ± 0.0022	0.7006 ± 0.0024	0.6516 ± 0.0032	0.6253 ± 0.0041	0.5589 ± 0.0059	0.4730 ± 0.0061

$\sigma_t \in$	(0.01, 0.021) ps	(0.021, 0.026) ps	(0.026, 0.032) ps	(0.032, 0.038) ps	(0.038, 0.044) ps	(0.044, 0.049) ps	(0.049, 0.054) ps	(0.054, 0.059) ps	(0.059, 0.064) ps	(0.064, 0.08) ps
$D_{\rm li}$, 2016	0.9162 ± 0.0063	0.9070 ± 0.0063	0.8374 ± 0.0038	0.8126 ± 0.0096	0.7922 ± 0.0091	0.7325 ± 0.0101	0.7110 ± 0.0104	0.6579 ± 0.0116	0.6115 ± 0.0128	0.5601 ± 0.0128
$f_{\rm li}$	0.0182	0.0215	0.0100	0.0279	0.0300	0.0311	0.0336	0.0361	0.0385	0.0410
$D_{\rm prompt}$, 2016	0.8759 ± 0.0028	0.8409 ± 0.0014	0.8015 ± 0.0009	0.7525 ± 0.0008	0.7015 ± 0.0009	0.6567 ± 0.0012	0.6102 ± 0.0015	0.5628 ± 0.0021	0.5197 ± 0.0029	0.4324 ± 0.0030
$D_{\rm li}$, 2017	0.9441 ± 0.0048	0.9255 ± 0.0057	0.8797 ± 0.0069	0.8320 ± 0.0088	0.8138 ± 0.0080	0.6760 ± 0.0061	0.7143 ± 0.0094	0.6713 ± 0.0102	0.6138 ± 0.0121	0.5624 ± 0.0112
$f_{\rm li}$	0.0169	0.0197	0.0228	0.0264	0.0284	0.0067	0.0311	0.0338	0.0372	0.0363
$D_{\rm prompt}$, 2017	0.8911 ± 0.0029	0.8569 ± 0.0015	0.8249 ± 0.0008	0.7751 ± 0.0007	0.7231 ± 0.0008	0.6694 ± 0.0010	0.6228 ± 0.0015	0.5670 ± 0.0022	0.5103 ± 0.0032	0.4327 ± 0.0035
$D_{\rm li}$, 2018	0.9341 ± 0.0059	0.9191 ± 0.0062	0.8849 ± 0.0075	0.8365 ± 0.0087	0.8166 ± 0.0082	0.7526 ± 0.0096	0.7244 ± 0.0106	0.6777 ± 0.0115	0.6197 ± 0.0134	0.5696 ± 0.0110
$f_{\rm li}$, 2018	0.0172	0.0203	0.0229	0.0263	0.0284	0.0304	0.0339	0.0365	0.0399	0.0353
$D_{\rm prompt}$, 2018	0.8903 ± 0.0029	0.8649 ± 0.0013	0.8295 ± 0.0008	0.7800 ± 0.0007	0.7282 ± 0.0008	0.6770 ± 0.0010	0.6219 ± 0.0015	0.5699 ± 0.0021	0.5161 ± 0.0031	0.4419 ± 0.0035

Table 4.10: Long-lived background dilution corrections for all data taking years.

Therefore, they have different kinematics than the $B_s^0 \to J/\psi K^+K^-$ candidates. Figure 4.27 shows a comparison of both samples for $p(B_s^0)$, $p_{\mathrm{T}}(B_s^0)$ and σ_t. The slight difference is apparent. This raises the question of how

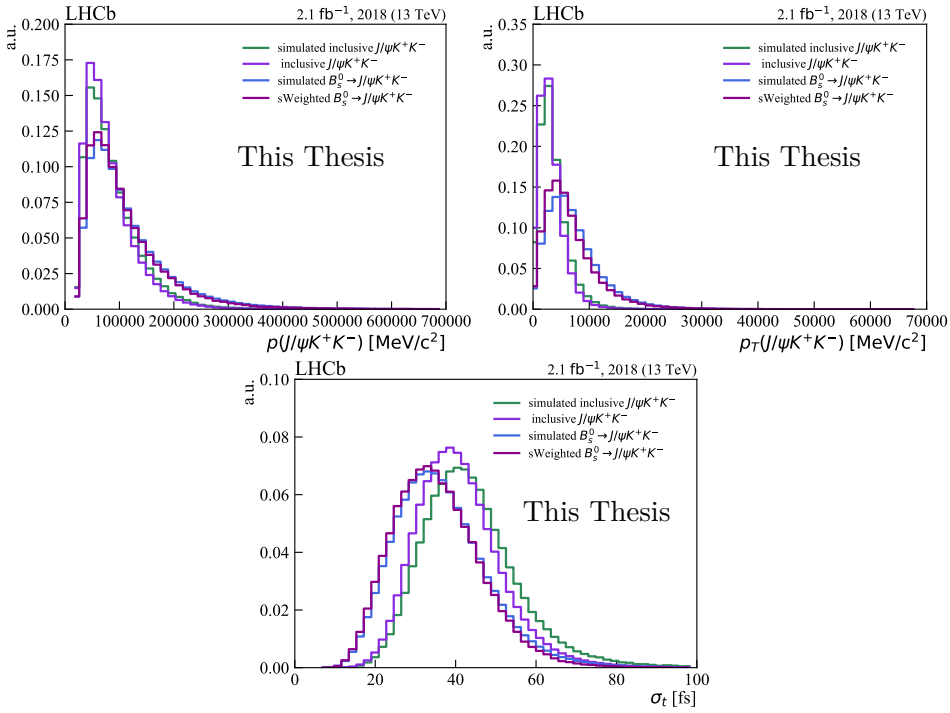
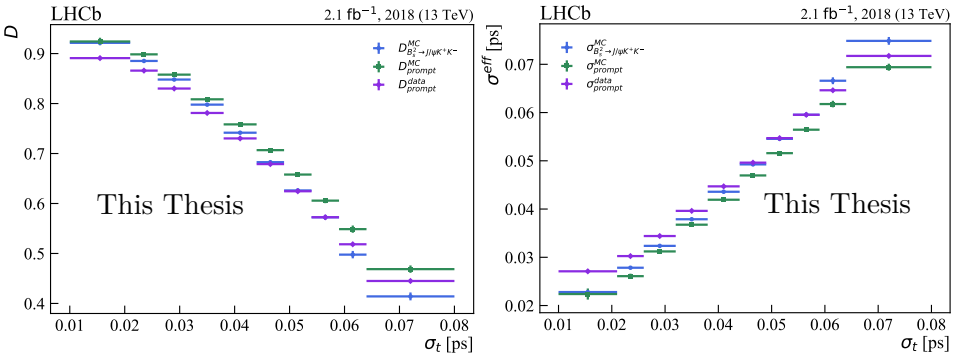


Figure 4.27: The inclusive $J/\psi K^+K^-$ and $B_s^0 \to J/\psi K^+K^-$ $p(B_s^0)$, $p_T(B_s^0)$ and σ_t for 2018.

representative the prompt $J/\psi K^+K^-$ candidate time resolution is of the signal $B_s^0 \to J/\psi K^+K^-$ sample decay time resolution. Figure 4.28 shows the effective decay time resolution width of simulated $B_s^0 \to J/\psi K^+K^-$ sample, simulated inclusive $J/\psi K^+K^-$ sample, and real inclusive $J/\psi K^+K^-$ sample. For the inclusive samples in both cases, the resolution of prompt candidates is shown.

To account for small deviations between samples, the dilution of the real data prompt candidates is translated to the signal sample dilution $D^{\mathrm{data}}_{B_s^0 \to J/\psi K^+K^-}$ as,

$$D^{\mathrm{data}}_{B_s^0 \to J/\psi K^+K^-} = D^{\mathrm{data}}_{\mathrm{prompt}} \cdot \frac{D^{\mathrm{MC}}_{B_s^0 \to J/\psi K^+K^-}}{D^{\mathrm{MC}}_{\mathrm{prompt}}} \tag{4.50}$$



(a) D (b) σ^{eff}

Figure 4.28: The effective dilutions (a) and decay time resolution (b) of the $B^0_s \to J/\psi K^+K^-$ simulated sample, the prompt candidates in the inclusive $J/\psi K^+K^-$ simulated and real samples.

where $D^{\text{data}}_{\text{prompt}}$ is the dilution of the prompt candidates in the inclusive $J/\psi K^+K^-$ sample; $D^{\text{MC}}_{B^0_s\to J/\psi K^+K^-}$ is the dilution of the simulated $B^0_s \to J/\psi K^+K^-$ candidates; $D^{\text{MC}}_{\text{prompt}}$ is the dilution of the simulated prompt $J/\psi K^+K^-$ candidates.

In Fig. 4.29, the translation factors $D^{MC}_{B^0_s\to J/\psi K^+K^-}/D^{MC}_{\text{prompt}}$ are shown in bins of the measured per-event decay time uncertainty. The correction factor is close to unity for most of the decay time uncertainty range.

Decay time resolution calibration for the $B^0_s \to J/\psi K^+K^-$ analysis

The effective decay time resolution width is derived from the corrected dilution $D^{\text{data}}_{B^0_s\to J/\psi K^+K^-}$ via the Eq. 4.37.The decay time uncertainty is then calibrated to the effective decay time resolution width using a first-order polynomial,

$$\sigma_{\text{eff}} = p_0 + p_1(\sigma_t - \sigma_t^0) \tag{4.51}$$

where $\sigma_t^0 = 0.036\,\text{ps}$ is the average σ_t of the simulated $B^0_s \to J/\psi K^+K^-$ sample.

The calibrations are shown in Fig. 4.30 and Tab. 4.11. The total effective decay time resolution $\langle\sigma_{\text{eff}}\rangle$ is used to compare decay time resolution between

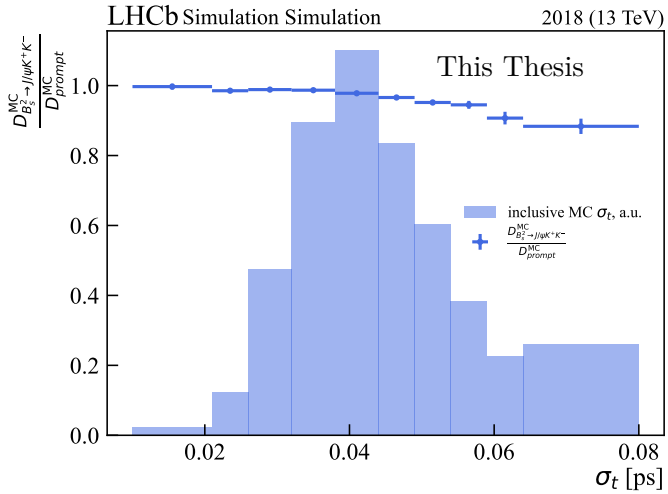


Figure 4.29: $D^{\mathrm{MC}}_{B_s^0\to J/\psi\, K^+K^-}/D^{\mathrm{MC}}_{\mathrm{prompt}}$ correction per decay time uncertainty bin for 2018.

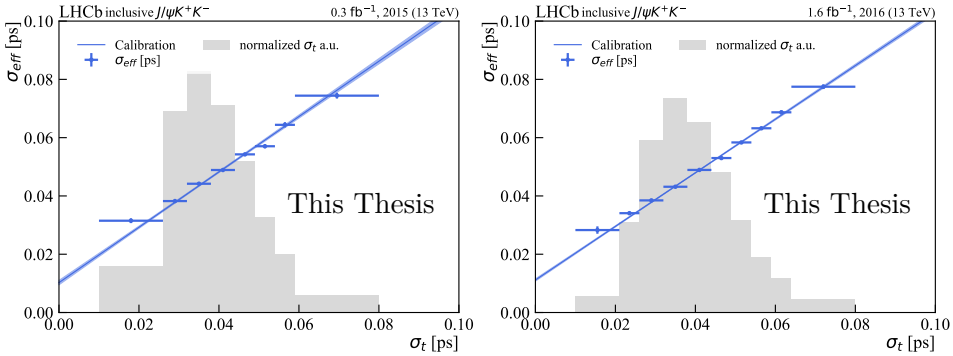
the data taking years in Tab. 4.12,

$$\langle\sigma_{\mathrm{eff}}\rangle = \frac{\sum_i^n w_i\sigma_i^{eff}}{\sum_i^n w_i}$$
$$\Delta\sigma_{\mathrm{eff}} = \left[\frac{\sum_i^n w_i^2(\sigma_i^{eff} - \langle\sigma_{\mathrm{eff}}\rangle)^2}{(\sum_i^n w_i)^2}\right]^{1/2} \tag{4.52}$$

where σ_i^{eff} is the effective decay time resolution of the i^{th} candidate computed using the calibration in Eq.4.51; w_i is the sWeight of the i^{th} candidate; n is the total number of candidates; $\Delta\sigma_{\mathrm{eff}}$ is the standard deviation of $\langle\sigma_{\mathrm{eff}}\rangle$. The effective dilutions D_{eff} are computed using Eq. 4.37. In 2017, it was found that the uncertainty assigned to tracking parameters in the LHCb VELO was not fully representative of the detector at that time. Around that time, the description of uncertainties in VELO was improved. An improvement in decay time resolution in 2017-2018, as seen in Tab. 4.12, is associated with this event.

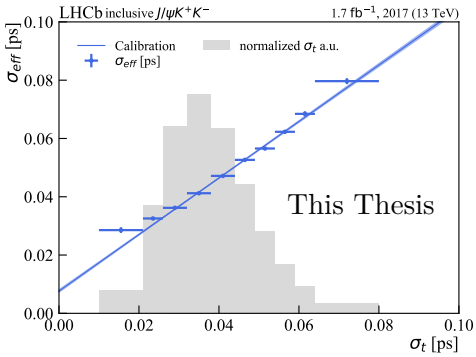
Is an exact description of the decay time resolution possible?

As was discussed before, the decay time distribution of the prompt $J/\psi\, K^+K^-$ candidates is a great proxy for the decay time resolution function. Analytically describing the distribution in Fig. 4.24 is a challenging task. The RooFit

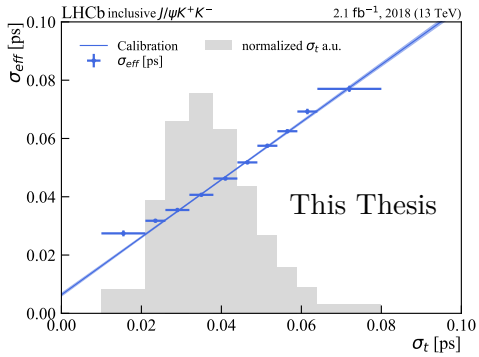


(a) 2015

(b) 2016



(c) 2017



(d) 2018

Figure 4.30: Decay time resolution calibration for the prompt candidates. Estimated with the inclusive $J/\psi K^+K^-$ sample.

year	p_0, ps	p_1	$\rho(p_0, p_1)$
2015	0.0445 ± 0.0002	0.9492 ± 0.0202	-0.3614
2016	0.0443 ± 0.0001	0.9194 ± 0.0102	-0.3440
2017	0.0425 ± 0.0001	0.9676 ± 0.0122	-0.2750
2018	0.0419 ± 0.0001	0.9846 ± 0.0130	-0.2669

Table 4.11: Decay time resolution calibration for translated $B_s^0 \to J/\psi K^+K^-$.

year	D_{eff}	σ^{eff} , fs
2015	0.7115 ± 0.1157	46.51 ± 11.11
2016	0.7090 ± 0.1147	46.75 ± 11.00
2017	0.7441 ± 0.1139	43.34 ± 11.22
2018	0.7536 ± 0.1122	42.40 ± 11.16

Table 4.12: Total effective dilution D_{eff} (Eq. 4.37) and decay time resolution width σ^{eff} (Eq. 4.52) for $B_s^0 \to J/\psi K^+K^-$ candidates.

framework [84], widely employed in particle physics for maximum likelihood fits, allows for the construction of probability distribution models from the binned data samples. This can be used to directly convert the decay time distribution from Fig. 4.24 into a resolution model by simply mirroring the left tail around the y-axis. The model can then be numerically convoluted with the physics PDF via fast Fourier transform. The main advantage of this approach is that the obtained decay time resolution function is exact and model-independent. The function preserves the non-Gaussian effects and their impact on the lifetimes, therefore potentially eliminating the limitation of the current procedure. However, a cautionary note is warranted: the control sample contains long-lived candidates. They should be statistically subtracted from the control sample if this approach is to be used.

4.5.2 Decay time bias

Due to the changes in the detector with time, the alignment applied to the VELO detector tracks was a tiny bit off. This led to a bias in the reconstructed decay length and, by extension, in the measured decay time in 2015-2018. This effect, known as *"decay time bias"*, was first described in the Δm_s measurement using $B_s^0 \to D_s^-\pi^+$ decays [81]. This bias can be observed in the inclusive $J/\psi K^+K^-$ sample, where the measured decay time distribution is not centered at 0 but at some shift μ. This shift is accounted for in the decay time resolution procedure by shifting the entire decay time distribution back to 0. The decay time resolution procedure estimates the width of the decay time resolution Gaussian in the final PDF. However, from the fit of Eq. 4.48, the observed shift for the prompt candidates in the inclusive $J/\psi K^+K^-$ sample is extracted. However, it is not clear how this shift translates to $B_s^0 \to J/\psi K^+K^-$. The time bias correction procedure described in this section calibrates the shift of the prompt candidates to the shift of $B_s^0 \to J/\psi K^+K^-$ candidates under the same simulated VELO

misalignment conditions. This calibration is then used to translate the shift observed in the fit of the inclusive $J/\psi K^+K^-$ sample with Eq. 4.48 to the shift in $B_s^0 \to J/\psi K^+K^-$ data. The mean of the decay time resolution Gaussian in the final PDF is set to the latter.

A set of $B_s^0 \to J/\psi K^+K^-$ samples is simulated with the constant misalignment ΔT_x of two VELO halves in x-coordinate (LHCb lab frame [8]). Figure 4.31 shows the difference between measured and true decay time ($\Delta t = t - t_{true}$), or decay time resolution distribution, in the $B_s^0 \to J/\psi K^+K^-$ simulated samples with a misalignment of $\pm 10\,\mu$m. The shift of the decay time resolution distribution in the Δt distributions is visible.

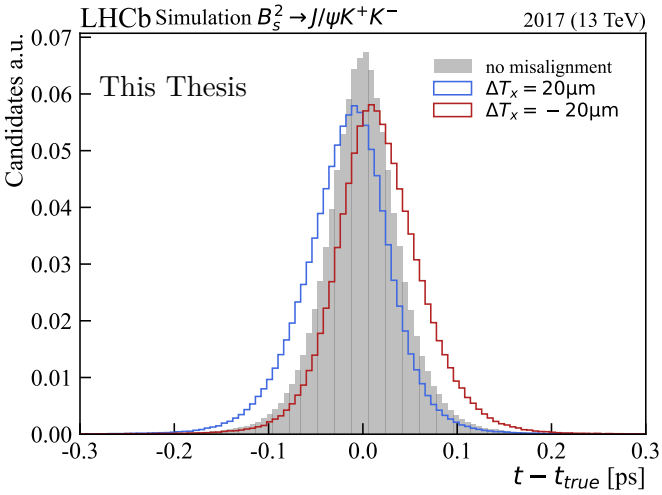


Figure 4.31: Distribution of $t - t_{true}$ for $B_s^0 \to J/\psi K^+K^-$ simulated samples with no misalignment, 20 μm and -20 μm misalignment in the x-coordinate of VELO halves in the LHCb lab frame [8].

In the Δm_s measurement with $B_s^0 \to D_s^-\pi^+$ decays [81] the misalignment between VELO-halves, led to a bias of $2.7\,\text{ps}^{-1}$ in Δm_s. The same effect biases the value of Δm_s in this analysis. It also strongly distorts the values of $\delta_\perp - \delta_0$ and $|\lambda|$. Both parameters are strongly correlated with Δm_s.

To account for the decay time bias, it is included as the mean of the decay time resolution function in the full PDF. The decay time bias is estimated following the procedure from [81],

1. A set of misaligned simulated inclusive $J/\psi K^+K^-$ and $B_s^0 \to J/\psi K^+K^-$ samples is used to get a calibration curve between the observed decay time bias of the prompt candidates and $B_s^0 \to J/\psi K^+K^-$

candidates.

2. The decay time bias is evaluated from the real prompt candidates by fitting the inclusive $J/\psi K^+K^-$ sample with Eq. 4.48. It is then translated to $B_s^0 \to J/\psi K^+K^-$ via the calibration curve in step 1.

3. The mean of the decay time resolution function in the decay time angular fit is set to the decay time bias of $B_s^0 \to J/\psi K^+K^-$.

The shift of the Δt distribution of the decay time distribution of the misaligned prompt $J/\psi K^+K^-$ candidates and $B_s^0 \to J/\psi K^+K^-$ candidates is plotted in Fig. 4.32. From [81], it is known that the VELO was biased in the positive direction of ΔT_x; therefore, only the simulated samples with $\Delta T_x \in [0, 15]$ µm were used.

The effect of misalignment on the prompt $J/\psi K^+K^-$ candidates is expected to be different from the effect the same misalignment has on the $B_s^0 \to J/\psi K^+K^-$ candidates. The estimated decay time depends on the decay lengths and momentum of a particle. When the VELO is misaligned, the tracks are reconstructed with an offset, which offsets the vertices positions and, therefore, the measured decay length. From Eq. 4.40, it can be seen that even if the offset is the same, particles with different momentum will experience a different bias in the estimated decay time. Figure 4.27 shows the difference between the momentum distributions in the signal $B_s^0 \to J/\psi K^+K^-$ sample and the inclusive $J/\psi K^+K^-$ sample, which is mainly populated with the prompt $J/\psi K^+K^-$ candidates. Figure 4.32 confirms that, indeed, the resulting decay time bias is different, even if the misalignment is the same.

The calibration curve from [67] is also shown in Fig. 4.32 and is given by,

$$\mu_{B_s^0 \to J/\psi K^+K^-} = (0.59 \pm 0.19)[\,\text{fs}] + (0.93 \pm 0.02) \times \mu_{\text{prompt}}. \tag{4.53}$$

Using this calibration, the decay time biases observed for the prompt candidates in the real inclusive $J/\psi K^+K^-$ sample are translated to $B_s^0 \to J/\psi K^+K^-$ sample. Their calibrated values for $B_s^0 \to J/\psi K^+K^-$ are reported in Tab. 4.13.

The VELO misalignment is, however, run dependent. Ideally, an individual shift correction should be derived and applied based on the run number. This, however, required additional information, which was not preserved. In particular, the raw banks of the tracking stations should be saved for all the collected events to allow for the correction of decay time biasing effects from misalignment. In future measurements, the complete original information of all tracking detectors should be retained.

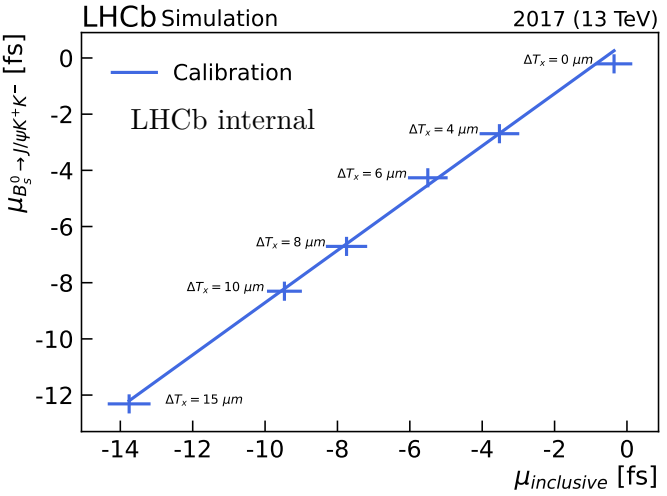


Figure 4.32: Decay time bias calibration curve for 2017 simulated samples. Taken from [67].

	$\mu_{\text{prompt } J/\psi K^+K^-}$, fs	$\mu_{B_s^0 \to J/\psi K^+K^-}$, fs
2015	-4.83±0.13	-3.90±0.10
2016	-6.14±0.05	-5.12±0.04
2017	-7.18±0.04	-6.09±0.04
2018	-5.88±0.05	-4.88±0.04

Table 4.13: Decay-time bias in the inclusive $J/\psi K^+K^-$ sample and the calibrated $B_s^0 \to J/\psi K^+K^-$ decay time bias. Taken from [67].

4.5.3 Reweighting for the acceptance

Parametrization for both decay time and angular acceptance is derived from the simulated samples. The simulated distributions differ from the measured distributions in real data. This can be due to either the detector description or the event generation. Therefore, the simulated samples are reweighted to represent the data. The reweighting procedure for the samples used in acceptance studies is complicated. It consists of many different weights, each correcting for a particular effect. For clarity, these weights are described here. The application of these weights and slight differences between their derivations for different samples are described in the acceptances estimation sections.

Kinematic weights - `kinWeight` Kinematics weights correct for the

differences in the kinematics of simulated and real data collected candidates. The weights are derived using the gradient-boosted decision tree, as in Sec. 4.2. These weights are usually applied to the simulated samples. However, the kinematics difference between the pion track in $B^0 \to J/\psi K^{*0}$ decay and the kaon track in $B_s^0 \to J/\psi K^+K^-$ decay leads to small differences in the observed decay time acceptance. To correct for those kinematics differences, the B-meson momentum and transverse momentum are reweighted in the $B^0 \to J/\psi K^{*0}$ background subtracted real data sample to the background subtracted $B_s^0 \to J/\psi K^+K^-$ real data sample.

Both $B_s^0 \to J/\psi K^+K^-$ and $B^0 \to J/\psi K^{*0}$ simulated samples have their B meson momentum and K^+K^- or $K^-\pi^+$ invariant mass reweighted. For $B_s^0 \to J/\psi K^+K^-$ simulation, the background subtracted $B_s^0 \to J/\psi K^+K^-$ real data sample is used as a target. For $B^0 \to J/\psi K^{*0}$ simulation, the background subtracted and kinematically reweighted to the $B_s^0 \to J/\psi K^+K^-$ real data sample of $B^0 \to J/\psi K^{*0}$ is used as a target.

PDF weights - `PDFWeight` Simulated events are generated using specific values for the parameters of interest in the underlying physics PDF. These parameters are not necessarily equal to those in real data. Because of these two discrepancies between the samples, the angular distributions of final state particles differ in the simulated sample from real data. This impacts both the angular and decay time acceptances. Given the assumption that angular and decay time acceptances factorize and can be independently extracted, a set of weights is derived to correct the above-mentioned discrepancies for the physics parameters. The weights are defined as,

$$w_{\text{pdf}} = \frac{\text{pdf}(\text{t}_{\text{gen}}, \Omega_{\text{gen}}|\theta^{\text{obs}})}{\text{pdf}(\text{t}_{\text{gen}}, \Omega_{\text{gen}}|\theta^{\text{gen}})} \tag{4.54}$$

where t_{gen} and Ω_{gen} are the generated decay time and angles; θ^{obs} are the observed PDF parameters and θ^{gen} are the parameters used during the generation of the simulated sample. The observed PDF parameters are taken from the previous LHCb measurements: [48] for $B_s^0 \to J/\psi K^+K^-$ and [85] for $B^0 \to J/\psi K^{*0}$.

Luminosity weights - `lumiWeight` The simulated samples do not match the luminosity that was collected in the real sample. In particular, the same number of events are generated with two LHCb magnet polarity (see Chapter 3), whereas the real data has a different ratio. Therefore, relative weights are applied to correct for this difference.

The combinatorial background is statistically subtracted in the real $B_s^0 \to J/\psi K^+K^-$ samples; see Sec. 4.3. The same procedure is applied to the $B^0 \to J/\psi K^{*0}$ sample and simulated samples.

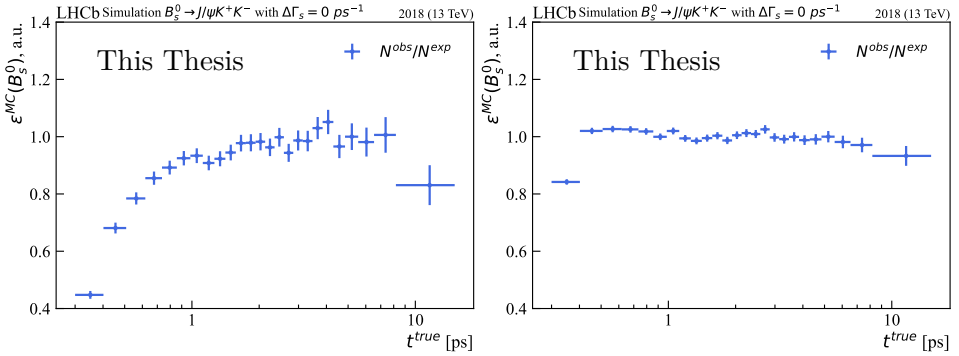
sWeights - `sw` Prior to the simulated sample kinematic reweighting (see above), backgrounds are subtracted from the real data samples using the sPlot technique. For $B_s^0 \to J/\psi K^+K^-$ real data sample, the weights are derived as described in Sec. 4.3. For $B^0 \to J/\psi K^{*0}$ the $J\psi K^-\pi^+$ mass is used as a discriminating variable. The details on the sPlot procedure for $B^0 \to J/\psi K^{*0}$ can be found in [67]. The sPlot technique is also applied to both simulation samples to remove the combinatorial background from the sample.

4.5.4 Decay time acceptance

Imperfect description decay time acceptance leads to bias in the estimated lifetime. To correct for it, the decay time acceptance has to be estimated accurately. The decay time acceptance can be estimated from simulated samples, for which the generated decay time distribution is known. The binned decay time acceptance $\varepsilon_i^{\rm MC}$ in simulation is,

$$\varepsilon_i^{\rm MC} = \frac{N_i^{\rm obs}}{N_i^{\rm exp}}, \tag{4.55}$$

where $N_i^{\rm obs}$ is the observed number of candidates in the i^{th} true decay time bin and $N_i^{\rm exp}$ is the number of candidates expected in that true decay time bin. Figure 4.33 shows the decay time acceptance in the simulated $B_s^0 \to J/\psi K^+K^-$ sample in two trigger categories.



(a) biased (b) unbiased

Figure 4.33: Ratio of reconstructed decay time distribution to the generated decay time distribution for $B_s^0 \to J/\psi K^+K^-$ 2018 simulated sample with $\Delta\Gamma_s = 0$ ps^{-1}. The normalisation is arbitrary and is chosen to have $\varepsilon^{\rm MC}_{B_s^0 \to J/\psi K^+K^-}(5\,{\rm ps}) = 1$.

The main contribution to the shape of the lower decay time acceptance ($t < 3$ ps) comes from the trigger requirements. The candidates in the **"biased"** category pass displacement requirements in both the HLT1 and HLT2 triggers, which results in the biased decay time distribution, as can be seen in Fig. 4.33. The second, **"unbiased"** category, has almost a perfect flat decay time acceptance. Due to a significant difference in the trigger categories' decay time acceptance shapes, the decay time acceptance is defined for the trigger categories separately.

The high decay times inefficiencies are introduced by the reconstruction of VELO tracks [86]. The VELO track reconstruction assumes that the VELO tracks originate at the beam axis. This assumption is not applicable for very displaced tracks, leading to a drop in the efficiency at the higher decay times.

Decay time acceptance estimation

Even though, in principle, the decay time acceptance could be estimated from simulation, because of the observed differences between the simulated sample and the real data sample, a data-driven method is preferable. The decay time acceptance $\varepsilon^{\text{data}}(t)$ is determined from the ratio of the reconstructed decay time distribution and the known expected decay time distribution with perfect decay time acceptance,

$$\varepsilon^{\text{data}}(t) \propto \frac{N^{obs}(t)}{e^{-\Gamma t} \otimes G(t|\mu,\sigma)} \tag{4.56}$$

where $N^{obs}(t)$ is the normalized number of observed events as a function of measured decay time; Γ is decay width; $G(t|\mu,\sigma)$ is Gaussian decay time resolution. It is assumed that the angular and decay time acceptances factorize; see Eq. 4.6.

Decay time acceptance control sample: $B^0 \to J/\psi K^{*0}$ sample

The $B^0 \to J/\psi K^{*0}$ control sample, where the K^{*0} decays into K^+ and π^-, is used to estimate the decay time acceptance from the data. The $B^0 \to J/\psi K^{*0}$ topology closely resembles $B_s^0 \to J/\psi K^+K^-$ and, therefore, has a similar decay time acceptance as the signal $B_s^0 \to J/\psi K^+K^-$ sample. The decay width difference in the B^0 system is negligible and can be essentially set to zero [87]. Assuming that $\Delta\Gamma_d$ can be set to zero, the expected decay time distribution for $B^0 \to J/\psi K^{*0}$ is a simple single exponential distribution with an average lifetime set to the world-average B^0 lifetime [87].

The $B^0 \to J/\psi K^{*0}$ selection criteria are similar to the $B_s^0 \to J/\psi K^+K^-$ ones, with differences related to having π^- being in the final state, rather than a K^-; and a K^{*0}, rather than a $\phi(1020)$ as intermediate resonance. The differences in selection can be found in [67].

Even though $B^0 \to J/\psi K^{*0}$ is topologically quite close to $B_s^0 \to J/\psi K^+K^-$, there are differences arising from selection and replacing a K^- by π^- in the final state. To correct for these differences, the decay time acceptance for $B_s^0 \to J/\psi K^+K^-$ ($\varepsilon^{\text{data}}_{B_s^0 \to J/\psi K^+K^-}(t)$) is estimated simultaneously from $B^0 \to J/\psi K^{*0}$ real data sample, simulated $B^0 \to J/\psi K^{*0}$ sample, and simulated $B_s^0 \to J/\psi K^+K^-$ sample,

$$\varepsilon^{\text{data}}_{B_s^0 \to J/\psi K^+K^-}(t) = \varepsilon^{\text{data}}_{B^0 \to J/\psi K^{*0}}(t) \times \frac{\varepsilon^{\text{MC}}_{B_s^0 \to J/\psi K^+K^-}(t)}{\varepsilon^{\text{MC}}_{B^0 \to J/\psi K^{*0}}(t)} \tag{4.57}$$

where t is reconstructed time, $\varepsilon^{\text{data}}_{B^0 \to J/\psi K^{*0}}(t)$ is the decay time acceptance of the real $B^0 \to J/\psi K^{*0}$ sample; $\varepsilon^{\text{MC}}_{B_s^0 \to J/\psi K^+K^-}(t)$ and $\varepsilon^{\text{MC}}_{B^0 \to J/\psi K^{*0}}(t)$ are decay time acceptances of reweighted simulated $B_s^0 \to J/\psi K^+K^-$ and $B^0 \to J/\psi K^{*0}$ respectively.

Contrary to a simple exponential expected decay time distribution of $B^0 \to J/\psi K^{*0}$ decays, $B_s^0 \to J/\psi K^+K^-$ expected decay time distribution for an untagged sample is,

$$\frac{d\Gamma(t)_{exp}}{dt} \propto e^{-\Gamma_s t} \sum_{k=1}^{10} \omega_k N_k \left(a_k \cosh(\Delta\Gamma_s t/2) + b_k \sinh(\Delta\Gamma_s t/2)\right), \tag{4.58}$$

where Γ_s is B_s^0 lifetime; $\sum_{k=1}^{10}$ is the summation over the polarization amplitudes terms; ω_k are the angular acceptance weights from Eq. 4.8; N_k, a_k and b_k are the coefficients from Tab. 2.1; $\Delta\Gamma_s$ is the B_s^0 decay width difference. It is connected to the observed decay time distribution as,

$$\frac{d\Gamma(t)_{obs}}{dt} \propto \frac{d\Gamma(t)_{exp}}{dt} \times \varepsilon(t). \tag{4.59}$$

Knowing these two distributions one can extract the $\varepsilon(t)$. However, the values of ω_k are not known and have to be somehow estimated. So instead, the Eq. 4.58 is simplified by using a simulated sample with $\Delta\Gamma_s$ set to 0 ps^{-1}. For this simulated sample, the expected decay time distribution is a single exponential distribution multiplied by a constant factor.

To account for differences between the simulated and real data samples, a complex reweighting scheme is done. Its schematic is shown in Fig. 4.34,

where the weights definitions can be found in Sec. 4.5.3. Firstly, the simulated samples are reweighted according to the magnet polarity of the sWeighted data samples with *polarity* weights. Secondly, the *PDF* weights are applied to account for the absence of S-wave in the simulation sample and to correct the difference between generated and observed[7] PDF parameters. Thirdly, the *kinematics* is corrected in the simulated samples: mother particle momentum distribution and $m(K^+K^-)$ or $m(K^+\pi^-)$ depending on the decay. The B^0 momenta in the real $B^0 \to J/\psi K^{*0}$ sample is reweighted with respect to B_s^0 one in the $B_s^0 \to J/\psi K^+K^-$ real data sample.

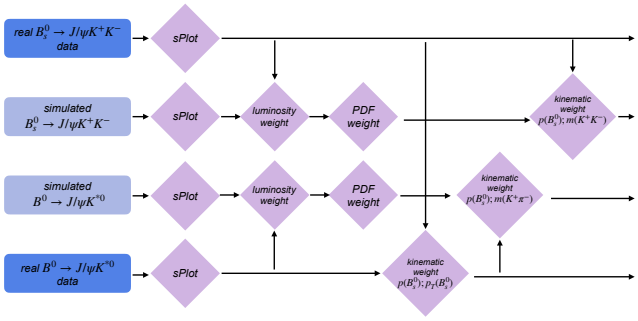


Figure 4.34: Reweighting scheme for the decay time acceptance studies. The weights are defined in Sec. 4.5.3. Adapted from [67].

Decay time acceptance modelling

A function to model decay time acceptance shape has to be a positively defined function. A cubic spline is a smooth function, which is positive if its coefficients are. In comparison to the normal polynomial interpolation, the spline is preferable as it is built from low-order polynomials. This makes a spline resistant to boundary oscillations, which can be seen when interpolating with normal high-order polynomials [88]. Therefore, a spline is chosen to model the decay time resolution. Moreover, the integral of a product of a cubic spline with decay time distribution has an analytical solution [89]. A spline function $S(t)$ is a smooth piecewise polynomial function. Any spline can be expressed as a linear combination of *basic*

[7] From previous LHCb measurements: [48], [85].

splines of order m, also known as *b-splines*,

$$S(t) = \sum_i \beta_i s_{i,m}(t) \tag{4.60}$$

where β_i are the spline coefficients. The basic splines $s_{i,m}(t)$ are piecewise polynomial functions defined on consecutive intervals, known as *knot* vector $(t_0, t_1, ..., t_n)$, as,

$$s_{i,m}(t) = \begin{cases} P_{0,m}(t), & \text{if } t_i \leq t < t_{i+1}, \\ ... & \\ P_{m,m}(t), & \text{if } t_{i+m} \leq t < t_{i+m+1}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.61}$$

where i is the knot index and $P_{j,m}(t)$ are polynomials that satisfy the continuity requirements:

$$\begin{aligned} P_{j,m}(t_i) &= P_{j+1,m}(t_i) \\ \frac{dP_{j,m}(t_i)}{dt} &= \frac{dP_{j+1,m}(t_i)}{dt} \\ &... \\ \frac{d^{m-2}P_{j,m}(t_i)}{dt^{m-2}} &= \frac{d^{m-2}P_{j+1,m}(t_i)}{dt^{m-2}}. \end{aligned} \tag{4.62}$$

The unity condition defines the overall b-spline scaling factor:

$$\sum_i^{n-(m+1)} s_{i,m}(t) = 1 \text{ for } t_m < t < t_{n-m}. \tag{4.63}$$

The exact b-spline of order m can be computed using a recursive Cox–de Boor formula [90],

$$s_{i,0}(t) \equiv \begin{cases} 1, & \text{if } t_i \leq t < t_{i+1}, \\ 0, & \text{otherwise} \end{cases}, \tag{4.64}$$

$$s_{i,m}(t) \equiv \frac{t - t_i}{t_{i+m} - t_i} s_{i,m-1}(t) + \frac{t_{i+m+1} - t}{t_{i+m+1} - t_{i+1}} s_{i+1,m-1}(t). \tag{4.65}$$

Figure 4.35 shows an example of $S(t)$ cubic spline function. Here, the cubic b-splines are used, so $m = 3$. The knots are chosen to be $[0.3; 0.91; 1.96; 9.0]$ ps. Between the last knot at $t = 9$ ps and the upper edge of decay time distribution (15 ps), the acceptance is extrapolated with a linear spline $\tilde{s}$. This is done to improve the numerical stability of the fit in the subcategories for the $t \in [9.0, 15.0)$. The continuity is ensured by these two conditions: $s_{2,3}(t_3 = 9 \text{ ps}) = \tilde{s}(t_3 = 9 \text{ ps})$ and $ds_{2,3}(t_3 = 9 \text{ ps})/dt = d\tilde{s}(t_3 = 9 \text{ ps})/dt$.

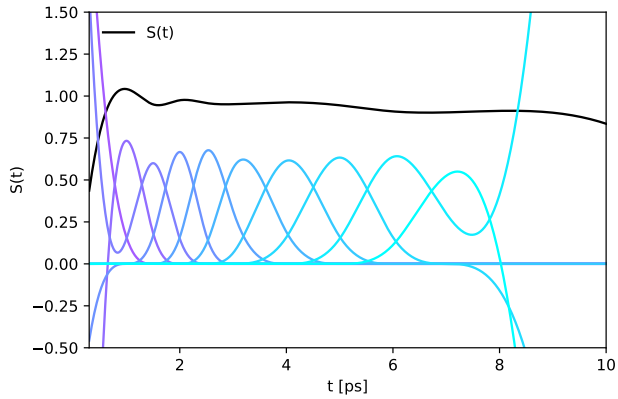


Figure 4.35: Example of a cubic spline function (Eq. 4.60). The basic splines are shown in colour.

Decay time acceptance procedure

The decay time acceptance procedure is split into two parts. The first part is reweighting of the necessary samples: simulated $B_s^0 \to J/\psi K^+K^-$ with $\Delta\Gamma_s = 0\,\mathrm{ps}^{-1}$, simulated $B^0 \to J/\psi K^{*0}$, and real $B^0 \to J/\psi K^{*0}$ sample. The simulated samples are reweighted with `sw` $\times$ `lumiWeight` $\times$ `PDFWeight` $\times$ `kinWeight`. The $B^0 \to J/\psi K^{*0}$ real data sample is reweighted with `sw` $\times$ `kinWeight`. The second part is a simultaneous fit of the three samples mentioned above. Table 4.14 summarizes the decay time acceptance models used in the fit. The simultaneous fit model is based on three key points:

	decay time acceptance model
simulated $B_s^0 \to J/\psi K^+K^-$	$S^{\mathrm{MC}}_{B_s^0 \to J/\psi K^+K^-}(t)$
simulated $B^0 \to J/\psi K^{*0}$	$S^{\mathrm{MC}}_{B^0/B_s^0}(t) \times S^{\mathrm{MC}}_{B_s^0 \to J/\psi K^+K^-}(t)$
real $B^0 \to J/\psi K^{*0}$	$S^{\mathrm{MC}}_{B^0/B_s^0}(t) \times S^{\mathrm{data}}_{B_s^0 \to J/\psi K^+K^-}(t)$

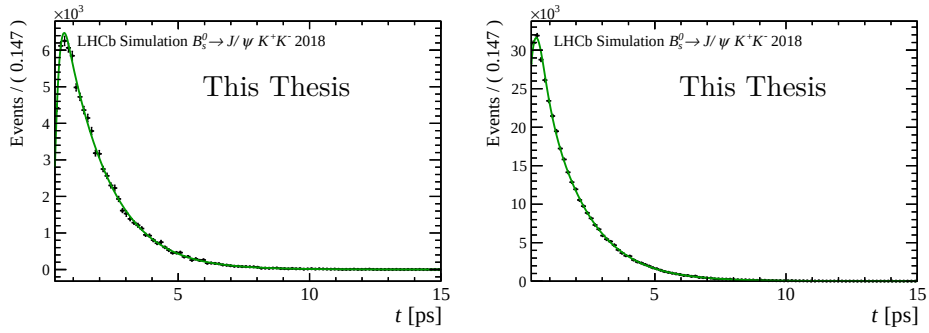
Table 4.14: Summary of decay time acceptance models. $S(t)$ is a cubic spline function. $S^{\mathrm{MC}}_{B^0/B_s^0}(t)$ is a spline of the simulated $B^0 \to J/\psi K^{*0}$ to simulated $B_s^0 \to J/\psi K^+K^-$ decay time distribution ratio.

1. The simulated $B_s^0 \to J/\psi K^+K^-$ ($\Delta\Gamma_s = 0\,\mathrm{ps}^{-1}$) decay time distribu-

tion is,

$$\frac{d\Gamma}{dt} = \frac{1}{N}\left[e^{-\Gamma_s t} \otimes G(t|\mu,\sigma)\right] \times \varepsilon^{\text{MC}}_{B^0_s \to J/\psi K^+K^-}(t) \tag{4.66}$$

where N is normalisation factor; Γ_s is the decay width, which is fixed to the generated value; $G(t|\mu,\sigma)$ is a Gaussian decay time resolution with mean set to zero and width set to decay time resolution estimated with simulated $B^0_s \to J/\psi K^+K^-$ sample; $\varepsilon^{\text{MC}}_{B^0_s \to J/\psi K^+K^-}(t)$ is the decay time acceptance of the simulated $B^0_s \to J/\psi K^+K^-$ sample, see Tab. 4.14. Figure 4.36 shows the simulated $B^0_s \to J/\psi K^+K^-$ decay time distribution with the model from Eq. 4.66.



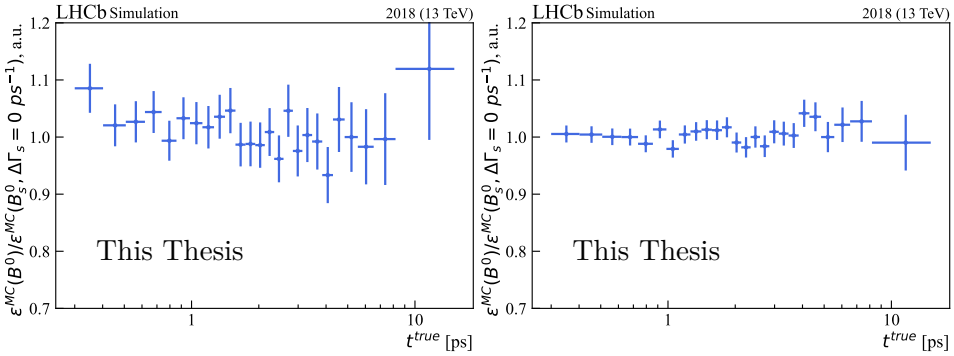
(a) biased (b) unbiased

Figure 4.36: Decay time distribution for the $B^0_s \to J/\psi K^+K^-$ simulated 2018 sample. Only candidates with $t \in [0.3, 15]$ ps are selected, see Sec. 4.2.

2. The simulated $B^0 \to J/\psi K^{*0}$ decay time distribution is described using the same model as in Eq. 4.66 with the decay width set to Γ_d and decay time resolution estimated the simulated $B^0 \to J/\psi K^{*0}$ sample, but with a decay time acceptance described with a product of two splines (see Tab. 4.14),

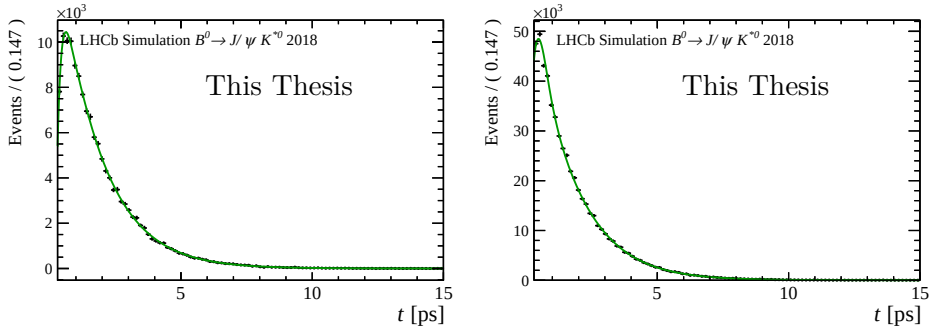
$$\varepsilon^{\text{MC}}_{B^0 \to J/\psi K^{*0}}(t) = S^{\text{MC}}_{B^0/B^0_s}(t) \times S^{\text{MC}}_{B^0_s \to J/\psi K^+K^-}(t) \tag{4.67}$$

where $S^{\text{MC}}_{B^0/B^0_s}(t)$ is a spline describing the ratio of the $B^0 \to J/\psi K^{*0}$ over $B^0_s \to J/\psi K^+K^-$ distributions. The ratio of the $B^0 \to J/\psi K^{*0}$ over $B^0_s \to J/\psi K^+K^-$ distributions is shown in Fig. 4.37. Figure 4.38 shows the simulated $B^0 \to J/\psi K^{*0}$ decay time distribution fit.



(a) biased (b) unbiased

Figure 4.37: Ratio of decay time acceptances in simulated $B^0 \to J/\psi K^{*0}$ and $B^0_s \to J/\psi K^+K^-$. The normalisation is arbitrary and is chosen to have $\varepsilon^{\text{MC}}_{B^0 \to J/\psi K^{*0}} / \varepsilon^{\text{MC}}_{B^0_s \to J/\psi K^+K^-}(5\,\text{ps}) = 1$. The line shows the ratio spline $S_{B^0/B^0_s}(t)$, see Eq. 4.67. Taken from [67].



(a) biased (b) unbiased

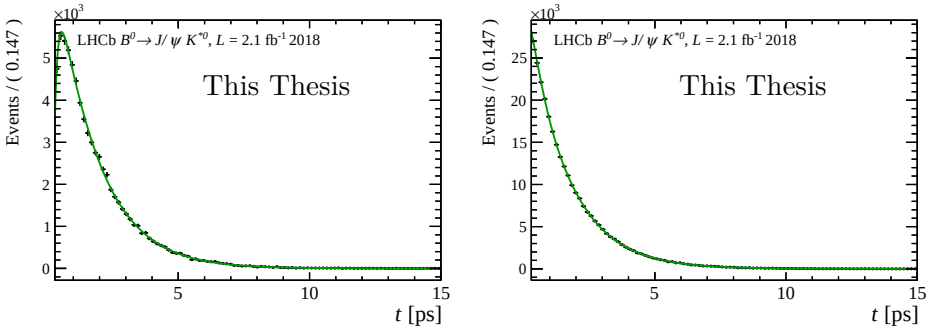
Figure 4.38: Decay time distribution for the $B^0 \to J/\psi K^{*0}$ simulated 2018 sample. Only candidates with $t \in [0.3, 15]$ ps are selected, see Sec. 4.2.

3. The model for the real $B^0 \to J/\psi K^{*0}$ sample is adjusted with the decay width set to the world average, and decay time resolution is taken from [91]. The decay time bias, described in Sec. 4.5.2, affects the $B^0 \to J/\psi K^{*0}$ sample as well. Therefore, the decay time resolution mean is set to the corresponding decay time biases. The decay time acceptance is defined as the product of the ratio spline $S_{B^0/B^0_s}(t)$ and

$B_s^0 \to J/\psi K^+K^-$ real data spline $S^{\text{data}}_{B_s^0 \to J/\psi K^+K^-}(t)$, see Tab. 4.14,

$$\varepsilon^{\text{data}}_{B^0 \to J/\psi K^{*0}}(t) = S^{\text{MC}}_{B^0/B_s^0}(t) \times S^{\text{data}}_{B_s^0 \to J/\psi K^+K^-}(t). \tag{4.68}$$

where $S^{\text{data}}_{B_s^0 \to J/\psi K^+K^-}(t)$ is a spline describing the decay time acceptance of the real data $B_s^0 \to J/\psi K^+K^-$ sample. Figure 4.39 shows the real data $B^0 \to J/\psi K^{*0}$ decay time distribution fit.



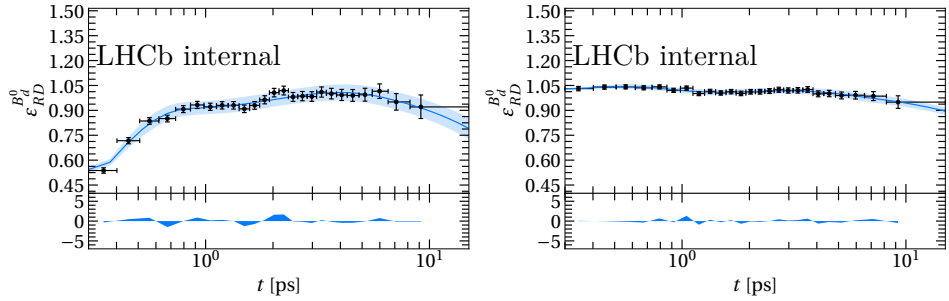
(a) biased (b) unbiased

Figure 4.39: Decay time distribution for the real $B^0 \to J/\psi K^{*0}$ 2018 sample. Only candidates with $t \in [0.3, 15]$ ps are selected, see Sec. 4.2.

As a result, when fitting the three samples simultaneously, the parameters of the $B_s^0 \to J/\psi K^+K^-$ decay time acceptance spline for real data are estimated. Figure 4.40 shows the decay time acceptance for $B^0 \to J/\psi K^{*0}$ in 2018, which is used to get the $B_s^0 \to J/\psi K^+K^-$ spline coefficients. The $B_s^0 \to J/\psi K^+K^-$ spline coefficients for real data can be found in Tab. 4.15.

4.5.5 Angular acceptance

As stated in Chapter 2, it is necessary to disentangle the different polarization amplitudes to measure ϕ_s. The contributions of these polarization amplitudes can be distinguished by their distinct angular dependence. The angular acceptance distorts the distributions of the measured helicity angles and, if not taken into account, would lead to a bias in the estimated polarization amplitudes, their phases, and, by extension, on ϕ_s. A combination of effects causes a non-uniform angular acceptance: the LHCb geometrical acceptance, detector hit resolution and efficiency, selection criteria, kinematics, and topology of the candidate.



(a) biased (b) unbiased

Figure 4.40: Decay time acceptance distributions for the $B^0 \to J/\psi K^{*0}$ data. Shown for both biased and unbiased trigger categories. The normalisation is arbitrary and is chosen to have $\varepsilon_{\mathrm{RD}}^{B^0 \to J/\psi K^{*0}}(5\,\mathrm{ps}) = 1$. Taken from [67].

	2015	2016	2017	2018
		unbiased		
β_0	1	1	1	1
β_1	0.99753 ± 0.04948	1.00096 ± 0.02594	1.03966 ± 0.02547	1.04094 ± 0.02469
β_2	0.95842 ± 0.03203	0.97213 ± 0.01693	0.97379 ± 0.01619	0.97953 ± 0.01569
β_3	1.04183 ± 0.06466	1.00713 ± 0.03373	1.03768 ± 0.03271	1.03234 ± 0.03113
β_4	0.98509 ± 0.07492	0.96068 ± 0.03966	0.95469 ± 0.03892	0.93487 ± 0.03552
β_5	0.88981 ± 0.05786	0.96478 ± 0.03227	0.96281 ± 0.03114	0.95973 ± 0.02939
		biased		
β_0	1	1	1	1
β_1	1.438 ± 0.1886	1.9726 ± 0.1179	1.7615 ± 0.1209	2.006 ± 0.2323
β_2	1.535 ± 0.1371	2.15632 ± 0.09378	1.96115 ± 0.0993	2.0365 ± 0.1833
β_3	1.5962 ± 0.2098	2.4164 ± 0.1682	2.1336 ± 0.1545	2.3711 ± 0.2505
β_4	1.7644 ± 0.2266	2.2361 ± 0.1721	1.9535 ± 0.152	2.179 ± 0.2349
β_5	1.2695 ± 0.1567	2.2421 ± 0.1508	1.7092 ± 0.1218	2.0033 ± 0.2088

Table 4.15: Decay time acceptance spline coefficients for the real $B_s^0 \to J/\psi K^+ K^-$ data sample. The first coefficient β_0 is chosen to be one.

The simulated samples describe the angular acceptance well. Therefore, contrary to the decay time acceptance, the angular acceptance is estimated from simulated samples. One could use the $B^0 \to J/\psi K^{*0}$ sample to build a data-driven procedure. However, because of a lighter pion track in the final state, it is a non-trivial task to ensure that the boost acceptance is representative of $B_s^0 \to J/\psi K^+K^-$.

It is assumed that the decay time acceptance and angular acceptance factorize. As there are distinctly different selection criteria in the two trigger categories, the angular acceptance is estimated separately for the trigger categories to account for possible differences in angular acceptance.

The binned angular acceptance ε is defined as the ratio of the observed number of candidates to the expected number of candidates in the i^{th} Ω bin,

$$\varepsilon_i = \frac{N_i^{obs}}{N_i^{exp}} \tag{4.69}$$

where N_i^{obs} is the number of candidates in i^{th} Ω bin and N_i^{exp} is expected number of candidates in this bin.

Angular acceptance estimation

As discussed in Sec 4.1, the angular acceptance only matters for the normalization of time-dependent angular PDF. In Eq. 4.14, angular acceptance enters as angular normalization weights w_k, which are defined in Eq. 4.8. Approximating the integral in Eq. 4.8 numerically, the weights can be computed as,

$$w_k = \frac{1}{N_{gen}} \sum_{i \in \text{accepted}} \frac{f_k(\Omega_i^{\text{reco}})}{\text{pdf}(\Omega_{\text{i}}^{\text{gen}} | \text{t}_{\text{i}}^{\text{gen}}, \theta^{\text{gen}})} \tag{4.70}$$

where N_{gen} is the number of generated events; i is accepted candidate index; f_k is angular distributions; Ω^{reco} are reconstructed helicity angles; t^{gen}, Ω^{gen} are generates decay time and true helicity angles; θ^{gen} are generated PDF parameters. There are 10 weights in total - one for each time-dependent angular PDF term so that $k \in 00, \|\|, \perp\perp, \|\perp, 0\,\|, 0\,\perp, SS, S\,\|, S\,\perp, S0$. With the first weight w_{00} set to 1, N_{gen} drops out from Eq. 4.70. The angular weights are discussed in detail in [92].

The angular acceptance determination mainly suffers from statistical uncertainty on the derived weights. To increase the sample size used, the angular acceptance weights are computed from both the simulated $B_s^0 \to J/\psi K^+K^-$ sample and the simulated $B_s^0 \to J/\psi K^+K^-$ $\Delta\Gamma_s = 0\,\text{ps}^{-1}$ sample. The obtained angular weights are then averaged over both samples.

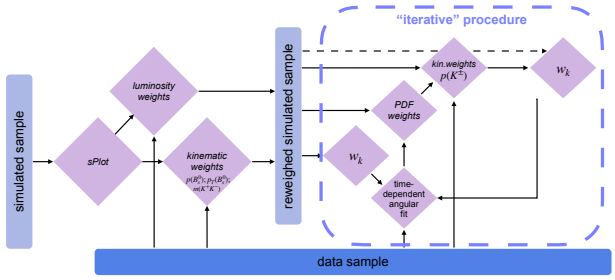


Figure 4.41: Schematics of the iterative angular acceptance procedure. Adapted from [67]

Angular acceptance procedure

The angular acceptance procedure is depicted in Fig. 4.41. If not mentioned otherwise, the weights are defined as in Sec. 4.5.3. The angular acceptance procedure is:

1. Simulated samples are reweighted with sWeights, polarity weights and kinematic weights, where kinematic weights correct the $p(B_s^0)$, $p_T(B_s^0)$ and $m(K^+K^-)$ distributions.

2. The first set of angular weights w_k^0 is computed on the reweighted simulated samples.

3. Next, the *iterative* procedure starts. For step i, where $i \in [1, n]$, angular weights w_k^{i-1} are used to compute the PDF weights. The PDF weights w_{pdf} are defined as,

$$w_{\text{pdf}} = \frac{\text{pdf}(\text{t}_{\text{gen}}, \Omega_{\text{gen}}|\theta^{\text{obs}})}{\int \text{pdf}(\text{t}_{\text{gen}}, \Omega|\theta^{\text{obs}})\text{d}\Omega} \Big/ \frac{\text{pdf}(\text{t}_{\text{gen}}, \Omega_{\text{gen}}|\theta^{\text{gen}})}{\int \text{pdf}(\text{t}_{\text{gen}}, \Omega|\theta^{\text{gen}})\text{d}\Omega} \tag{4.71}$$

4. The kinematic weights are computed to correct the $p(K^{\pm})$ distributions on the PDF weighted simulated sample.

5. A new set of angular weights w_k^i is derived. These weights are used as inputs to step 3.

The angular weights are found using angular weights as input, which means that one solves the equation in the form $F(x) = x$. The equation $F(x) = x$ is solved by defining the $f(x) \equiv F(x) - x$ and finding its root $\hat{x}$,

such that $f(\hat{x}) = 0$. This can be done with Newton's method [93, 94] and Newton-like methods, where

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{4.72}$$

where n is the iteration index and $f'(x_n)$ is the derivative of f. This is repeated until the x_n satisfies $|x_{n+1} - x_n| < \varepsilon$, where ε is the chosen tolerance level. Then the solution $\hat{x}$ is simply x_n. The disadvantage of Newton's method is that it requires the knowledge of the derivative of f. Steffensen's method is a Newton-like method that does not require the knowledge of the derivative of f [95,96]. Moreover, a common implementation of the Steffensen method includes a "built-in" convergence acceleration with Aitken's Δ^2 method [97]. This is extremely useful for problems where, instead of converging to a single solution, oscillate around the $\hat{x}$. Aitken's Δ^2 method increases the robustness of the algorithm against such oscillations. The oscillations around $\hat{x}$ were spotted when finding the root for the angular weights equation with the old algorithm in [49] and motivated the switch to Steffensen's method with Aitken's Δ^2 acceleration. It can be summarized in a few steps:

1. Pick the initial guess x_0.

2. Compute x_1 and x_2 using Steffensen method, such that:

$$x_{n+1} = x_n - \frac{f(x_n)}{g(x_n)}, \tag{4.73}$$

   where $g(x) = \frac{f(x+f(x))}{f(x)} - 1$ is the slope function. If the $|x_{n+1} - x_n| < \varepsilon$, where ε is desired tolerance, then iterations stop.

3. Compute x_3 using the Aitken's Δ^2 acceleration method as,

$$x_3 = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0}. \tag{4.74}$$

4. Set x_0 to x_3 and repeat.

The iterations are repeated until the angular weights satisfy the following conditions:

$$|w_k^n - w_k^{n+3}| < \varepsilon \tag{4.75}$$

where w_k^n are angular weights at iteration n and ε is the desired tolerance. Details on the iterative procedure can be found in [67].

The values of the angular weights for all years are given in Tab. 4.16 for completeness. Weights close to one correspond to $|A_0|^2$, $|A_\perp|^2$, $|A_\parallel|^2$ and $|A_S|^2$ terms of time-dependent angular PDF. The weights close to 0 are the weights in front of the interference terms.

	2015	2016	2017	2018
		unbiased		
w_{00}	1.0	1.0	1.0	1.0
$w_{\parallel\parallel}$	1.0319 ± 0.0023	1.0332 ± 0.0017	1.0352 ± 0.0020	1.0288 ± 0.0020
$w_{\perp\perp}$	1.0334 ± 0.0023	1.0338 ± 0.0016	1.0335 ± 0.0019	1.0291 ± 0.0019
$w_{\parallel\perp}$	0.0026 ± 0.0018	0.0005 ± 0.0013	0.0013 ± 0.0016	-0.0008 ± 0.0016
$w_{0\parallel}$	-0.0021 ± 0.0011	0.0002 ± 0.0008	0.0019 ± 0.0010	0.0006 ± 0.0009
$w_{0\perp}$	-0.0007 ± 0.0011	0.0006 ± 0.0008	-0.00004 ± 0.0010	0.0022 ± 0.0009
w_{SS}	1.0111 ± 0.0015	1.0123 ± 0.0011	1.0099 ± 0.0013	1.0083 ± 0.0013
$w_{S\parallel}$	-0.0022 ± 0.0014	0.0008 ± 0.0010	0.0009 ± 0.0012	0.0016 ± 0.0012
$w_{S\perp}$	-0.0002 ± 0.0014	0.0013 ± 0.0010	0.0004 ± 0.0013	0.0009 ± 0.0012
w_{S0}	0.0002 ± 0.0029	-0.0022 ± 0.0022	-0.0030 ± 0.0025	0.0032 ± 0.0025
		biased		
w_{00}	1.0	1.0	1.0	1.0
$w_{\parallel\parallel}$	1.0318 ± 0.0039	1.0269 ± 0.0031	1.0226 ± 0.0039	1.0267 ± 0.0039
$w_{\perp\perp}$	1.0286 ± 0.0038	1.0245 ± 0.0031	1.0234 ± 0.0039	1.0207 ± 0.0038
$w_{\parallel\perp}$	0.0004 ± 0.0031	0.0029 ± 0.0024	-0.0009 ± 0.0035	0.0003 ± 0.0031
$w_{0\parallel}$	0.0015 ± 0.0018	0.0020 ± 0.0015	0.0020 ± 0.0019	0.0024 ± 0.0018
$w_{0\perp}$	0.0012 ± 0.0018	0.0007 ± 0.0015	0.0015 ± 0.0019	-0.0017 ± 0.0019
w_{SS}	1.0197 ± 0.0027	1.0170 ± 0.0021	1.0143 ± 0.0028	1.0125 ± 0.0027
$w_{S\parallel}$	-0.0059 ± 0.0024	-0.0016 ± 0.0019	0.0005 ± 0.0025	-0.0008 ± 0.0024
$w_{S\perp}$	0.0012 ± 0.0024	0.0012 ± 0.0020	0.0022 ± 0.0025	-0.0020 ± 0.0024
w_{S0}	-0.0312 ± 0.0049	0.0016 ± 0.0041	-0.0048 ± 0.0052	-0.0055 ± 0.0050

Table 4.16: Angular weights for $B_s^0 \to J/\psi K^+K^-$ in two trigger categories and four data-taking years. Taken from [67].

4.6 Cross-checks and systematics

The robustness of the final result is as important, if not more, as the precision of it. To guarantee robustness, multiple cross-checks are done for both the individual analysis components and the time-dependent angular fit. For robustness, the measurement must be compatible between different subsamples of the data or when an equivalent procedure is used. If an unexpected result is spotted, it can be used to fix possible mistakes in different parts of the analysis. Therefore, cross-checks play an important role in developing the analysis strategy.

Firstly, there are some known assumptions that the analysis strategy is based on, like that $m(J/\psi K^+K^-)$ and the observables of interest are uncorrelated. Secondly, control samples, such as the inclusive $J/\psi K^+K^-$ control sample, have finite sample sizes. Hence nuisance parameters that are obtained from those samples and which are required for the analysis are subject to both statistical uncertainties and possible biases. The impact of

these on both the precision and possible bias of the estimated parameters of interest must be quantified, and this may result in the assignment of systematic uncertainty.

The line between a cross-check and systematic uncertainty is often blurry [98]. A common rule of thumb is to define a tolerance level which will define whether the check has passed or not. The tolerance level is defined by the $\mathcal{O}(1\%)$ of statistical uncertainty on the parameters of interest. This definition, although not conceptually satisfactory, is practical. However, it should always be used with caution, and if any unexpected result is seen, it should be investigated further, and the origin should be understood sufficiently to gauge its possible effect on the final result.

The first part of the section describes cross-checks for all the analysis components. The second part describes systematic uncertainties. The parts related to the author's work are described in greater detail.

4.6.1 Cross-checks

Selection : $\Lambda_b^0 \to J/\psi p K^-$ weights

As described in Sec. 4.2.4, the left-over $\Lambda_b^0 \to J/\psi p K^-$ background candidates are removed before the statistical background subtraction with the *sPlot* method. This is done by injecting simulated $\Lambda_b^0 \to J/\psi p K^-$ candidates into the $B_s^0 \to J/\psi K^+ K^-$ sample with negative weights. The sum of the weights is equal to the number of expected $\Lambda_b^0 \to J/\psi p K^-$ candidates in the $B_s^0 \to J/\psi K^+ K^-$ sample. The number of $\Lambda_b^0 \to J/\psi p K^-$ candidates is varied within the 1σ of their uncertainty, and the analysis is repeated. No deviations exceeding the tolerance level are observed.

4.6.2 Decay time resolution cross-checks

One of the largest changes with respect to the 2015-2016 measurement [49] is the procedure to determine the decay time resolution. Therefore, multiple cross-checks are done to verify the robustness of this novel procedure.

The new procedure differs qualitatively from the previous one in [49] by only using the negative decay time distribution of the inclusive $J/\psi K^+ K^-$ control sample and effectively including a contribution of the non-Gaussian tails, see Sec. 4.5.1. Because the non-Gaussian tails are dominated by wrong primary vertex associations, the subsample with only one primary vertex is expected to have a significantly smaller non-Gaussian contribution. Therefore, the two most important checks are the robustness of results when

dividing the sample according to the number of primary vertices and when changing the range of negative decay times utilized.

Compatibility with the 2015-2016 method

In the 2015-2016 analysis, the decay time resolution was estimated by fitting the entire decay time distribution of the full inclusive $J/\psi K^+K^-$ sample, see [49]. The decay time distribution is modeled with Eq. 4.48, and the effective width is computed from a double-Gaussian decay time resolution model. This method, however, did not include decay time resolution effects from the non-Gaussian tails and, therefore, is dropped in this iteration of the analysis. Nevertheless, the compatibility is checked between the new method and the full decay time distribution fit from [49] for all the data taking years. The final parameters estimated using the new method and the 2015-2016 method are found to be compatible within their quoted statistical uncertainty. The new method is chosen as the baseline as it accounts for the non-Gaussian effects.

Compatibility between simulated and real inclusive $J/\psi K^+K^-$ samples

Figure 4.27 shows a comparison between the distributions of the inclusive $J/\psi K^+K^-$ samples taken from simulation and real data. Some deviations are observed. However, from looking at Fig. 4.27, it is not clear if this discrepancy is due to the differences between detector resolutions in the simulated sample and real data or something else. The decay time resolution of the prompt candidates from the simulated inclusive $J/\psi K^+K^-$ sample is found to be compatible with real sample results within uncertainties.

In future iterations, the simulated samples can be made more representative of real data with kinematical reweighting; for example, see kinematic reweighting in Sec. 4.2. The reweighting should be based on at least $p_z(B_s^0)$ and $p_T(B_s^0)$. The simulated samples should be generated with real data hit resolution so that the decay time uncertainty σ_t in the simulated sample is representative of real data.

Translation from prompt $J/\psi K^+K^-$ candidates to $B_s^0 \to J/\psi K^+K^-$ candidates

The decay time resolution of the prompt candidates is translated to the $B_s^0 \to J/\psi K^+K^-$ sample via the ratio of dilutions obtained from simulation,

see Eq. 4.50. As an alternative, two other possibilities are studied:

$$D^{\text{data}}_{B^0_s\to J/\psi K^+K^-} = D^{\text{data}}_{prompt\ J/\psi K^+K^-} + D^{\text{MC}}_{B^0_s\to J/\psi K^+K^-} - D^{\text{MC}}_{prompt\ J/\psi K^+K^-} \tag{4.76}$$

and

$$\sigma^{\text{data}}_{B^0_s\to J/\psi K^+K^-} = \sigma^{\text{data}}_{prompt\ J/\psi K^+K^-} + \sigma^{\text{MC}}_{B^0_s\to J/\psi K^+K^-} - \sigma^{\text{MC}}_{prompt\ J/\psi K^+K^-} \tag{4.77}$$

where $D^{\text{data}}_{B^0_s\to J/\psi K^+K^-}$ is the effective dilution determined from the real $B^0_s \to J/\psi K^+K^-$ candidates, and $\sigma^{\text{data}}_{B^0_s\to J/\psi K^+K^-}$ is their effective decay time resolution; $D^{\text{data}}_{prompt\ J/\psi K^+K^-}$ and $\sigma^{\text{data}}_{prompt\ J/\psi K^+K^-}$ are the effective dilutions of prompt $J/\psi K^+K^-$ candidates, and their effective decay time resolution; $D^{\text{MC}}_{B^0_s\to J/\psi K^+K^-}$ and $\sigma^{\text{MC}}_{B^0_s\to J/\psi K^+K^-}$ are the effective dilution of simulated $B^0_s \to J/\psi K^+K^-$ candidates, and their effective decay time resolution; $D^{\text{MC}}_{prompt\ J/\psi K^+K^-}$ and $\sigma^{\text{MC}}_{prompt\ J/\psi K^+K^-}$ are the effective dilution and decay time resolution of the simulated prompt $J/\psi K^+K^-$ candidates.

The results of both procedures are found to be equivalent to the one in Eq. 4.50, as no deviation for the parameters of interest is observed.

Robustness with respect to decay time range

The decay time resolution is estimated from the negative decay time tail of the inclusive $J/\psi K^+K^-$ sample. Four different decay time ranges are compared between each other: $t \in [-10, \mu]$ ps, $t \in [-7, \mu]$ ps, $t \in [-4, \mu]$ ps and $t \in [-1, \mu]$ ps, where μ is a decay time bias for the prompt candidates seen in the inclusive real $J/\psi K^+K^-$ sample. For the simulated samples, $\mu = 0$ ps. Figure 4.42 shows the mentioned decay ranges for comparison. The non-Gaussian tails are relatively less populated than the core of the decay time distribution, which results in all of the ranges giving compatible results. The $t \in [-4, \mu]$ ps is used in the baseline result for the analysis, which is in agreement with [49].

Robustness with respect to the number of primary vertices in the sample

About 40% of the $B^0_s \to J/\psi K^+K^-$ candidates are produced in events with one primary vertex. Figure 4.43 shows the decay time distribution for the 2018 inclusive $J/\psi K^+K^-$ subsamples with one primary vertex and multiple primary vertices. It is obvious that the non-Gaussian tail is more apparent in the sample with multiple primary vertices, respectively. The nature of

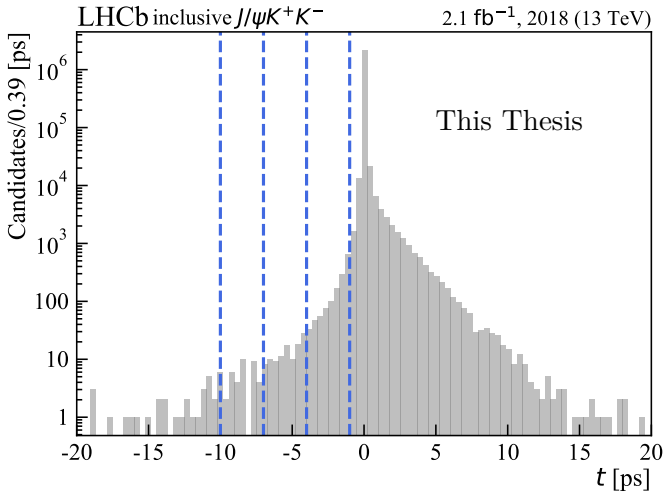


Figure 4.42: Decay time distribution of 2018 inclusive real $J/\psi K^+K^-$ sample with overlayed decay time ranges: $t \in [-10, \mu]$ ps, $t \in [-7, \mu]$ ps, $t \in [-4, \mu]$ ps, $t \in [-1, \mu]$ ps.

this tail is associated with the events for which the primary vertex is not correctly associated with the B_s^0 decay vertex. For events with only one primary vertex, this is highly unlikely and can only happen if the correct primary vertex is not reconstructed.

To cross-check the difference between the two subsamples, the decay time resolution procedure is repeated in both. The final fit was repeated for the calibration from the control sample with one primary vertex per event and for the calibration from the control sample with multiple primary vertices per event. No deviations in the values of the parameters of interest in comparison to the baseline were found.

Since 2022, the LHCb detector environment has, on average, five visible proton-proton collisions in the event. This means that the tails observed in the multiple primary vertices subsample will populate a larger fraction of the collected data. In future iterations of the analysis, a division of the sample into subsamples by different primary vertex multiplicities might be necessary to describe the decay time resolution. As an alternative, the division can be done in subsamples of low and high B_s^0 pseudorapidity η. The B_s^0 candidates with low pseudorapidity have larger angles with respect to the LHCb z-axis, which makes it easier to point to the correct primary vertex. And vice versa, the more forward B_s^0 candidates are, the more likely they are to be misassociated with the wrong primary vertex. Figure 4.44

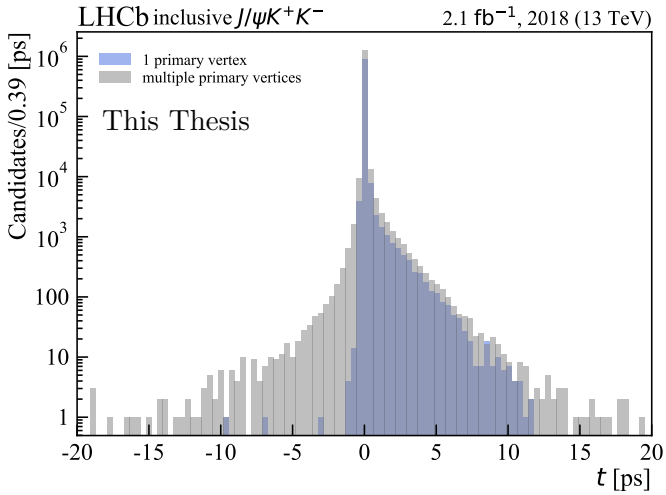


Figure 4.43: Decay time distribution of 2018 inclusive real $J/\psi K^+K^-$ sample for a subsample with on primary vertex and with multiple primary vertices.

illustrates this effect: the inclusive $J/\psi K^+K^-$ candidates with $\eta < 3.5$ have less non-Gaussian tails, that candidates with $\eta \geq 3.5$.

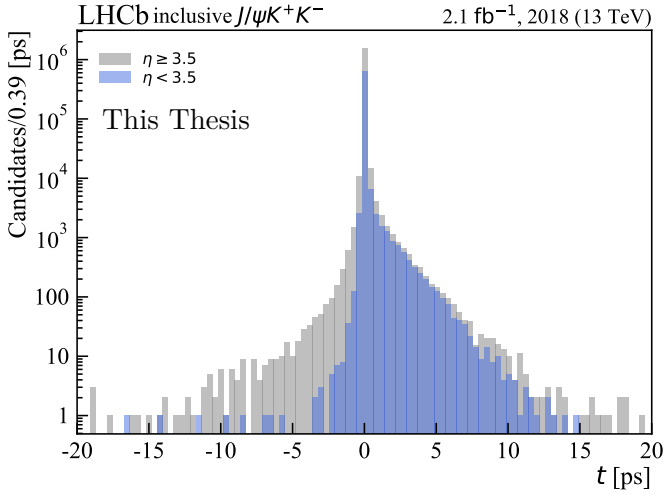


Figure 4.44: The decay time distribution of the inclusive $J/\psi K^+K^-$ candidates split between the low and high pseudorapidity.

Robustness with respect to the decay time uncertainty binning

The bins of decay time uncertainty σ_t are chosen such that the sample of simulated $B_s^0 \to J/\psi K^+K^-$ candidates populates each bin equally. An alternative binning is defined by requiring the inclusive $J/\psi K^+K^-$ candidates to populate bins equally. The effect of a binning change is negligible.

Dependence on the Δm_s value

The effective decay time resolution is defined by Eq. 4.37. It requires using an estimate of Δm_s to perform the translation from dilution to resolution. The value of Δm_s is varied within its one σ uncertainty, and the difference with the baseline is found to be negligible.

Quadratic decay time uncertainty calibration

The baseline calibration of the effective decay time uncertainty σ_{eff} is a first-order polynomial, as defined in Eq. 4.51. A second-order polynomial is tested as an alternative. It is parameterized as

$$\sigma_{\text{eff}} = p_0 + p_1(\sigma_t - \sigma_t^0) + p_2(\sigma_t - \sigma_t^0)^2, \tag{4.78}$$

where p_0, p_1, p_2 are polynomial parameters; σ_t is the measured decay time uncertainty and $\sigma_t^0 = 0.036$ ps is the average measured decay time uncertainty in the simulated $B_s^0 \to J/\psi K^+K^-$ sample.

A negligible difference with the first-order polynomial calibration is found.

Robustness of the long-lived candidate's dilution correction in inclusive $J/\psi K^+K^-$

The dilution correction for the long-lived candidates is described in Sec. 4.5.1. The dilution is computed using a representative sample. This representative sample is taken from the decay time distribution of long-lived candidates, which in turn is taken from the decay time distribution fit of the inclusive $J/\psi K^+K^-$ sample. The robustness corresponding to changes in the sample statistics and variations in the shape of the long-lived distribution is assessed. For this, the dilution is sampled multiple times with varied long-lived distribution model parameters. As well as, samples with more statistics (10^5 and 10^6 versus 10^4) are checked. The differences are negligible.

4.6.3 Decay time acceptance

The $B^0 \to J/\psi K^{*0}$ control sample is used to estimate the decay time acceptance; see Sec. 4.5.4. As an alternative, a $B^+ \to J/\psi K^+$ sample is used as a control instead of a $B^0 \to J/\psi K^{*0}$ sample in the analysis. Using the decay time acceptance estimated with the $B^+ \to J/\psi K^+$ sample, the B_s^0 meson lifetime agrees with the one estimated using $B^0 \to J/\psi K^{*0}$ as the control sample with p-value ≈ 0.15.

For the next check, $B^+ \to J/\psi K^+$ sample is used as a signal sample and $B^0 \to J/\psi K^{*0}$ as a control sample. The B^+ lifetime is measured in agreement within two σ with the current world average [20].

For the last lifetime check, the $B^0 \to J/\psi K^{*0}$ real data and simulated samples are split into two statistically independent subsamples, where one is used as a "signal" and one as a "control" sample. Additional selection criteria are applied to the "signal" sample to mimic differences between the real $B_s^0 \to J/\psi K^+K^-$ (signal) and $B^0 \to J/\psi K^{*0}$ (control) samples. The selection criteria can be found in [67]. For the "control" sample, the B^0 lifetime is set to the world average [20]. The B^0 lifetime is measured in per-year and per-trigger category subsamples. It is found to be compatible within $0.01 - 2\ \sigma$ with the input world average value.

The choice of using continuous splines, instead of the ratio from Eq. 4.57, is checked by using Eq. 4.57 as a proxy for the decay time acceptance. The effect on the final parameters of interest is found to be very small and, therefore, is ignored.

As a cross-check, the number and position of the spline knots number and position are varied. The difference with the baseline estimates of parameters of interest was found to be negligible.

Figure 4.34 shows the weighting schematics for the decay time acceptance procedure. As a cross-check, the kinematic weights in Fig. 4.34 were not used. Negligible deviations are observed.

The `L0` trigger cuts change multiple times per year to control the total data rate in LHCb. However, the simulated samples use only one representative set of cuts at `L0` trigger level. This leads to discrepancies between the samples of the selected simulated and real candidates. The fractions of candidates with specific `L0` trigger cuts are compared between $B_s^0 \to J/\psi K^+K^-$ and $B^0 \to J/\psi K^{*0}$ simulated samples. The fractions are found to be similar, and no influence on the decay time acceptance is expected. This is further verified by splitting the sample into subsamples that passed through the `L0` trigger representative cut set from simulation and the subsample that passed through the other `L0` trigger cuts. The parameters of interest estimated

from both subsamples were found to be compatible.

4.6.4 Angular acceptance

The key assumption of the angular acceptance procedure is that the angular and decay time acceptances factorize. Strictly speaking, however, they are correlated. The effect of this correlation on the parameters of interest is studied by comparing angular acceptance in bins of decay time. It was found that the angular acceptance is compatible between bins of decay time [67]. Therefore, this correlation is ignored.

The angular acceptance procedure is validated on the simulated $B_s^0 \to J/\psi\, K^+K^-$ sample. The simulated sample is split randomly into control and test samples. The control sample is used to estimate the angular acceptance. The test sample is used to measure the parameters of interest using the angular acceptance from the control sample. The parameters estimated on the simulated samples are found to be compatible with the parameters the simulation was generated with.

4.6.5 Decay time angular fit

As a first cross-check, the 2015-2016 results [49] are reproduced with this analysis baseline. Overall, 88% of events overlap between samples used in [49] and 2015-2016 samples used here. The observed differences in the final parameters are compatible within one σ with the result published in [49] and are attributed to the changes in the methodology. The most notable changes include improvements in the selection, the different decay time resolution procedure, the use of different stripping versions, and the removal of some incorrectly simulated samples.

Secondly, the robustness of the estimated parameters of interest is checked in subsamples, split by: magnet polarity; "biased" and "unbiased" trigger category; years; B_s^0 transverse momentum bins; B_s^0 pseudorapidity bins; taggers; number of primary vertices; bins of ProbNN_K. The estimated parameters in the subsamples are found to be compatible.

Thirdly, the background subtraction procedure with the *sPlot* method is known to lead to an underestimated statistical uncertainty. As a cross-check, the statistical uncertainty coverage is checked with the bootstrapping method. In the bootstrapping method, the dataset is resampled multiple times and is used to remeasure parameters of interest. The variances of the distribution of the parameters of interest obtained by this method give a good estimate of their standard deviation. More can be found in [99].

The asymptotically correct errors from [100] were also computed using the RooFit implementation of the final fit and were found to be compatible with the errors from the bootstrapping method.

Lastly, the selection criteria in `HLT1` change during data taking. However, simulated samples are only simulated with a single representative set of selection criteria. A special simulated $B_s^0 \to J/\psi K^+K^-$ sample is produced with realistic `HLT1` selection criteria, which is more representative of real data. The estimated parameters of interest are compared between a standard representative `HLT1` selection criterion simulated sample and a special multi-criteria `HLT1` simulated sample. The difference between parameters of interest is $\sim 0.001 - 0.1$ of corresponding statistical uncertainty, with the exception of $\Gamma_s - \Gamma_d$, for which it is $0.2\sigma_{stat}$.

4.7 Systematic uncertainties

Systematic uncertainties quantify a systematic bias that can be introduced by a multitude of effects: the analysis strategy, the irreducible detector effects, the uncovered backgrounds, etc. Tables 4.17-4.18 summarize the systematic uncertainties in this analysis. The largest uncertainty comes from the background subtraction procedure which assumes that the $m(J/\psi K^+K^-)$ mass and the observables of interest used in the angular decay time fit are uncorrelated while there is an observed, non-negligible correlation between them, see Fig. 4.14 and Fig. 4.15. This section describes each row in Tables 4.17-4.18 separately.

Including multiple candidates

As was mentioned before, the presence of multiple candidates can lead to uncovered biases in the estimated parameters [101]. About $1.3\% - 1.7\%$ of the collected events contain multiple $B_s^0 \to J/\psi K^+K^-$ candidates. All of these candidates share some of the final state tracks with other candidates. The majority share the J/ψ candidate but have different kaon tracks. Multiple candidates are removed from the sample by randomly picking one of them. The parameters of interest are then estimated using multiple candidate-free samples. The difference with respect to the baseline results is assigned as "multiple candidates" systematic uncertainty.

Source	$\lvert A_0\rvert^2$	$\lvert A_\perp\rvert^2$	ϕ_s [rad]	$\lvert\lambda\rvert$	$\delta_\perp - \delta_0$ [rad]	$\delta_\parallel - \delta_0$ [rad]	$\Gamma_s - \Gamma_d$ [ps^{-1}]	$\Delta\Gamma_s$ [ps^{-1}]	Δm_s [ps^{-1}]
Including multiple candidates	0.0001	-	0.0027	0.0022	0.0090	0.0041	0.0001	0.0001	0.0024
Particle identification correction	0.0006	0.0009	0.0027	0.0027	0.0131	0.0051	0.0005	0.0015	0.0046
Decay time resolution calibration	-	-	0.0003	0.0002	0.0011	0.0007	-	-	0.0005
Decay time resolution translation	-	-	0.0008	0.0003	0.0026	0.0009	-	-	0.0009
Lifetime bias from decay time resolution	-	-	-	-	-	-	-	+0.0003	-
Imperfect decay time bias correction	0.0004	−0.0005	−0.0006	−0.0005	0.0077	−0.0011	0.0003	0.0005	−0.0044
Decay time acceptance	0.0001	0.0001	-	-	-	-	0.0003	−0.0004	-
Angular acceptance: simulated sample size	0.0005	0.0014	0.0025	0.0032	0.0042	0.0044	0.0001	0.0002	0.0013
Correlation of $m(J/\psi K^+K^-)$ with observables	0.0011	0.0010	0.0042	0.0019	0.0054	0.0060	0.0012	0.0016	0.0018
sWeights: sample size	0.0004	0.0004	0.0005	0.0009	0.0062	0.0033	0.0002	0.0001	0.0011
Parameterization of signal width in $m(J/\psi K^+K^-)$ model	0.0004	0.0003	0.0003	0.0002	0.0015	0.0012	0.0002	0.0004	0.0003
C_{SP} factors	-	0.0001	0.0001	0.0003	0.0073	0.0041	-	0.0001	0.0004
Limited knowledge of magnetic field	0.0008	0.0004	0.0004	-	0.0007	0.0011	0.0001	-	0.0013
Limited knowledge of LHCb length scale	-	-	0.0005	0.0002	0.0008	-	0.0002	0.0001	0.0042
Ignoring decays via $f_2(1270)$ resonant	0.0004	0.0004	0.0002	-	0.0007	0.0013	0.0001	0.0003	0.0002
Kinematic corrections configuration	0.0024	0.0009	0.0001	0.0001	0.0098	0.0086	0.0002	0.0003	0.0031
Peaking candidates in the simulated combinatorial background	0.0007	0.0004	0.0002	0.0010	0.0018	0.0018	0.0002	-	0.00010
Ignoring angular resolution	0.0001	0.0001	0.0002	0.0001	0.0002	0.0008	-	0.0001	0.0002
Validation with simulated sample and ignored correlations	-	-	-	-	0.042	0.0496	-	-	-
Lifetime bias from $B_c^+ \to B_s^0\pi^+$ decays	0.0004	0.0005	-	0.0002	-	0.0017	0.0007	0.0003	-
Total (quadratic sum)	0.0032	0.0024	0.0064	0.0053	0.048	0.052	0.0014	0.0024	0.0090
Stat.	0.0017	0.0023	0.022	0.011	0.075	0.061	0.0013	0.0044	0.033

Table 4.17: Summary of the systematic uncertainties for polarization amplitudes and B_s^0 parameters.

Source	$\delta_S^1 - \delta_\perp$ [rad]	$\delta_S^2 - \delta_\perp$ [rad]	$\delta_S^3 - \delta_\perp$ [rad]	$\delta_S^4 - \delta_\perp$ [rad]	$\delta_S^5 - \delta_\perp$ [rad]	$\delta_S^6 - \delta_\perp$ [rad]	$\lvert A_S^1\rvert^2$	$\lvert A_S^2\rvert^2$	$\lvert A_S^3\rvert^2$	$\lvert A_S^4\rvert^2$	$\lvert A_S^5\rvert^2$	$\lvert A_S^6\rvert^2$
Including multiple candidates	0.026	0.021	0.14	0.0078	−0.0053	−0.010	−0.0090	−0.0004	−0.0003	−0.0011	−0.0041	−0.0016
Particle identification correction	0.138	0.026	0.10	0.041	0.033	0.031	0.0106	0.0026	0.0002	0.0002	0.0031	0.0081
Decay time resolution calibration	0.0008	0.0018	0.0040	0.0007	0.0004	0.0004	-	-	0.0003	-	-	-
Decay time resolution translation	0.0027	0.0001	0.0019	0.0003	0.0032	0.0020	-	0.0001	-	-	0.0004	0.0004
Imperfect decay time bias correction	0.014	−0.0016	0.004	0.005	0.017	−0.002	−0.0035	0.0007	−0.00007	−0.0002	0.0002	−0.0009
Decay time acceptance	-	-	0.0004	-	0.0001	0.0001	0.0001	-	-	-	-	0.0001
Angular acceptance : simulated sample size	0.0072	0.0057	0.0361	0.2368	0.0142	0.0109	0.0027	0.0009	0.0002	0.0005	0.0021	0.0025
Correlation of $m(J/\psi K^+K^-)$ with observables	0.0535	0.0092	0.0163	0.0070	0.0072	0.0320	0.0070	0.0007	0.0001	0.0002	0.0007	0.0038
sWeights: sample size	0.0082	0.0055	0.0082	0.0036	0.0283	0.0305	0.0108	0.0005	-	0.0001	0.0043	0.0113
Parameterization of signal width in $m(J/\psi K^+K^-)$ model	0.0079	0.0042	0.0139	0.0068	0.0096	0.0101	0.0102	0.0005	0.0001	0.0003	0.0016	0.0038
C_{SP} factors	0.0743	0.0044	0.0036	0.0012	0.004	0.0437	0.0006	0.0012	0.0002	0.0002	0.0003	0.0023
Limited knowledge of magnetic field	0.0017	0.0052	0.0009	0.0050	0.0019	0.0014	0.0016	0.0004	-	-	0.0002	0.0010
Limited knowledge of LHCb length scale	0.0137	-	-	0.0014	-	-	0.0044	-	-	-	-	-
Ignoring decays via $f_2(1270)$ resonant	0.094	0.17	0.052	0.034	0.017	0.004	0.0043	0.0002	0.0005	-	-	0.0007
Kinematic corrections configuration	0.0028	0.0012	0.0002	0.0005	0.0024	0.0026	0.0011	0.0061	0.0232	0.0312	0.0212	0.0131
Peaking candidates in the simulated combinatorial background	0.0002	0.0002	0.0001	0.0003	0.0014	0.0009	0.0130	0.0062	0.0016	0.0060	0.0142	0.0013
Ignoring angular resolution	0.0083	0.0015	0.0075	0.0022	0.0021	0.0065	0.0076	-	0.0001	0.0001	0.0002	0.0016
Validation with simulated sample and ignored correlations	-	-	-	-	0.07	-	-	-	-	-	-	-
Lifetime bias from $B_c^+ \to B_s^0\pi^+$ decays	-	−0.24	−0.28	−0.026	+0.20	−0.15	-	−0.0041	−0.0007	-	−0.0015	−0.0053
Total (quadratic sum)	0.19	0.41	0.19	0.31	0.17	0.07	0.027	0.010	0.023	0.032	0.027	0.051
Stat.	0.13	0.20	0.35	0.14	0.073	0.079	0.024	0.005	0.001	0.002	0.007	0.011

Table 4.18: Summary of the systematic uncertainties for the S-wave parameters.

Particle identification correction

The particle identification correction uncertainty has two sources. The first one is limited statistics of the calibration samples and the second one comes from the choice of particle identification distribution hyperparameters in `PIDCalib`. For each parameter of interest, the largest effect of the two is assigned as systematic uncertainty.

The influence of limited statistics of the correction samples on the final parameter values is estimated using the bootstrapping technique. New corrections are derived from the bootstrapped samples and are used to repeat the analysis.

The hyperparameters of the particle identification distributions in `PIDCalib` are varied around their optimal value. The analysis is repeated with these modified hyperparameters, and the final parameters of interest are compared to the baseline.

Decay time resolution calibration

The new decay time resolution calibration is sampled from a two-dimensional Gaussian distribution centered on the optimal values of the parameters with a width given by their uncertainties. The decay time angular fit is redone with each set of new values. The standard deviations of the resulting distributions of the parameters of interest are taken as systematic uncertainty.

Decay time resolution translation

To access the effect of the translation procedure, it is skipped altogether, and instead, the decay time resolution of prompt candidates is used directly in the decay time angular fit. The difference with respect to the baseline procedure with translated decay time resolution is taken as uncertainty.

Lifetime bias from decay time resolution

In general, the decay time resolution mainly affects the $B_s^0 - \overline{B}_s^0$ oscillation frequency Δm_s. However, a small bias in the estimated lifetimes is introduced by the non-Gaussian effects. Assuming that the decay time distribution for an untagged $B_s^0 \to J/\psi K^+K^-$ with $\Delta\Gamma_s = 0\,\text{ps}^{-1}$ is a simple exponent, a sample is generated, and the bias from the decay time resolution is measured using this sample. The lifetime estimator in the absence of the decay

time resolution bias is an average shifted by the integration limits[8], see Appendix A.3. Any deviation from the expected value of the estimator, see Eq. A.6, is assigned as a bias.

The decay time resolution proxy is defined as the negative decay time distribution of the inclusive $J/\psi K^+K^-$ sample, which is mirrored around zero. Its dilution is effectively represented as a single Gaussian. Firstly, the single Gaussian, which effectively represents the dilution of the decay time resolution proxy, is numerically convolved with the decay time distribution. No bias is found with respect to the expected lifetime. However, when the decay time resolution proxy is convolved with the decay time distribution, the bias of about 2.3 fs with respect to Eq. A.6 is observed. The main source of this bias is linked to the non-Gaussian decay time resolution tails. The exercise is repeated for the B^0 and B_H, B_L lifetimes, see Sec. 2.3. Table 4.19 shows the biases summary. Note that the biases are estimated conservatively,

	estimated bias fs
B^0	2.35
B_s^0	2.35
B_H	2.31
B_L	2.38

Table 4.19: The biases from the expected lifetime MLE in Eq. A.6.

assuming that no correction from the decay time acceptance is present, even though the bias on B^0 and B_s^0 lifetimes are of the same order.

The biases from Tab. 4.19 are translated to the effective biases on lifetimes measured in the decay time angular fit. The B_s^0 lifetime in the decay time angular fit is measured with respect to B^0 lifetime as $\Gamma_s - \Gamma_d$. The bias on Γ_d is similar in size as on Γ_s, and therefore, the total bias is just $10^{-6}\,\mathrm{ps}^{-1}$, which is negligible. For $\Delta\Gamma_s$, the bias is $3 \cdot 10^{-4}\,\mathrm{ps}^{-1}$, which is about 7% of the statistical uncertainty on $\Delta\Gamma_s$. Therefore, the $\Delta\Gamma_s$ bias is included in systematics as an asymmetric effect. Generally, it is more appropriate to include such asymmetric effects in the MLE fit rather than quoting it as an uncertainty. However, due to the absence of the full study with decay time acceptance included, it was decided to present it as a part of systematic uncertainties rather than modifying the MLE itself.

[8]In this measurement the decay time $t \in [0.3, 15)$ ps as can be seen from selections in Tab. 4.2.

Imperfect decay time bias correction

When the correcting decay time bias procedure from Sec. 4.5.2 is applied to the simulated $B_s^0 \to J/\psi K^+K^-$ samples with an artificial misalignment, it does not fully unbias the estimated values of parameters of interest so that they are equal to their generated values. The residual bias after applying the decay time bias procedure on the simulated $B_s^0 \to J/\psi K^+K^-$ sample with the average misalignment observed in data is assigned as the systematic uncertainty of the decay time bias correction. The residual bias mainly affects Δm_s and $\delta_\perp - \delta_0$, as shown in Tab. 4.17.

Decay time acceptance

Two sources of systematic uncertainty come from the simulated samples used in the decay time acceptance estimation. The first is the limited size of the samples. It is accounted for by constraining the decay time resolution spline parameters according to their statistical uncertainties.

The second uncertainty comes from differences between the $\Delta\Gamma_s = 0$ ps^{-1} $B_s^0 \to J/\psi K^+K^-$ simulated samples and decay time acceptance estimated using the $\Delta\Gamma_s \neq 0$ ps^{-1} simulated $B_s^0 \to J/\psi K^+K^-$ sample. The $\Delta\Gamma_s = 0$ ps^{-1} $B_s^0 \to J/\psi K^+K^-$ simulated sample is used to estimate the baseline decay time acceptance, as described in Sec. 4.5.4. The differences between the $\Delta\Gamma_s = 0$ ps^{-1} and $\Delta\Gamma_s \neq 0$ ps^{-1} simulated $B_s^0 \to J/\psi K^+K^-$ sample lead to a bias in the parameters of interest. This bias is estimated by using the alternative decay time acceptance in the MLE fit. The alternative decay time acceptance is evaluated on the $\Delta\Gamma_s \neq 0$ ps^{-1} $B_s^0 \to J/\psi K^+K^-$ simulated sample, reweighted to have $\Delta\Gamma_s = 0$ ps^{-1}.

Angular acceptance: simulated sample size

Alternative angular weights are sampled from a multi-Gaussian distribution with widths corresponding to the angular weights uncertainties; see Sec. 4.5.5. The analysis is repeated with these new sets of angular weights. The widths of the parameters of interest distributions are assigned as systematic uncertainties.

Correlation of $m(J/\psi K^+K^-)$ with observables

The dominant systematic uncertainty on ϕ_s comes from the background subtraction with the *sPlot* method. As was mentioned in Sec. 4.3, the core assumption for the *sPlot* method is that the discriminating variable

($m(J/\psi K^+K^-)$) is independent of the observables of interest (decay time and helicity angles). It is assumed that this holds true for this analysis. However, strictly speaking, this assumption is not true, as there is a direct correlation between the mass resolution and helicity angles (see Eq. A.3), as well as between the decay time and the mass (see Eq. 4.40). These correlations can be seen from Fig. 4.14-4.15, and ignoring them leads to a bias in the parameters of interest. This bias is estimated from deriving the sWeights in bins of observables. The analysis is then repeated on the full sample with new sWeights derived from the binned samples.

Next, the background is subtracted in the asymmetrical mass ranges: $m(J/\psi K^+K^-) \in [5200, 5440]\,\mathrm{MeV}/c^2$ and $m(J/\psi K^+K^-) \in [5300, 5550]\,\mathrm{MeV}/c^2$ and the analysis is repeated.

The final parameters of interest are compared to the baseline result. The largest deviation is assigned as a systematic uncertainty, which in the case of ϕ_s and $\Gamma_s - \Gamma_d$ comes from the correlation with the decay times.

sWeights: sample size

A part of the background subtraction procedure is a maximum likelihood fit of a $m(J/\psi K^+K^-)$ model described by Eq. 4.22. The estimated model parameters are varied within their associated statistical uncertainty and are used to recompute the sWeights. Then, the analysis is repeated with the new sWeights. The largest difference between the new values of parameters of interest and the baseline values of parameters of interest is assigned as systematic uncertainty.

Including per event mass uncertainty parametrization for the signal tail parameters in $m(J/\psi K^+K^-)$ fit

As mentioned in Sec. 4.3, the $m(J/\psi K^+K^-)$ model is a conditional PDF. The signal in this model is described by a double-sided Crystal Ball distribution or DCB, which is defined Eq. 4.20. Its width depends on the per-event mass uncertainty σ_m. The *"tail"* parameters are considered uncorrelated. This assumption between the *tail* parameters of DCB and the per-event mass uncertainty is, however, not fully justified. Figure 4.45 shows both the left-side and right-side tail parameters of DCB estimated in bins of σ_m.

It is obvious that there is a correlation. To estimate the effect of ignoring this correlation, the *tail* parameters are parametrized as a function of σ_m. For α_R, n_R, α_L the first-order polynomial $s_0 + s_1\sigma_m$ is used. For n_L the second-order polynomial $s_1 + s_2(\sigma_m - s_0)^2$ is used. The polynomial parameters of

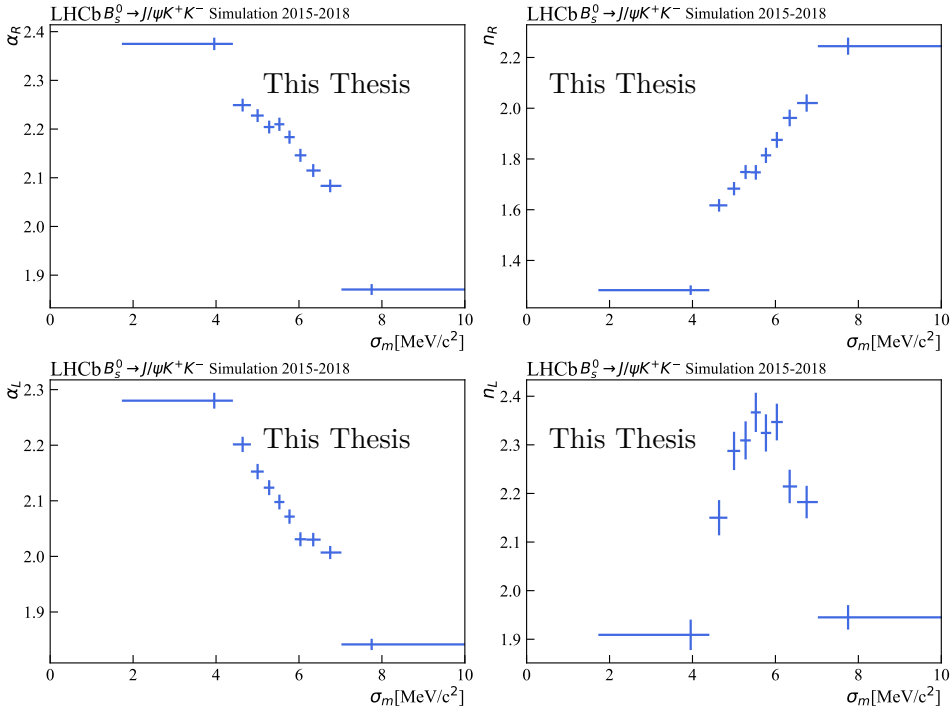


Figure 4.45: Dependence of *"tail"* parameters of signal $m(J/\psi K^+K^-)$ model on the per event mass uncertainty σ_m. Plotted for the 2015-2018 simulated $B_s^0 \to J/\psi K^+K^-$ sample.

α_R and n_R are fixed from the simulated samples. Parameters describing α_L and n_L are left floating in the real data fit. No improvement in the mass fit quality is observed. The **sWeights: sample size** systematics covers deviations in the final parameters.

Unsurprisingly, the correlations between *tail* parameters and σ_m would have been naturally included in the model if $m(J/\psi K^+K^-)$ is fitted in bins of σ_m, as described in Sec. 4.3.

Parameterization of signal width in $m(J/\psi K^+K^-)$ model

The width of the signal shape in the $m(J/\psi K^+K^-)$ model is parameterized as a function of the per event mass uncertainty σ_m, see Eq. 4.22. Alternative parameterizations are used for the width of signal shape in Eq. 4.22,

1. linear : $s_0 + s_1 \cdot \sigma_m$;

2. quadratic shifted to the average per event mass uncertainty $\langle\sigma_m\rangle$: $s_0 + s_1 \cdot (\sigma_m - \langle\sigma_m\rangle) + s_2 \cdot (\sigma_m - \langle\sigma_m\rangle)^2$;
3. cubic : $s_1 \cdot \sigma_m + s_2 \cdot \sigma_m^2 + s_3 \cdot \sigma_m^3$.

The analysis is repeated for each of the parametrizations. The largest deviation between the new final parameters of interest and the baseline ones is assigned as systematic uncertainty.

C_{SP} factors

As described in Sec. 4.1, multiple amplitudes contribute $B_s^0 \to J/\psi K^+K^-$ transition, which is described by the C_{SP} factors. The C_{SP} factors are determined with Eq. 4.12, which uses the P-wave line shape (decay via $\phi(1020)$) and S-wave line shape (decay without resonance or via the $f_0(980)$). Both line-shapes depend on the externally measured parameters [67]. The P-wave line shape parameters are accurately measured, and their uncertainty does not strongly affect the C_{SP} factors. The S-wave line shape, however, is less precisely measured and has higher uncertainties [67]. Therefore, the P-wave line shape uncertainty is ignored, but the S-wave line shape uncertainty is propagated as systematics. The S-wave line shape parameters are varied within their uncertainties, and the C_{SP} factors are recomputed. The differences between the parameters of interest estimated with the new C_{SP} factors and the parameters of interest estimated with the baseline C_{SP} factors are assigned as systematic uncertainty.

Limited knowledge of magnetic field

Due to the limited knowledge of the magnetic field, the momentum is known up to a certain precision in LHCb. As fine-tuning the magnetic field description is rather complex, a data-driven correction procedure, known as "momentum scaling", is used. The corrections are derived in bins of track angles using known resonances. These corrections are of permille size and, if not included, lead to a couple of MeV/c^2 bias in B_s^0 mass. The full description of the correction procedure can be found elsewhere [102].

However, the derived corrections suffer from the limited statistics of the calibration sample. Their statistical uncertainty is propagated to the final parameters of interests by rescaling tracks momenta with correction factors s varied within uncertainties Δs as

$$\frac{p}{q} = \frac{p}{q}\frac{1}{s \pm \Delta s}, \tag{4.79}$$

where q is the charge of the track and p is its momentum. The difference between the parameters of interest estimated with the rescaled $B_s^0 \to J/\psi K^+K^-$ simulated sample and the baseline $B_s^0 \to J/\psi K^+K^-$ simulated samples is assigned as systematics.

Limited knowledge of LHCb length scale

The LHCb length scale is known up to a 0.022% uncertainty [103]. This uncertainty is propagated to the decay times, and the parameters of interest are measured with the scaled decay time. In [67], this uncertainty is estimated with re-reconstructed simulated $B_s^0 \to J/\psi K^+K^-$ candidates with scaled final state tracks. Using the first or the second approach does not affect the total uncertainty as published in [63]. However, following the approach in [67], the Δm_s uncertainty from the LHCb length scale would go down from 0.0042 to 0.0031.

Ignoring decays via $f_2(1270)$ resonant

In addition to the S-wave contribution, which includes non-resonant $B_s^0 \to J/\psi K^+K^-$ decay and decay via $f_0(980)$, the $B_s^0 \to J/\psi K^+K^-$ final state can be reached via the $f_2(1270)$, known as D-wave. It's rather low, $\mathcal{O}(10^{-6})$, [20]. However, it can still interfere with the measurement, introducing a bias. This bias is evaluated conservatively with a special $f_2(1270)$-enhanced $B_s^0 \to J/\psi K^+K^-$ simulated sample. The bias is assigned from the difference with the generated values of parameters of interest.

Kinematic corrections configuration

In Sec. 4.5.3, the kinematic corrections of the simulated samples are described. They are done using machine learning techniques, which depend on the training configuration. Alternative configurations are used to repeat the analysis. The difference between the parameters of interest estimated with alternative configurations and the baseline parameters of interest is assigned as systematic uncertainty.

Peaking candidates in the simulated combinatorial background

About 4.4% of all reconstructed candidates in the $B_s^0 \to J/\psi K^+K^-$ simulated samples are made of daughter tracks that do not originate from signal B_s^0. These candidates are usually excluded from the simulated sample and are only used as a proxy of the combinatorial background in the decay time and

angular acceptance procedures. However, some of them peak close to the B_s^0 meson mass.

To evaluate the bias from ignoring those peaking candidates, the *sPlot* background subtraction is repeated exclusively for the candidates with daughter tracks not from signal B_s^0. After assigning the sWeights, these candidates are merged with all other simulated candidates. The differences between the obtained parameters of interest and the parameters of interest estimated on the baseline simulated samples are taken as the bias. Note that this systematics might be overestimated, as in principle, part of the peaking candidates would be picked up in the baseline mass model with floating width parameters, see Sec. 4.3.

Ignoring angular resolution

As described in Sec. 4.1, the effect of the finite angular resolution is ignored. To quantify the bias by ignoring this angular resolution, the analysis is repeated on the simulated $B_s^0 \to J/\psi K^+K^-$ sample. The angular acceptance weights twice: first, using generated helicity angles and then using reconstructed helicity angles. The final fit is repeated for two sets of weights. The difference in estimated parameters of interest between two fits is taken as a systematic bias. A note of warning: the angular resolution is partially taken into account using the reconstructed angles in Eq. 4.8 rather than generated ones. Therefore, this is a rather conservative estimation of this uncertainty.

Ignored correlations

The bias from ignored correlations is partially covered by **"Correlation of $m(J/\psi K^+K^-)$ with observables"**. However, this does not include non-mass-related correlations. To catch any uncovered bias from ignored correlations, the analysis is repeated fully on the simulated sample of a mixture of $B_s^0 \to J/\psi K^+K^-$ decaying via $\phi(1020)$, $f_0(980)$ and without a resonance. The simulated sample does not contain background contributions, but the largest of them, the combinatorial background, is additionally generated and merged into the sample. The technique to generate the combinatorial background is described in Sec. 5.2. of [104]. The differences between the generated parameters of interest and estimated parameters of interest are assigned as systematic uncertainty. If the difference is smaller than the one obtained with the method described in the section **"Correlation of $m(J/\psi K^+K^-)$ with observables"**, it is ignored.

Lifetime bias from $B_c^+ \to B_s^0\pi^+$ decays

About 2% of the collected $B_s^0 \to J/\psi K^+K^-$ candidates are initially produced as B_c^+ mesons which subsequently decay into $B_s^0\pi^+$ [67]. Because the B_c^+ meson has an average lifetime of $0.501 \pm 0.009\,\text{ps}$ [20], the observed decay time of these B_s^0 candidates is, on average longer, since it originates from the B_c^+ decay vertex rather than the primary vertex. Such B_s^0 candidates introduce bias in estimated lifetimes. The size of this bias is estimated using a simulated $B_c^+ \to B_s^0(\to J/\psi\,\phi)\pi^+$ sample. The details on the bias estimation can be found in [67]. The resulting biases are the following: for $\Delta\Gamma_s$, the bias $\delta\Delta\Gamma_s$ is $-0.00145 \pm 0.00021\,\text{ps}^{-1}$; for $\Gamma_s - \Gamma_d$, the bias $\delta(\Gamma_s - \Gamma_d)$ is $-0.00168 \pm 0.00068\,\text{ps}^{-1}$. Both biases are taken into account by parametrizing in the decay time angular maximum likelihood fit as

$$\begin{aligned} \Delta\Gamma_s &\to \Delta\Gamma_s + \delta\Delta\Gamma_s + \sigma_{\delta\Delta\Gamma_s} \times s \\ \Gamma_s - \Gamma_d &\to \Gamma_s - \Gamma_d + \delta(\Gamma_s - \Gamma_d) + \sigma_{\delta(\Gamma_s-\Gamma_d)} \times s \end{aligned} \tag{4.80}$$

where $\sigma_{\delta\Delta\Gamma_s,\delta(\Gamma_s-\Gamma_d)}$ are the bias uncertainties, and s is the correlation coefficient between the two biases, which is left free in the maximum likelihood fit.

4.8 Results

The maximum likelihood fit of the model described in Eq. 4.13-4.14 is performed on a sample of approximately $5.5 \cdot 10^5$ candidates collected in $2015 - 2018$ by LHCb. The data sample corresponds to about 5.7 fb^{-1} of luminosity. It is divided into 48 subsamples: by year, by trigger category, and by K^+K^- mass bins, as defined in Tab. 4.1. The PDF model is configured specifically for each of the subsamples. For example, the S-P wave interference coefficients C_{SP} differ per K^+K^- mass bin, whereas the time acceptance spline coefficients are different per year and per trigger category. Table 4.20 summarizes the PDF configurations according to the subsample.

The collected candidates are weighted with sWeights as described in Sec. 4.3. About $3.5 \cdot 10^5$ signal $B_s^0 \to J/\psi K^+K^-$ candidates are found among all candidates. Applying sWeights statistically subtracts the background candidates from the sample and reveals the signal $B_s^0 \to J/\psi K^+K^-$ candidates distributions. This allows simplification of the final maximum likelihood fit, as only a PDF describing the signal has to be used. The likelihood from

	year	trigger category	K^+K^- mass bin	Value
Γ_d	-	-	-	0.65789 [70]
Decay time resolution	✓	-	-	Tab. 4.11
Decay time bias	✓	-	-	Tab. 4.13
Decay time acceptance	✓	✓	-	Tab. 4.15
Angular weights	✓	✓	-	Tab. 4.16
Flavour tagging calibration	✓(shared for 2015 and 2016)	-	-	Tab. 4.7
S-P wave interference effective coupling	-	-	✓	Tab. 4.1

Table 4.20: Description of the subsample specific and shared configurations of the maximum likelihood fit.

Eq. 4.1 then becomes weighted with per candidate weights w_i as,

$$\mathcal{L}^{weighted}(\theta|\vec{x}) = \prod_{i=1}^{n} P(x_i|\theta)^{w_i}, \tag{4.81}$$

where i is the candidate index; n is the number of candidates; $P(x_i|\theta)$ is the PDF for candidate x_i computed with parameters θ. However, the error estimates computed with Eq. 4.3 for this weighted likelihood are underestimated. To correct for this, the weighted likelihood is scaled with a factor α, which is defined as [105],

$$\alpha = \frac{\sum_{i}^{n} w_i}{\sum_{i}^{n} w_i^2} \tag{4.82}$$

where i is the candidate index and w_i is a weight. The coverage of errors with this α correction applied is validated with the bootstrapping method as mentioned in Sec. 4.6.5. Asymptotically correct covariance C_{ij} for the weighted maximum likelihood estimator can be computed [100] as,

$$C_{ij} = \sum_{k,l=1}^{M} H_{ik}^{-1} \left(\frac{\partial \ln \mathcal{L}^{weighted}(\theta|\vec{x})}{\partial \theta_k} \frac{\partial \ln \mathcal{L}^{weighted}(\theta|\vec{x})}{\partial \theta_l} \right) H_{lj}^{-1}\Big|_{\hat{\theta}}, \tag{4.83}$$

where $\hat{\theta}$ is the maximum likelihood estimator of the PDF parameters and H_{ik} is the second derivative of the weighted maximum likelihood,

$$H_{ij} = \frac{\partial^2 \ln \mathcal{L}^{weighted}(\theta|\vec{x})}{\partial \theta_i \partial \theta_j}. \tag{4.84}$$

This approach is known as “asymptotic errors”. As was described in Sec. 4.6.5, this was tested in the RooFit implementation of the fit and was found to give

similar results to the bootstrapping procedure. However, in future iterations of this analysis, the "asymptotical errors" should be preferred over Eq. 4.82.

The flavour tagging calibration is included in the maximum likelihood as a Gaussian constraint. Therefore, the uncertainties from the flavour tagging calibration are naturally propagated into the statistical uncertainties of the estimated parameters of interest. Additionally, a correction is applied to account for the $B_c^+ \to B_s^0\pi^+$ contribution as described in Sec. 4.7.

4.8.1 Baseline results

This thesis shows results from the maximum likelihood fit that is based on the RooFit framework [84] [106]. Two other independent implementations of this fit were used to cross-check the results of the RooFit implementation [67]. All three agree with each other and the results of the fitter based on Graphic Processor Unit architecture are quoted in the public results [63]. The results of the RooFit implementation are quoted in Tab. 4.21, where the first uncertainty is the statistical one and the second is the systematic one from Tab. 4.17-4.18. The correlation matrix can be found in Tab. A.2. The quoted statistical uncertainty corresponds to the second derivative of likelihood with respect to the parameters of interest, assuming the parameter is equal to its maximum likelihood estimator. Figures 4.46 show projections of the PDF onto the observables for background-subtracted $B_s^0 \to J/\psi\, K^+K^-$ candidates.

4.8.2 Polarization dependence

Table 4.21 shows the estimated values of parameters described in Sec. 2.4, assuming that $|\lambda|$ and ϕ_s are polarization independent. However, the penguin contributions to $|\lambda|$ and ϕ_s can be polarization-dependent, see Sec. 2.6. The observation that $|\lambda| \approx 1$, see Tab. 4.21, is consistent with the hypothesis that the penguin contributions are small, see Sec. 2.5. The interference between the penguin and tree decay amplitudes in Fig. 2.7 would generally generate a direct CP asymmetry, which would cause lambda to deviate from unity. Given the observed result, one may conclude that the penguin contribution is not important.

However, if the penguin contributions interfere destructively when combined, it may be that they are more pronounced in a polarization-dependent fit. In this case, both $|\lambda|$ and ϕ_s are assumed to depend on polarization. Results of the polarization-dependent fit are shown in Tab. 4.22. Because of the non-parabolic likelihood, the asymmetric uncertainties are used, rather than Eq. 4.3. The asymmetric statistical uncertainties $\Delta_\pm$ are defined as a

Parameter	Value $\pm$ stat. $\pm$ syst.
ϕ_s, rad	$-0.038 \pm 0.021 \pm 0.006$
$\lvert\lambda\rvert$	$1.001 \pm 0.010 \pm 0.005$
$\Delta\Gamma_s$, ps^{-1}	$0.0848 \pm 0.0044 \pm 0.0024$
$\Gamma_s - \Gamma_d$, ps^{-1}	$-0.0058 \pm 0.0013 \pm 0.0014$
Δm_s, ps^{-1}	$17.742 \pm 0.032 \pm 0.008$
$\lvert A_0\rvert^2$	$0.5179 \pm 0.0017 \pm 0.0032$
$\lvert A_\perp\rvert^2$	$0.2463 \pm 0.0023 \pm 0.0024$
$\delta_\parallel - \delta_0$, rad	$3.147 \pm 0.061 \pm 0.052$
$\delta_\perp - \delta_0$, rad	$2.906 \pm 0.074 \pm 0.048$
F_S , $m(K^+K^-) \in [990, 1008)$, MeV/$c^2$	$0.472 \pm 0.024 \pm 0.027$
F_S , $m(K^+K^-) \in [1008, 1016)$, MeV/$c^2$	$0.041 \pm 0.005 \pm 0.010$
F_S, $m(K^+K^-) \in [1016, 1020)$, MeV/$c^2$	$0.003 \pm 0.001 \pm 0.023$
F_S, $m(K^+K^-) \in [1020, 1024)$, MeV/$c^2$	$0.004 \pm 0.002 \pm 0.032$
F_S, $m(K^+K^-) \in [1024, 1032)$, MeV/$c^2$	$0.051 \pm 0.007 \pm 0.027$
F_S, $m(K^+K^-) \in [1032, 1050]$, MeV/$c^2$	$0.152 \pm 0.011 \pm 0.051$
$\delta_S - \delta_\perp$, rad, $m(K^+K^-) \in [990, 1008)$, MeV/$c^2$	$2.06 \pm 0.13 \pm 0.19$
$\delta_S - \delta_\perp$, rad, $m(K^+K^-) \in [1008, 1016)$, MeV/$c^2$	$1.63 \pm 0.20 \pm 0.41$
$\delta_S - \delta_\perp$, rad, $m(K^+K^-) \in [1016, 1020)$, MeV/$c^2$	$1.12 \pm 0.35 \pm 0.19$
$\delta_S - \delta_\perp$, rad, $m(K^+K^-) \in [1020, 1024)$, MeV/$c^2$	$-0.21 \pm 0.14 \pm 0.31$
$\delta_S - \delta_\perp$, rad, $m(K^+K^-) \in [1024, 1032)$, MeV/$c^2$	$-0.66 \pm 0.073 \pm 0.17$
$\delta_S - \delta_\perp$, rad, $m(K^+K^-) \in [1032, 1050]$, MeV/$c^2$	$-1.03 \pm 0.079 \pm 0.07$

Table 4.21: Parameters of interest estimated with a RooFit fitter on background-subtracted $B_s^0 \to J/\psi K^+K^-$ sample collected in LHCb between 2015-2018, where $|A_S|^2 = F_S/(1 - F_S)$. Systematic uncertainties are taken from Tab. 4.17-4.18.

difference between MLE $\hat{\theta}$ and $\theta^{'}$ such that,

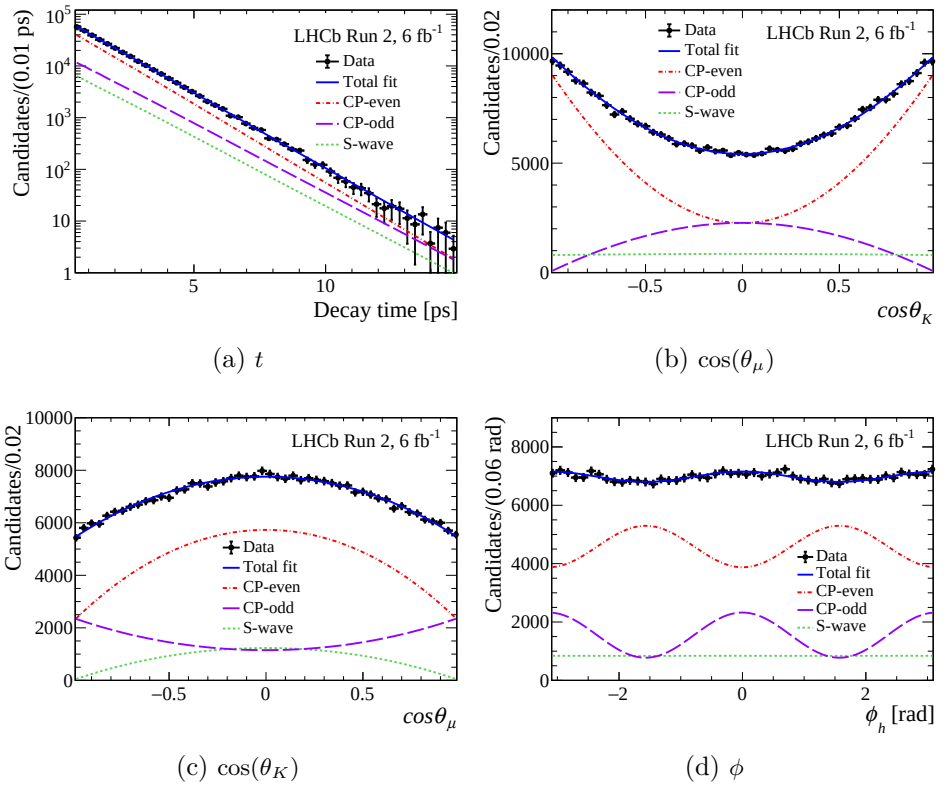
$$\ln \mathcal{L}(\theta^{'}|\vec{x}) - \ln \mathcal{L}(\hat{\theta}|\vec{x}) = 0.5, \tag{4.85}$$

where there are two possible solutions for $\theta^{'}$: for $\theta^{'} > \hat{\theta}$ the uncertainty is $\Delta_+ = \theta^{'} - \hat{\theta}$ and for $\theta^{'} < \hat{\theta}$ the uncertainty is $\Delta_- = \hat{\theta} - \theta^{'}$. If the errors are the same within the precision of the reported parameter, an average uncertainty is reported. Tab. 4.21 and Tab. 4.22 are found to be consistent; see [67] for the details when accounting for the difference in the number of degrees of freedom. Therefore, no penguin contribution has yet been observed even in the polarization-dependent fit.

As described in Sec. 2.4.1, the time-dependent $\mathcal{CP}$-asymmetry between $\overline{B}_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi \phi(1020)$ is proportional to $\sin\phi_s \sin(\Delta m_s t)$.

Parameter	Value ± stat.
$\phi_s{}^0$, rad	-0.034 ± 0.023
$\phi_s{}^{\parallel} - \phi_s{}^0$, rad	-0.002 ± 0.021
$\phi_s{}^{\perp} - \phi_s{}^0$, rad	$-0.001^{+0.020}_{-0.019}$
$\phi_s{}^S - \phi_s{}^0$, rad	$0.023^{+0.027}_{-0.026}$
$\lvert\lambda^0\rvert$	$0.969^{+0.024}_{-0.024}$
$\lvert\lambda^{\parallel}/\lambda^0\rvert$	$0.982^{+0.055}_{-0.052}$
$\lvert\lambda^{\perp}/\lambda^0\rvert$	$1.106^{+0.081}_{-0.075}$
$\lvert\lambda^S/\lambda^0\rvert$	$1.122^{+0.084}_{-0.079}$
$\Delta\Gamma_s$, ps^{-1}	$0.0848^{+0.0044}_{-0.0050}$
$\Gamma_s - \Gamma_d$, ps^{-1}	$-0.0058^{+0.0016}_{-0.0014}$
Δm_s, ps^{-1}	$17.749^{+0.033}_{-0.032}$
$\lvert A_0\rvert^2$	0.5284 ± 0.0101
$\lvert A_{\perp}\rvert^2$	0.2267 ± 0.0141
$\delta_{\parallel} - \delta_0$, rad	3.146 ± 0.061
$\delta_{\perp} - \delta_0$, rad	2.916 ± 0.077
F_S, $m(K^+K^-) \in [990, 1008)$, MeV/$c^2$	$0.448^{+0.029}_{-0.030}$
F_S, $m(K^+K^-) \in [1008, 1016)$, MeV/$c^2$	0.037 ± 0.005
F_S, $m(K^+K^-) \in [1016, 1020)$, MeV/$c^2$	0.002 ± 0.001
F_S, $m(K^+K^-) \in [1020, 1024)$, MeV/$c^2$	0.0035 ± 0.002
F_S, $m(K^+K^-) \in [1024, 1032)$, MeV/$c^2$	0.046 ± 0.007
F_S, $m(K^+K^-) \in [1032, 1050]$, MeV/$c^2$	$0.140^{+0.013}_{-0.012}$
$\delta_S - \delta_{\perp}$, rad, $m(K^+K^-) \in [990, 1008)$, MeV/$c^2$	$2.08^{+0.12}_{-0.14}$
$\delta_S - \delta_{\perp}$, rad, $m(K^+K^-) \in [1008, 1016)$, MeV/$c^2$	$1.62^{+0.20}_{-0.19}$
$\delta_S - \delta_{\perp}$, rad, $m(K^+K^-) \in [1016, 1020)$, MeV/$c^2$	$1.12^{+0.39}_{-0.29}$
$\delta_S - \delta_{\perp}$, rad, $m(K^+K^-) \in [1020, 1024)$, MeV/$c^2$	$-0.20^{+0.12}_{-0.16}$
$\delta_S - \delta_{\perp}$, rad, $m(K^+K^-) \in [1024, 1032)$, MeV/$c^2$	$-0.65^{+0.07}_{-0.08}$
$\delta_S - \delta_{\perp}$, rad, $m(K^+K^-) \in [1032, 1050]$, MeV/$c^2$	$-1.00^{+0.08}_{-0.09}$

Table 4.22: Polarization-dependent parameters of interest estimated with a RooFit fitter on background-subtracted $B_s^0 \to J/\psi K^+K^-$ sample that was collected in LHCb between 2015-2018, where $|A_S|^2 = F_S/(1 - F_S)$ and P-wave amplitudes are normalized such that $|A_0|^2 + |A_{\perp}|^2 + |A_{\parallel}|^2 = 1$. Only statistical uncertainty is shown.



(a) t (b) $\cos(\theta_\mu)$

(c) $\cos(\theta_K)$ (d) ϕ

Figure 4.46: Projections of the decay-time angular PDF from Eq. 4.13 with MLE parameters on the observables. The background-subtracted $B_s^0 \to J/\psi K^+K^-$ candidates are also shown. Taken from [63].

A work-in-progress method of time-dependent $\mathcal{CP}$-asymmetry visualization is presented in Appendix A.5.

4.8.3 Combination with other LHCb results and comparison with previous measurements

$B_s^0 \to J/\psi K^+K^-$ is the cleanest decay channel to probe ϕ_s, see Chapter 1. Nevertheless, the obtained ϕ_s in Tab. 4.21 is statistically limited. To reduce the statistical uncertainty, the results of this analysis from Tab. 4.21 were combined with all other LHCb measurements sensitive to ϕ_s. These include measuring ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays for $m_K K > 1050$ MeV/c^2 [50]; with $\overline{B}_s^0 \to D_s^+ D_s^-$ decays [54]; with $\overline{B}_s^0 \to J/\psi \pi^+\pi^-$ [52]; with $B_s^0 \to$

$\psi(2S)\phi$ [53]; with $B_s^0 \to J/\psi K^+K^-$ decays collected in 2015-2016 [49]; with $B_s^0 \to J/\psi (e^+e^-)\phi$ decays [51] ; with $B_s^0 \to J/\psi K^+K^-$ decays collected in 2011-2012 [48].

The combination results are

$$\begin{aligned}
\phi_s &= -0.031 \pm 0.018 \text{ rad}, \\
|\lambda| &= 0.990 \pm 0.010, \\
\Gamma_s &= 0.6563 \pm 0.0020 \text{ ps}^{-1}, \\
\Delta\Gamma_s &= 0.0846 \pm 0.0039 \text{ ps}^{-1}, \\
|A_\perp|^2 &= 0.2471 \pm 0.0031, \\
|A_0|^2 &= 0.5175 \pm 0.0035, \\
\delta_\perp - \delta_0 &= 2.94 \pm 0.07 \text{ rad}, \\
\delta_\parallel - \delta_0 &= 3.150 \pm 0.062 \text{ rad}, \\
\Delta m_s &= 17.740 \pm 0.027 \text{ ps}^{-1},
\end{aligned} \tag{4.86}$$

where the uncertainties include statistical and systematical uncertainties; the correlations between systematical uncertainties are included; Γ_d is Gaussian constrained to the world-average from [62] to obtain Γ_s from $\Gamma_s - \Gamma_d$. Figure 4.47 shows the 68% confidence levels[9] in $\phi_s - \Delta\Gamma_s$ plane for both the combined results and results obtained in separate measurements [48–54, 63]. All the ϕ_s measurements with $B_s^0 \to J/\psi K^+K^-$ decays, excluding the one in the high K^+K^- mass ($m_{KK} > 1050$ MeV/c^2), are shown together.

Figures 4.48-4.49 show a comparison of ϕ_s, $\Delta\Gamma_s$ and Γ_s obtained in this thesis to the 3 fb^{-1} LHCb Run 1 [48], CMS combined [61] and ATLAS combined [59]. Table 4.23 shows the results of this thesis, the LHCb Run 1, the latest up-to-date CMS, and the latest up-to-date ATLAS measurements for completeness.

[9]The confidence level is reported for the two-dimensional Gaussian likelihood.

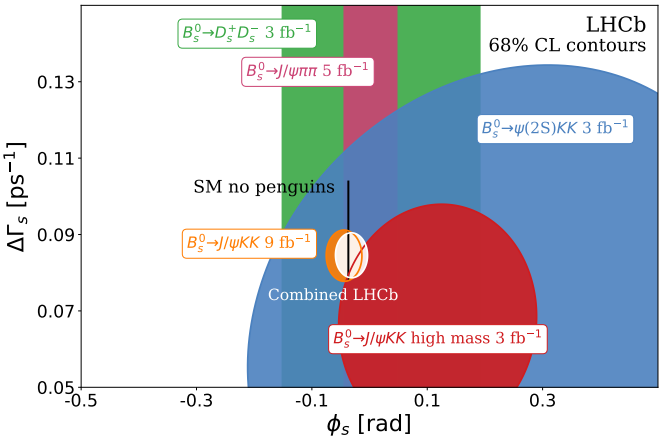


Figure 4.47: The 68% confidence levels in $\phi_s - \Delta\Gamma_s$ plane for the combined and separate LHCb measurements. The SM indicates a tree-level SM prediction. Individual measurements are the following: $\overline{B}{}_s^0 \to D_s^+ D_s^-$ [54]; $\overline{B}{}_s^0 \to J/\psi\,\pi^+\pi^-$ [52]; $B_s^0 \to \psi(2S)\phi$ [53]; $B_s^0 \to J/\psi\,K^+K^-$ high mass [50]; $B_s^0 \to J/\psi\,K^+K^-$ [48,49,51,63]. Taken from [63].

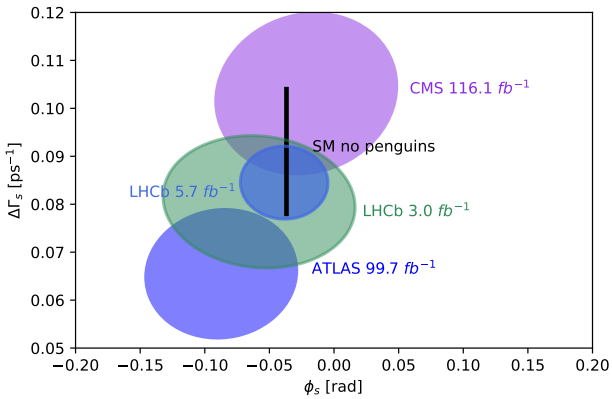


Figure 4.48: The 68% confidence levels in $\phi_s - \Delta\Gamma_s$ plane for the ATLAS combination [59], CMS combination [61], 3 fb^{-1} LHCb Run 1 [48] and this thesis (LHCb 5.7 fb^{-1}). The Standard Model prediction for ϕ_s is from [14] for ϕ_s and $\Delta\Gamma_s$ is from [107].

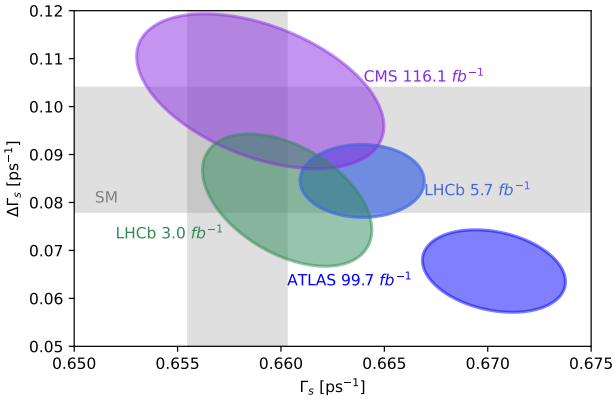


Figure 4.49: The 68% confidence levels in $\Gamma_s - \Delta\Gamma_s$ plane for the ATLAS combination [59], CMS combination [61], 3 fb^{-1} LHCb Run 1 [48] and this thesis (LHCb 5.7 fb^{-1}). The Standard Model prediction for $\Delta\Gamma_s$ is from [107] and for Γ_s is as quoted by [62].

Parameter	This thesis, 5.7 fb^{-1}	LHCb Run 1, 3.0 fb^{-1}	ATLAS, 80.5 fb^{-1}	CMS, 96.4 fb^{-1}
ϕ_s, rad	$-0.038 \pm 0.021 \pm 0.006$	$0.058 \pm 0.049 \pm 0.006$	$-0.081 \pm 0.041 \pm 0.022$	$-0.011 \pm 0.050 \pm 0.010$
$\lvert\lambda\rvert$	$1.001 \pm 0.010 \pm 0.005$	$0.964 \pm 0.019 \pm 0.007$	fixed to 1	$0.972 \pm 0.026 \pm 0.008$
$\Delta\Gamma_s$, ps^{-1}	$0.0848 \pm 0.0044 \pm 0.0024$	$0.0805 \pm 0.0091 \pm 0.0032$	$0.0607 \pm 0.0047 \pm 0.0043$	$0.6531 \pm 0.0042 \pm 0.0026$
Γ_s, ps^{-1}	$0.6637 \pm 0.0013 \pm 0.0014$	$0.6603 \pm 0.0027 \pm 0.0015$	$0.6687 \pm 0.0015 \pm 0.0022$	$0.6531 \pm 0.0042 \pm 0.0026$
Δm_s, ps^{-1}	$17.742 \pm 0.032 \pm 0.008$	$17.711^{+0.055}_{-0.057} \pm 0.011$	fixed to [108]	$17.51^{+0.10}_{-0.09} \pm 0.03$
$\lvert A_0\rvert^2$	$0.5179 \pm 0.0017 \pm 0.0032$	$0.5241 \pm 0.0034 \pm 0.0067$	$0.5131 \pm 0.0013 \pm 0.0038$	$0.5350 \pm 0.0047 \pm 0.0049$
$\lvert A_\perp\rvert^2$	$0.2463 \pm 0.0023 \pm 0.0024$	$0.2504 \pm 0.0049 \pm 0.0036$	$0.2656 \pm 0.0023 \pm 0.0044$	$0.2337 \pm 0.0063 \pm 0.0045$
$\delta_\parallel - \delta_0$, rad	$3.147 \pm 0.061 \pm 0.052$	$3.26^{+0.10+0.06}_{-0.17-0.07}$	$2.94 \pm 0.05 \pm 0.09$ $3.35 \pm 0.05 \pm 0.09$	$3.18 \pm 0.12 \pm 0.03$
$\delta_\perp - \delta_0$, rad	$2.906 \pm 0.074 \pm 0.048$	$3.08^{+0.14}_{-0.15} \pm 0.06$	$2.91 \pm 0.11 \pm 0.06$ $3.12 \pm 0.11 \pm 0.06$	$2.77 \pm 0.16 \pm 0.05$

Table 4.23: Comparison of the parameters of interest between this thesis (and [63]), LHCb Run 1 [48], ATLAS [59] and CMS [61]. Both $\delta_\parallel - \delta_0$ and $\delta_\perp - \delta_0$ have two minima in ATLAS [59].

5 Preparation for future data taking

The measured value of ϕ_s from Tab. 4.21 is statistically limited and is compatible with zero. Unfortunately, that means that mixing-induced $\mathcal{CP}$-violation has not yet been observed in $B_s^0 \to J/\psi K^+K^-$ decays. To observe it in the future, more data is needed. Luckily, within the following years, the LHCb collaboration plans to collect 50 fb^{-1} of integrated luminosity. Assuming that the detector performance and flavour tagging remain the same, the statistical uncertainty on the ϕ_s measurement should be reduced by a factor of two compared to the current result. Assuming the central value does not change, the first hint of the mixing-induced $\mathcal{CP}$-violation in $B_s^0 \to J/\psi K^+K^-$ decays (aka 3 σ effect) is within LHCb reach.

Such an unprecedented amount of data is possible because of the higher level of instantaneous luminosity: $2 \times 10^{33}\text{s}^{-1}\text{cm}^{-2}$ [109] and a redesign of the trigger system. Because of an increase of instantaneous luminosity by a factor of five when compared to the Run 1/2 detector described in Chapter 3, the detector had to be upgraded to cope with the higher occupancy and an environment with an increased radiation fluence. In the redesigned trigger system the hardware trigger `L0` is entirely removed and the first software trigger `HLT1` has to operate at the full p-p collision rate, which is 30 MHz[1]. The detector's readout electronics have to be replaced entirely to provide the information to `HLT1` at the necessary speed.

Overall, four factors drive the new LHCb detector development:

1. the increased occupancy due to the increased luminosity, which defines the granularity of the new detector;

[1]The 30 MHz is the average frequency of the non-empty proton-proton bunch crossings at LHCb, which is a bit lower than the 40 MHz bunch crossing rate of the LHC beams.

2. the continuous readout at 30 MHz, which means fast electronics;

3. the increased fluence, which requires readout electronics and sensors that can cope with more harsh radiation conditions;

4. maintaining physics performance, such as the resolution of observables, efficiencies on signal, and rejection of background, as good as before.

After an extensive R&D, most of the old LHCb detector from Fig. 3.2 was replaced with a new one, which has been in commissioning and operation since 2022. The schematics of the new detector can be seen in Fig. 5.1.

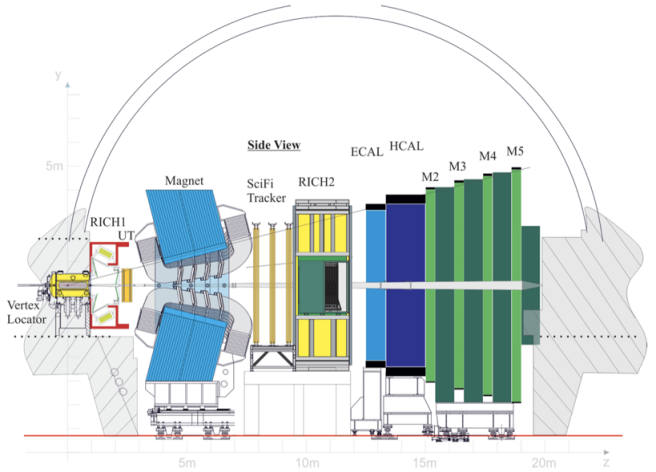


Figure 5.1: Schematics of new LHCb detector (since 2022). From the interaction point on the left: the new pixel Vertex Locator (VELO, tracking), upgraded Ring Imaging Cherenkov detector (RICH1, particle identification), the new Upstream tracker (UT, tracking), magnet, the new Scintillating tracker (SciFi, tracking), upgraded second Ring Imaging Cherenkov detector (RICH1, particle identification), electromagnetic calorimeter (ECAL) with new readout electronics, hadronic calorimeter (HCAL) with new readout electronics and Muon stations (M2-M5, muon identification) with new readout electronics. In comparison with the old detector shown in Fig. 3.2, the tracking is entirely replaced, Cherenkov detectors are upgraded, and all the readout electronics are replaced. The first muon station M1, the scintillating pad detector, and the preshower detector are gone.

This chapter summarizes work done by the author in preparation for the new data-taking period with the upgraded detector.

Firstly, the contribution to the new pixel Vertex Locator commissioning is described in Sec. 5.1. The results presented reflect the state of the art

as of the second half of 2022. As improvements occur rapidly during any commissioning period, some parts are likely to be superseded by the time this thesis is read. The author has developed the so-called "noise-rate" based equalization and has performed the first equalization scans of the installed detector in 2022.

Secondly, the author of this thesis has provided a monitoring algorithm that compares reconstructed objects at the different trigger stages in Sec. 5.2. This is important for both reconstruction algorithms fine-tuning and limiting the fake negative selection decisions made with the partially reconstructed events at `HLT1` level. To expand on the latter, if the event was disregarded based on the partially reconstructed objects, it should still be disregarded if the same cuts are applied to fully reconstructed objects. If this would not be the case, then it implies that a potential event of interest was lost prematurely in a way that may introduce non-trivial efficiencies, which may affect any subsequent measurement utilizing these events.

5.1 VELO front-end ASIC equalisation

As was described in Sec. 4.1, to measure ϕ_s three important requirements must be satisfied:

1. The initial flavour of the B_s^0 meson (aka flavour tagging) must be known;
2. The decay time must be measured with the precision necessary to measure Δm_s;
3. The angular observables must be determined to disentangle polarization amplitudes.

For the flavour tagging, both the signal candidates and the tagging object need to be reconstructed and matched to their primary vertex. The decay time is measured from the distance between the primary vertex and the B_s^0 decay vertex, see Eq. 4.40. On average, the boosted B_s^0 meson travels about 7 mm before decaying, which is within the coverage of the VELO stations. Moreover, the decay time resolution is mainly defined by the resolution on the decay vertex position [25]. So for the first two conditions necessary to measure ϕ_s, the vertex locator near the interaction point of proton beams is crucial.

About 1.7 p-p interactions were visible per event in the original LHCb detector. With the increase of instantaneous luminosity, about five visible

p-p interactions are expected per event. As a result, the combinatorics of the vertex reconstruction increases and makes it more complicated to distinguish one primary vertex from another. The original VELO, see Chapter 3, does not have the granularity necessary to cope with the higher occupancy of the detector and successfully reconstruct vertices. Moreover, the electronics were not developed for the higher fluence expected during the operation of the detector.

The primary way to discriminate b-particles from the background is by applying detachment requirements on their daughter tracks. As described in Sec. 4.2.1, for $B_s^0 \to J/\psi K^+K^-$, the χ^2_{IP} of the daughter tracks is used to filter the $B_s^0 \to J/\psi K^+K^-$ candidates from the background. The achievable impact parameter resolution is another driving force for the development of the new interaction point tracking stations. The impact parameter resolution depends on the VELO aperture radius (see Eq. 1 in [110]) and on the material budget of the detector. The VELO aperture radius is the distance between the VELO sensors and the beam axis. For the upgrade of the detector, it was chosen to be as close as ~ 5.1 mm to the beam [110]. Such an unprecedentedly small distance to the beam results in an environment with large radiant exposure, which requires suitable radiation-hard electronics. The maximum fluence in the closest point to the beam reaches 8×10^{15} 1 MeV$n_{eq} \cdot cm^{-2}$, which is 19 times more than the original VELO could withstand [110] [35]. The VELO detector consists of 52 modules. Each module hosts four silicon sensor tiles. A tile is glued to three front-end ASICs. Each ASIC has a grid of 256×256 pixels, and the pixel size is $55{\times}55$ μm^2 [111]. To reduce the radiation damage to electronics, the modules are cooled to -25 $^\circ C$. The front-end ASICs, called VeloPix, are based on TimePix3 technology and were developed to meet the specifications of the new VELO detector. VeloPix is a zero-suppressed ASIC, meaning that only the hits are read out, whereas empty pixels are dropped. It can sustain the 900 MHit/s rate, which is required for the 30 MHz continuous readout in LHCb. Figure 5.2 shows the VeloPix schematics, where only the single-pixel logic is shown. First, from left to right, the input signal is amplified in the analog part. A radiation-induced leakage current is compensated by the leakage-current compensator [112]. Then, the input signal is digitized with a discriminator and time-stamped in the digital part. The hit is accepted based on the time-over-threshold (ToT) value, which constitutes the end of the single-pixel logic part. Afterward, pixels are combined into a 2×4 super-pixel. The super-pixel logic is not described here in detail but can be found in [111].

The main job of the VeloPix discriminator is to separate noise from signal

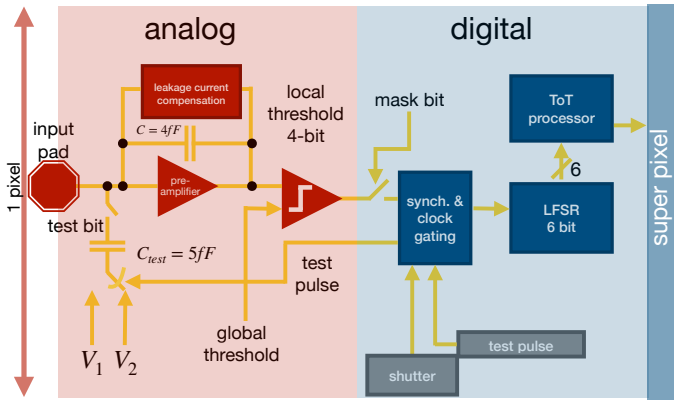


Figure 5.2: The simplified VeloPix ASIC schematics. This schematic was adapted from [111], where the detailed schematics can be found.

hits. For this, the signal level is compared to the threshold value, called Th, see Fig. 5.3. The value of the threshold is configured per ASIC and is known as "global threshold" Th_{global}. However, it can be adjusted on a per-pixel basis with a 4-bit digital-analog converter code (DAC code). This adjusted threshold is called a "local threshold" Th_{local} The adjustment is necessary because each pixel reacts slightly differently, even if the same input signal arrives. Additionally, there is a possibility of masking pixels that are too noisy to read out.

Equalization is a procedure that aims to homogenize the pixel behavior according to the chosen metric in the presence of the same input. An output of the equalization procedure is a set of per-pixel 4-bit DAC codes and a pixel mask.

There are two categories of equalization techniques. The first is equalization on a test signal, which is possible because VeloPix has a dedicated testing circuit that can send a controlled test pulse to the beginning of the digital part, see Fig. 5.2. The second category of equalization techniques is called "equalization on noise". The noise here refers to the absence of a signal, which is expected in the absence of a beam in the LHC. Here, the second category is discussed in detail.

The main role of the discriminator in Fig. 5.2 is to register a signal hit, separating it from noise, which means that ideally, a global threshold is

set to the value that the noise would never be able to reach. However, the higher the threshold, the lower the efficiency of the detector. To set an optimal global threshold, the ASIC should be equalized. For an equalized ASIC, all pixels end up with about the same "representative" thresholds, and the global threshold can then be used to set the above-mentioned average "representative" threshold. The representative threshold is chosen based on the noise rate, either relative rate or absolute. The rate is defined as the number of hits registered per second. Figure 5.3 shows schematically how VeloPix counts the hits. A continuous analog signal crosses the threshold

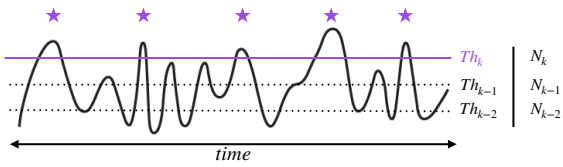


Figure 5.3: Schematics of hit counting with VeloPix. Stars correspond to the counted hits.

multiple times, and it is the crossing upward and downward that is counted as a hit. The number of hits N_k (or the number of threshold crossings) is counted within the time window t_{shutter}, which is defined by the shutter on Fig. 5.2. Figure 5.4 shows the threshold scan, where the number of hits is counted per a chosen shutter time for different values of the global threshold. The left edge of Fig. 5.4 appears because of VeloPix saturation. At the low threshold, the noise rate is high, and the VeloPix does not count individual hits anymore, as the downward crossing of the threshold is unlikely (see Fig. 5.3). The right edge of Fig. 5.4 falls because the upward crossing becomes less and less likely. The left and right edges are expected to be similar if the jitter around the baseline is symmetric and the time over the threshold (ToT) is within the shutter time.

If the representative threshold is chosen based on the absolute noise rate, then the procedure is referred to as "noise rate" equalization. In this case, the representative threshold is chosen such that the number of hits within the time the shutter was open t_{shutter} corresponds to the chosen absolute noise rate N_k/t_{shutter}. The rate of 23 Hz was chosen, which is roughly 0.001 ‰ of the $p-p$ collision rate of LHC [2]. Figure 5.7 shows an example of the representative threshold for the "noise rate". The threshold scan is taken

[2]This corresponds to $t_{\text{shutter}} = 1.3\ s$ and $N_k = 30$. Both can and should be fine-tuned to get 30 Hz.

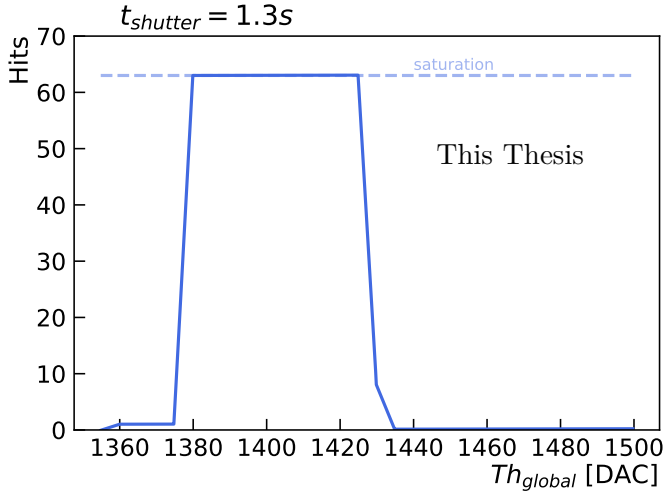


Figure 5.4: An example of a scan of thresholds for a single pixel. A jargon name of this plot is "s-curve", coming from the typical shape of such scans. The VeloPix hit requires upward and downward crossing of the threshold (see Fig. 5.3). At the low threshold, only a few downward crossings of the threshold are observed, which explains the left edge. Due to the limited readout space, the histogram is saturated at 63 counts.

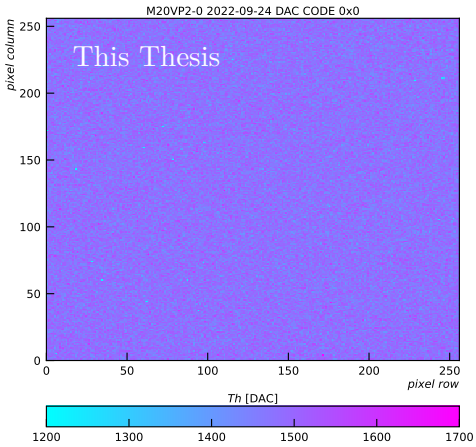
with a defined threshold step. The representative threshold is then computed from linear interpolation between the two nearest to $N_k = 30$ measured points.

When the relative rate is used, the procedure is called "noise baseline" equalization. The representative threshold for the "noise baseline" corresponds to the maximum number of noise hits seen for a pixel, which corresponds to the mean of Fig. 5.4, as can be seen from the Fig. 5.7.

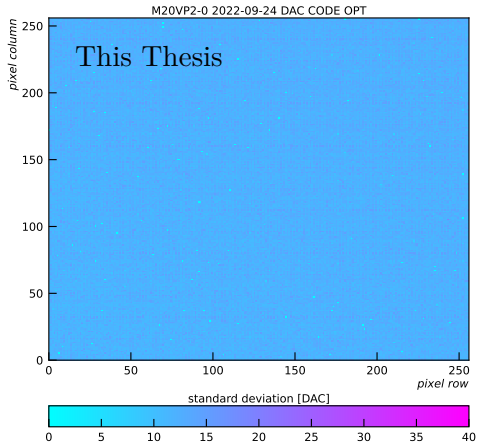
Figure 5.5(a) shows the representative thresholds that correspond to the noise baseline for one ASIC. The width of Fig. 5.4 is called *noise width* and is shown for the same ASIC in Fig. 5.5(b). The width represents the spread of the thresholds for which noise hits are counted, see Fig. 5.6. Figure 5.5(c) shows the representative thresholds that correspond to 23 Hz noise rate.

The choice of representative threshold is the main difference between the noise baseline and noise rate-based equalization procedures. The following steps are the same for both.

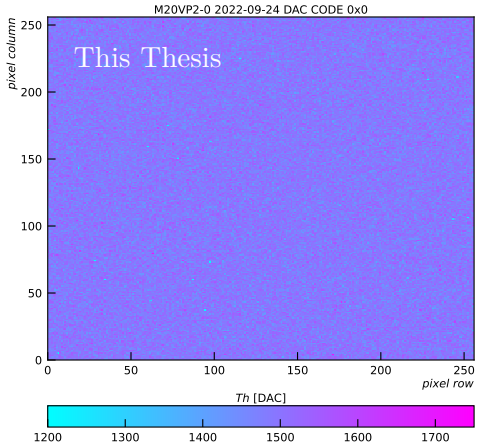
As was mentioned before, the output of the equalization procedure is a set of 4-bit DAC codes and a mask. The 4-bit DAC codes adjust the global threshold on a per-pixel level, which leads to a homogeneous noise behavior



(a) Noise baseline



(b) Noise width



(c) Noise rate

Figure 5.5: Representative thresholds of noise baseline and noise rate procedures taken with extreme DAC code zero and noise width parameters taken with optimized DAC codes. Also, the noise width in thresholds is shown. The $t_{\text{shutter}} = 1.3\ s$.

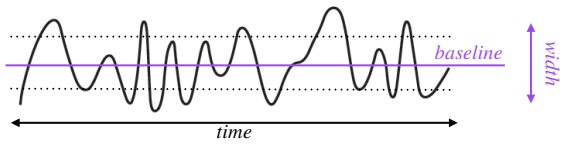


Figure 5.6: Schematics representing the noise baseline and noise width parameters for a pixel.

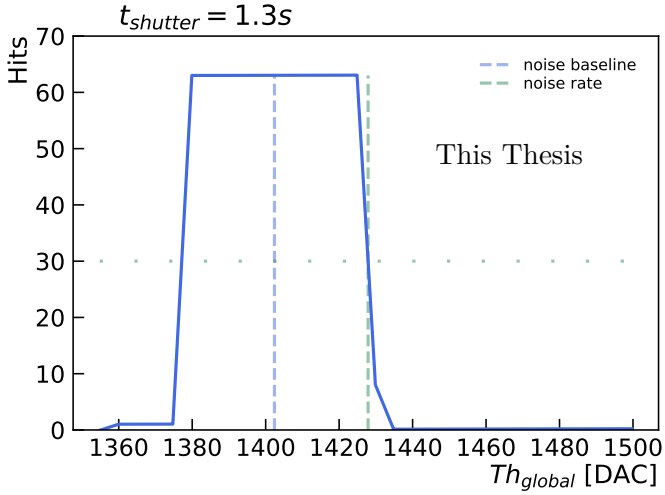


Figure 5.7: Comparison of the pixels representative threshold values for noise baseline and noise rate procedures.

across the pixels.

The optimal DAC codes are taken from the interpolation between the measured representative thresholds with two extreme DAC code values: `0x0` and `0xF`[3]. Figure 5.8 shows the pixels' noise rate representative thresholds for one VeloPix at the edge DAC codes. At DAC code zero, the unchanged representative thresholds of pixels are shown, and at `0xF`, the maximally shifted representative thresholds are shown.

The interpolation between representative thresholds at `0x0` and `0xF` is assumed to be linear,

$$Th_k = Th_k^{\texttt{0x0}} + s \cdot \text{DAC code} \tag{5.1}$$

where $Th_k^{\texttt{0x0}}$ is the representative threshold of pixel k measured with `0x0` DAC code; s is the slope parameter. However, it was observed in the past that the dependence between the DAC code values and the representative thresholds is quadratic. To interpolate a quadratic function at least three points are necessary. Unfortunately, the communication protocol used to take threshold scans in 2022 was very slow[4], and thus, for practical reasons, it was necessary to restrict the procedure and use a linear approximation.

[3] `0xF` is a hex value of 15. DAC code is a 4-bit word, so only values $\in [0, 2^4 - 1]$ can be represented.

[4] It should be mentioned that the LHCb VELO team is developing a faster communication protocol, and therefore, in the future, a more nuanced calibration can be done.

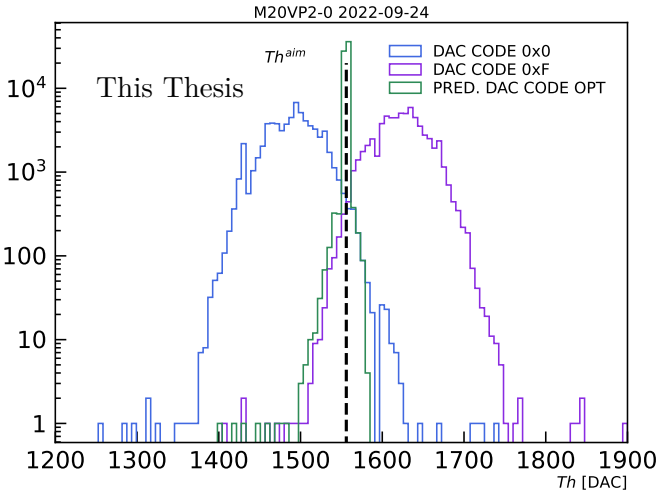


Figure 5.8: Distributions of the representative threshold for two extreme values of DAC code: `0x0` and `0xF`. `PRED. DAC CODE OPT` is the distribution of the predicted representative threshold for optimized DAC code values. It is used to derive the mask. For the noise rate procedure. Th^{aim} is the threshold that defines the optimal DAC code, as defined in Eq. 5.2.

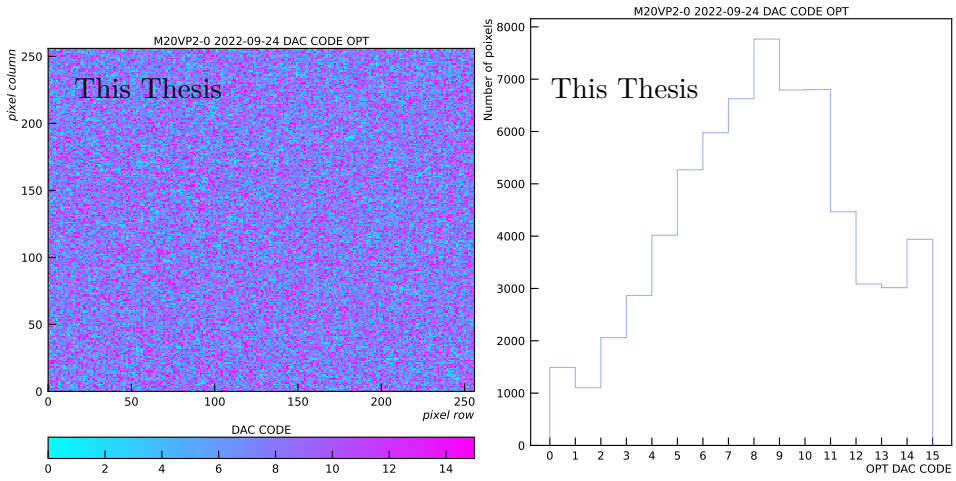
The optimal DAC code is chosen such that the pixel response is close to a chosen Th^{aim}. The Th^{aim} is defined as,

$$Th^{aim} = \frac{\sum_{k=1}^{N_{pixels}} Th_k^{\texttt{0xF}}/N_{pixels} - \sum_{k=1}^{N_{pixels}} Th_k^{\texttt{0x0}}/N_{pixels}}{2}, \tag{5.2}$$

where k is the pixel index; $N_{pixels} = 256 \times 256$, and is the total number of pixels of one VeloPix; $Th_k^{\texttt{0x0}}$ is the representative threshold of pixel k measured with `0x0` DAC code; $Th_k^{\texttt{0xF}}$ is the representative threshold of pixel k measured with `0xF` DAC code; The DAC code is then computed as,

$$\text{DAC code} = \frac{Th^{aim} - Th_k^{\texttt{0x0}}}{s}. \tag{5.3}$$

However, all values are integers. Therefore, on top of using the linear interpolation instead of a quadratic one, an additional imperfection of this calibration comes from rounding to the integer values. If the value of the DAC code exceeds the allowed range of $[0, 2^4 - 1]$, the edge DAC value is assigned. Figure 5.9 shows an example of the optimal DAC codes for the noise rate equalization.



(a) Geometrical location of pixels (b) Distribution of the optimal DAC codes

Figure 5.9: An example of the optimized value for DAC code from the noise rate procedure.

Because of all these little imperfections, not all of the pixels will end up at exactly Th^{aim}. Figure 5.8 also shows the predicted Th_k at the optimal DAC codes (PRED. DAC code OPT). The spread of the predicted $Th_k^{optimal}$ at the optimal DAC codes is smaller than for the ones measured with the edge DAC codes. However, some pixels have $Th_k^{optimal} >> Th^{aim}$, which indicates that they are more noisy than others and should be masked.

Defining fair masking conditions is crucial for the detector's performance. A factor of ten difference in the number of masked pixels can be observed depending on how strict these conditions are. For the study in this thesis, the average spread of the $Th_k^{optimal}$ is used as a tolerance level for masking and is chosen to be 35 DAC counts. Pixels with $Th_k^{optimal} - Th^{aim} > 35$[DAC] are then masked. On average, 1 ‰ of pixels are masked in each VeloPix ASIC. Figure 5.10 shows in red the masked pixels for one VeloPix ASIC.

Figure 5.11 shows an example of representative thresholds taken with optimal DAC codes (see Fig. 5.9) and mask applied (see Fig. 5.10). If compared to Fig. 5.5(c), where a DAC code `0x0` is used, the representative thresholds are homogeneous. The only exception is the masked pixels that are shown in light blue in Fig. 5.11.

Both the mask and the optimal DAC codes were derived from the noise rate procedure in 2022. However, this procedure has a few things that could

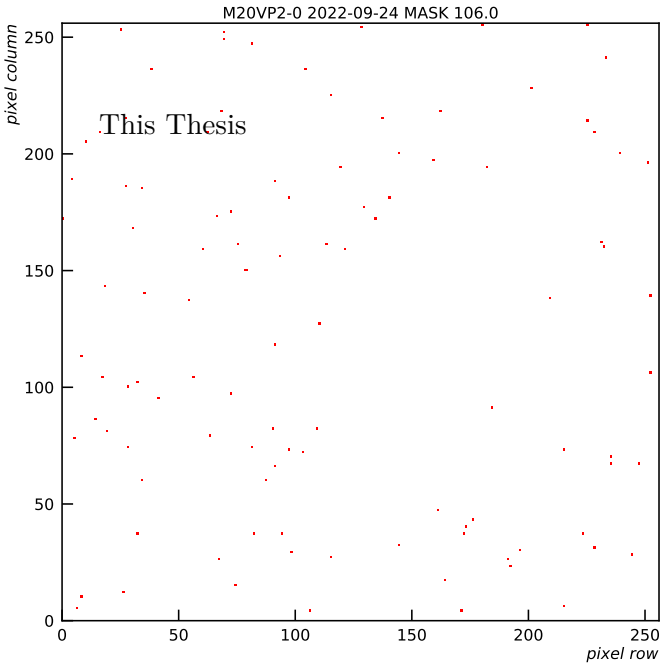


Figure 5.10: An example of a mask applied to one VeloPix ASIC, taken from the outlined noise rate procedure.

be improved. There are two other important issues to address besides the already discussed linear calibration curve (which can be interpolated more realistically with more points measured).

Firstly, as was mentioned above, the masking conditions are a very sensitive configuration parameter. The mask described here is rather unrefined, and no attempt was made to study the underlying causes for the need to mask pixels. Another problem is using the predicted $Th_k^{optimal}$ for masking. Retaking a scan of the threshold with optimal DAC codes applied would give a real $Th_k^{optimal}$[5]. The masking should then be done with the results of the new scan, which represents the pixel responses more realistically. It is especially crucial, as the predicted $Th_k^{optimal}$ used for masking here relies on the linear calibration curve, which is known to be an approximation.

Secondly, crosstalk between pixels is possible and can lead to observing an increased amount of noise in some pixels. To avoid counting crosstalk signals, neighboring pixels should be isolated from one another. This can be

[5]For future reference, this is known in the VELO jargon as the “control” scan.

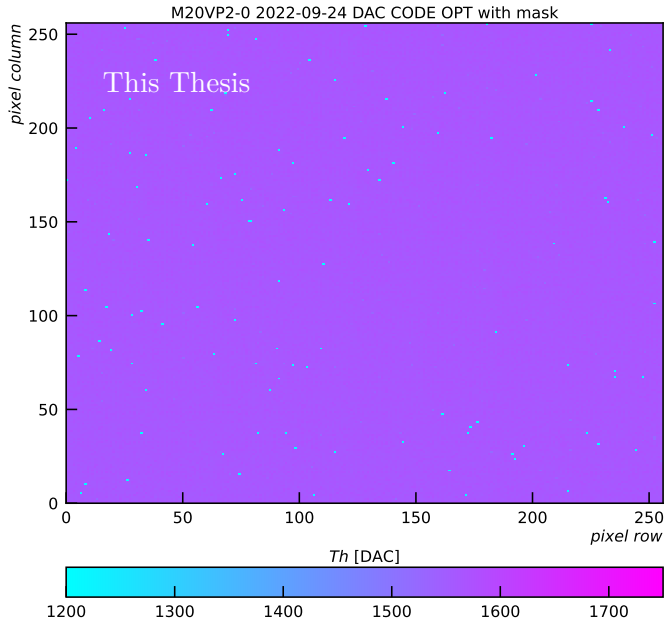


Figure 5.11: Representative thresholds for an equalized VeloPix ASIC with the mask applied. Obtained from the noise rate procedure.

achieved by masking individual pixels when taking the threshold scans. The pixels can then be divided into groups of four and masked sequentially one by one, as shown in Fig. 5.12.

In this way, the pixels to be measured are surrounded only by de-activated pixels and, therefore, isolated. The only downside of this approach is that four times as many scans are required. Practically, this means that there has to be sufficient no-beam time to perform all the scans.

5.2 Comparison of reconstructed objects at the different trigger stages

To profit from the increased luminosity of LHC, the LHCb trigger system had to be redesigned [113]. The hardware trigger stage `L0`, which reduced the rate from 30 MHz to 1.1 MHz in the original LHCb trigger system, was necessary because of the limited rate of readout electronics of the original detector. In the upgraded LHCb detector, the hardware trigger stage `L0` has been dropped in favor of a triggerless readout system. As discussed in the introduction to this chapter, this is possible because the readout electronics

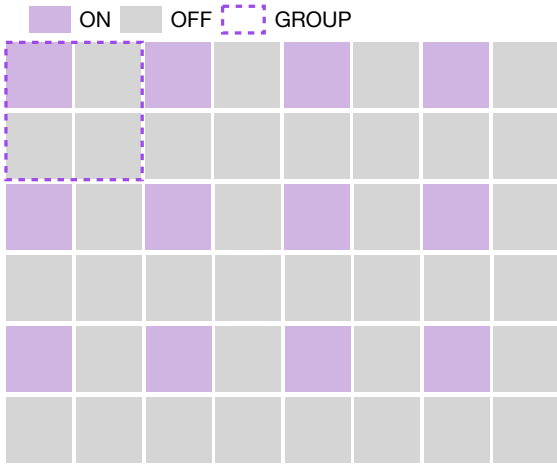


Figure 5.12: A scheme of the per-pixel masking to take scans of thresholds without cross-talk between pixels.

are replaced to deal with the higher rate. Figure 5.13 shows the schematic representation of the upgraded trigger system of LHCb [113, 114].

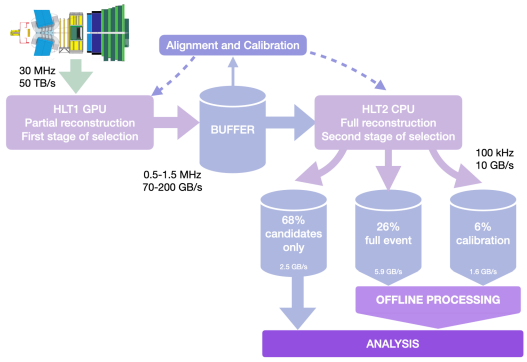


Figure 5.13: The upgraded trigger system of LHCb. Adapted from [115]. Values are taken from [113, 114].

Without the hardware trigger, the first software stage `HLT1` has to process the 30 MHz of the $p-p$ collision rate coming from the detector. There is only enough time to partially reconstruct the events and apply some loose basic selection criteria. These selection criteria reduce the rate to $0.5 - 1.5$ MHz. Selected events are placed into a disk buffer. It is at this point that the

real-time alignment and calibration are performed. The fastest parts of real-time alignment and calibration, such as the global spacial alignment of the tracking detectors, are done within minutes and can be propagated back to HLT1. But the full alignment and calibration are only available in the next stage of software trigger, HLT2. Depending on the performance and the schedule of the LHC collider, events can spend days or weeks in the buffer. HLT2 fully reconstructs the selected events. At this stage, more exclusive selection criteria are applied, similar to the Stripping selection in Sec. 4.2.2. The bandwidth is optimized in two ways: using the high-rate triggers, which only store the information directly related to the trigger, and using lower-rate triggers, where the larger part of the event is stored. The LHCb trigger system is flexible and allows to specify the exact amount of data to be saved per triggered event.

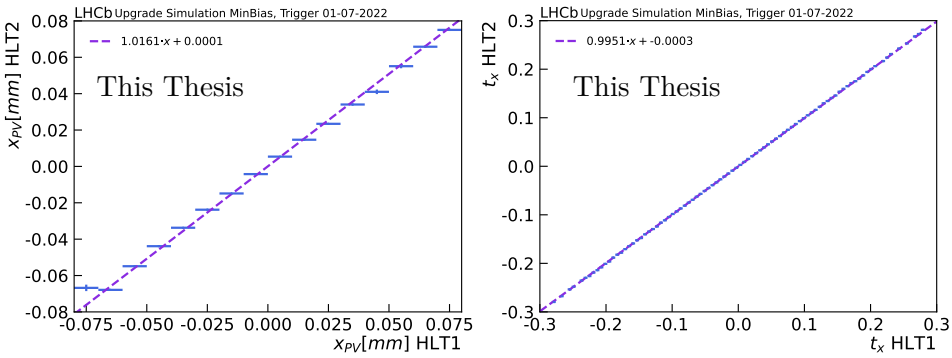
This results in a reduced data volume and consequently, more candidates can be saved. Only about 30 % of the bandwidth is reserved for the events that have all information stored. Both full events and calibration events go through offline processing, similar to Stripping in Sec. 3.2.7.

To perform the initial selection, HLT1 performs a partial reconstruction of tracks and primary vertices. Reconstruction of tracks and vertices is done with the same algorithm that is executed multiple times on different data. Moreover, track and vertices are reconstructed independently, i.e., a track does not need information from other tracks to be reconstructed. The graphic-processing units or GPUs, are built to perform basic operations on multiple input data streams in parallel and independently. A hybrid architecture that combines GPU and computing processing units or CPUs (more suitable for sequential execution of complex computations) was chosen to be used for HLT1 [116]. However, GPUs require the development of specialized software, which means that it has a dedicated framework, distinct from the one used for HLT2.

The HLT2 level, on the other hand, is based on CPUs and includes full track and vertex reconstruction, as well as reconstruction of the calorimeter and RICH objects.

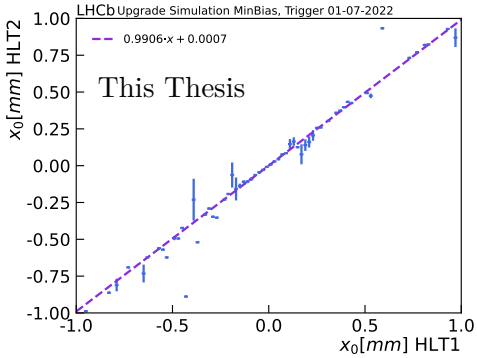
Both HLT1 and HLT2 select what looks like interesting events, but due to the different algorithms and architectures, potentially different decisions may be made for the same event. This is an undesirable behavior because if the event is rejected by HLT1 based on the partially reconstructed event, it should be rejected by HLT2 based on the fully reconstructed event. The selection decisions are made based on reconstructed variables. Monitoring possible deviations of these variables is, therefore, crucial, as non-trivial deviations are a source of concern and should be resolved. The

`TriggerObjectCompatibilityChecker` algorithm, which the author developed, is an algorithm that allows the comparison of the reconstructed objects for the same generated particles at `HLT1` and `HLT2` levels. Figure 5.14 shows examples of the track and vertex parameters comparison between `HLT1` and `HLT2` levels. All plots are created with minimum bias simulated samples, which contain multiple simulated interactions and decays. The 01.07.2022 version of `HLT1` and `HLT2` is used.



(a) Primary vertex x-position

(b) First track segment slope in x



(c) First track segment x-position

Figure 5.14: Profiles of track and vertex parameters at `HLT1` and `HLT2` trigger. The tracks are created by the same simulated particle.

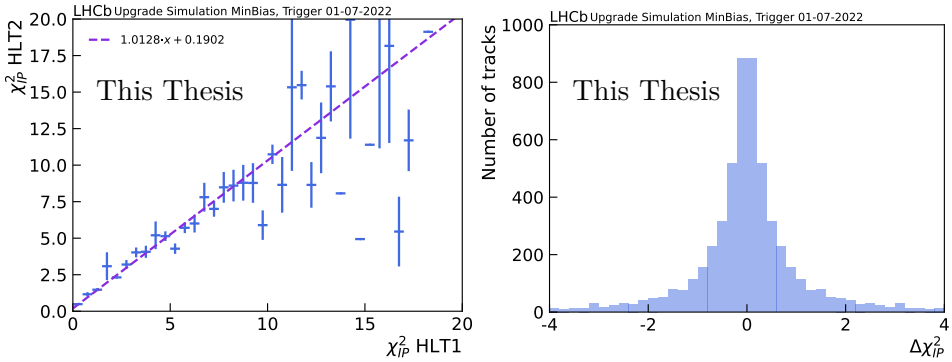
As was mentioned before, part of the requirements for selecting B_s^0 mesons is the detachment requirement, which requires a significant decay length of B_s^0 and the daughter tracks to be incompatible with the primary vertex (large χ^2_{IP}). Table 5.1 shows the number of the rejected events at the HLT1 and HLT2 for a set of typical HLT1 χ^2_{IP} selection criteria. Some events

are exclusively rejected by HLT1 or HLT2, which means that the HLT1 and HLT2 χ^2_{IP} are not equivalent. Notably, the HLT1 rejects more events than HLT2, which could be partially compensated by applying different χ^2_{IP} criteria at the HLT2 level. The current trigger developers work on improving this discrepancy, but ultimately, a 100% overlap between HLT1 and HLT2 might not be achievable.

$\chi^2_{\text{IP}} >$	At HLT1	At HLT2	At both HLT1 and HLT2
7.4	312	282	230
10.0	414	386	334
13.0	496	469	422

Table 5.1: Number of rejected events based on the typical HLT1 χ^2_{IP} selection criteria. Measured on 3000 events.

Figure 5.15 shows a comparison between χ^2_{IP} for the same tracks at `HLT1` and `HLT2` levels.



(a) Profile of reconstructed χ^2_{IP} (b) $\Delta\chi^2_{\text{IP}} = \chi^2_{\text{IP}}(HLT2) - \chi^2_{\text{IP}}(HLT1)$

Figure 5.15: The χ^2_{IP} of reconstructed tracks at `HLT1` and `HLT2` trigger. The tracks originate from the same generated particle.

The full list of monitored values can be found in the source code [117]. This monitoring tool is included in the central LHCb monitoring system and can be rerun for any chosen trigger software version and simulated sample. The hope is that it will be able to spot the problems in reconstruction before valuable data is lost.

6 The future of measuring $\mathcal{CP}$-violating phase ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays

The measured ϕ_s in Tab. 4.21 is compatible with the SM tree level prediction in Eq. 2.62 within half statistical σ. At the same time, it is also compatible with zero within two σ, and the estimated magnitude of λ is consistent with no direct $\mathcal{CP}$-violation. Therefore, no $\mathcal{CP}$-violation was observed in $B_s^0 \to J/\psi K^+K^-$ decays. Both values of ϕ_s and $|\lambda|$ are statistically limited. LHCb plans to collect about ten times more data than the currently available sample by 2032 and thirty times more than that by 2041[1]. This chapter omits discussing the LHCb upgrade necessary to achieve these goals. Instead, it concentrates on analysis-related issues, assuming the projected integrated luminosity can be attained.

There are four main points to consider in the future measurements of ϕ_s with $B_s^0 \to J/\psi K^+K^-$:

1. Flavour tagging becomes increasingly difficult in higher pile-up conditions. As a result, the performance of the standard algorithms will degrade, reducing the effective statistical power of the collected sample.

2. With a higher pile-up, it becomes harder to associate primary and decay vertices correctly, which impacts the determination of the decay time. As a consequence, the decay time resolution is expected to become less Gaussian due to the increased chance of vertex mis-association.

3. With 30 times more data by 2041 than collected by LHCb now, the expected statistical uncertainty of ϕ_s, assuming the same detector performance, will improve by a factor of five or so. The penguin contributions to ϕ_s are currently estimated to be $0.003^{+0.0010}_{-0.0012}$ rad [6],

[1] Both dates are taken from the LHC long-term schedule as of April 2023.

which is of approximately the same size as the expected uncertainty on ϕ_s. Therefore, keeping the penguin contributions under control is more important than ever.

4. The measured $\Gamma_s - \Gamma_d$, see Tab. 4.21, has the same size statistical and systematical uncertainties. The dominant systematics comes from the correlations between $m(J/\psi K^+K^-)$ and the decay time and angular observables used in the final fit. The reduction of this effect is necessary not only for the determination of $\Gamma_s - \Gamma_d$, but also for ϕ_s, for which the systematical uncertainty is also dominated by these correlations.

Future of flavour tagging

Section 4.4 introduces a key component of measuring ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays - tagging of the initial B_s^0 meson flavor. The tagging power $\varepsilon_{tag}D_{tag}^2$ together with the sample size determines the statistical power of the ϕ_s measurement. The total tagging power achieved in the current ϕ_s measurement is about 4.2%, see Tab. 4.8, which means that only about 4.2% of the candidates collected effectively contribute to the measured ϕ_s value. Loss of flavour tagging performance thus negates the benefit of increased sample size. Therefore, developing the flavour tagging algorithms for future data-taking is crucial for ϕ_s analysis.

As described in Sec. 4.4, the flavour tagging relies on being able to associate the signal b-quark b_{sig} and the tagging b-quark b_{tag}. The flavour tagging performance degrades with a higher number of pile-up vertices and track multiplicity [34], [118]. Several studies of flavour tagging algorithms are ongoing in preparation for the data-taking conditions described in Chapter 5 to ensure performance at least as good as Tab. 4.8, and hopefully even better.

Wrong association of primary and decay vertices

Because of the higher pile-up in the upcoming data-taking years and more visible p-p collisions, see Chapter 5, naively more wrong primary vertices associations are expected. These wrongly associated primary vertices contribute to the non-Gaussian tails of the decay time resolution distribution, see Fig. 4.43 and increase the sample's overall decay time resolution dilution. Even worse, the non-Gaussian contribution to the decay time resolution can introduce biases in estimated lifetimes; see **Lifetime bias from decay time resolution** in Sec. 4.7, if not accounted for correctly. This bias is a

source of greater concern as it is an uncovered "mistake" in the estimated parameters of interest.

As the relative contribution of the non-Gaussian effect increases and the average effective resolution no longer represents the resolution experienced by the majority of the candidates in the sample, the single Gaussian approximation of the resolution function in the final fit breaks down. To continue using the single Gaussian resolution function approximation, the candidates that have a larger probability of being reconstructed mis-associated vertex can be filtered out. As illustrated in Fig. 4.44, this can be done by selecting B_s^0 with small pseudorapidity η. The B_s^0 mesons with a bigger angle to the z-axis (smaller η) have a lower probability of primary vertex disassociation. Moreover, because these B_s^0 will also have bigger p_{T}, the opposite-side flavour tagging is expected to perform better. Another way is to apply the isolation criteria on selected primary vertices, which will reduce the chances of this vertex being confused with another one. However, applying vertex isolation cuts introduces bias in decay time acceptance and, therefore, should be avoided. Therefore, in the future, it is strongly recommended to use the negative decay time tail of the control sample (inclusive $J/\psi K^+K^-$) as an exact proxy for decay time resolution distribution and convolute it with the full PDF from Eq. 4.13 directly. The long-lived contribution must then be statistically subtracted from the control sample for this method to work. Moreover, the exact proxy of the resolution function accounts for the uncovered lifetime bias from the non-Gaussian tails of the decay time resolution correctly.

ϕ_s statistical uncertainty in the future

Figure 6.1 summarizes the current and expected *statistical uncertainty* of ϕ_s with respect to the collected integrated luminosity $\mathcal{L}$. Only the latest LHCb measurements are shown: LHCb 15-16 [49] and LHCb 15-18: this thesis and [63]. For completion, the latest measurements of ϕ_s from $B_s^0 \to J/\psi K^+K^-$ decays of ATLAS (ATLAS 15-17, [59]) and CMS (CMS 17-18 [61]) are reported. The HFLAV 2021 [62] world average for ϕ_s measured with $B_s^0 \to J/\psi K^+K^-$ decays, which does not yet include results of this thesis and [63], is shown with a line. The precision of SM tree-level only prediction for ϕ_s from [14] is shown for comparison. The expected statistical uncertainty is projected for LHCb Run 4 (50 fb^{-1}, [34]), LHCb Run 6 (300 fb^{-1}, [34]), CMS Run 6 (3000 fb^{-1}, [120]) and ATLAS Run 6 (3000 fb^{-1}, [119]) data samples. Both Run 4 and Run 6 LHCb uncertainty projections assume the detector performance of 2010 - 2018 and only show a

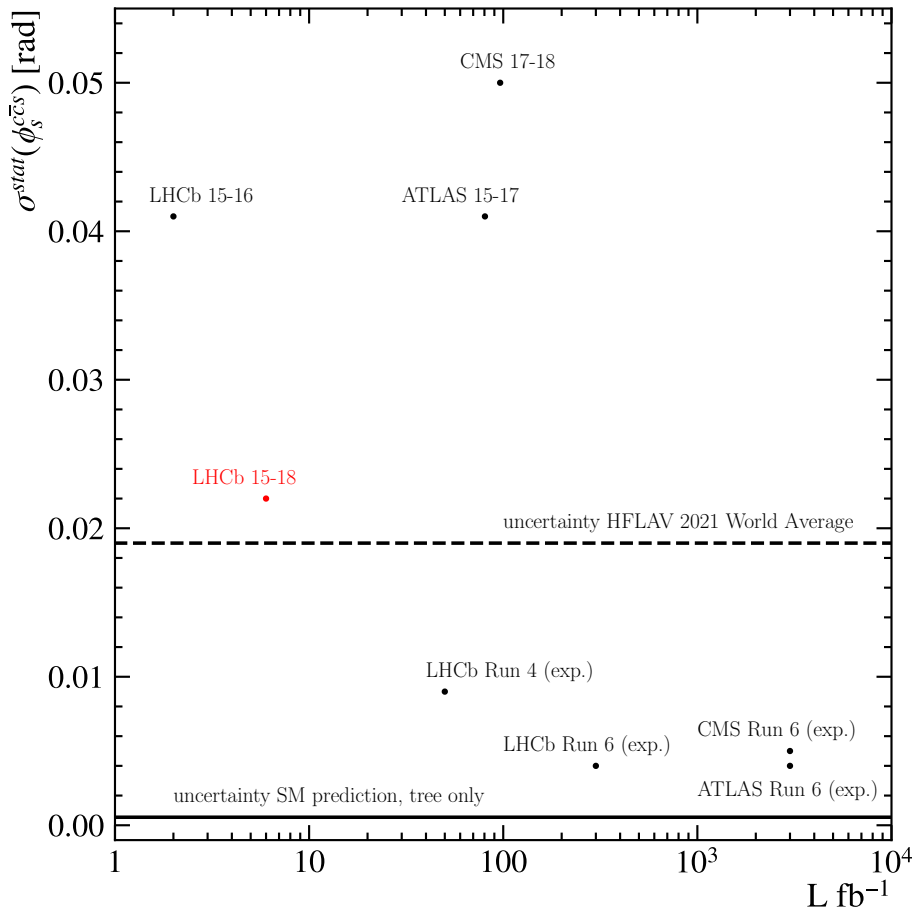


Figure 6.1: The projected statistical uncertainty on the ϕ_s: LHCb 15-16 [49]; LHCb 15-18 is this thesis and [63]; CMS 17-18 [61]; ATLAS 15-17 [59]; LHCb Run 4 and Run 6 projections from [34]; ATLAS Run 6 from [119]; CMS Run 6 from [120]; SM prediction with tree level decay only is from 2021 update of [14].

relative gain in statistics. However, in principle, some improvements in the mass and decay time resolutions may be expected. The biggest limitation in the future measurement as was described above will most likely come from a limited flavour tagging power. Therefore, effort should be put into developing flavour tagging algorithms.

The ATLAS Run 6 projection assumes an improvement in both decay time and mass resolutions; 1.5% flavour tagging power; loose muon trigger with

$p_{\mathrm{T}}(\mu^{\pm}) > 6$ GeV/c. The CMS Run 6 projection also assumes improvements in decay time and mass resolutions and about 2% flavour tagging power, keeping selection close to [121]. For both experiments, the biggest limitation comes from the trigger. B-mesons are predominantly produced at low-p_{T} values, whereas ATLAS and CMS programs concentrate on high-p_{T} physics mainly. Fortunately, dedicated trigger developments in CMS have opened a new path for triggering low-p_{T} physics in these experiments. In the new CMS trigger, the p_{T} cuts are getting reduced to allow for more low-p_{T} physics to be saved, as luminosity drops and more space in the trigger rate is available. These events are not immediately processed, but only when the resources are available offline. This is known as data "parking". The dedicated B-physics trigger in CMS is known as B-parking and is described in [122].

Penguin contributions

The penguin contribution to ϕ_s, $\Delta\phi_s^{penguin}$, is suppressed by a factor 0.05 in $B_s^0 \to J/\psi K^+K^-$ decays, see Chapter 2. Barel et al [6] estimate it to be $\Delta\phi_s^{penguin} = 0.003^{+0.0010}_{-0.0012}$ rad. The central value is of the size of expected statistical uncertainty on ϕ_s in LHCb Run 6, see Fig. 6.1, and has uncertainty about twice larger than the tree-level SM prediction from [14], see [6]. To be able to disentangle possible beyond the SM effects from the measured ϕ_s, see Sec. 2.5, the estimation of penguin-level contribution has to be improved. Luckily, the penguin contribution can be constrained from $B^0 \to J/\psi \rho^0$ and $B_s^0 \to J/\psi \overline{K}^{*0}$ [123]. LHCb has determined ranges for penguin contributions with $B^0 \to J/\psi \rho^0$ decays [124] and $B_s^0 \to J/\psi \overline{K}^{*0}$ [125] using 3 fb^{-1} dataset. An update with a larger dataset for both decays is needed.

Controling systematics

The biggest systematical effect for both ϕ_s and lifetimes comes from ignoring the correlations between $m(J/\psi K^+K^-)$ and final observables in background subtraction with *sPlot* method, see Tab. 4.17. Unfortunately, this systematics puts a limitation on the applicability of using *sPlot* method for background subtraction in the future when the systematical uncertainty becomes more important. As mentioned in Sec. 4.3.3, Custom Orthogonal Weights [77], can accommodate such correlations without introducing a bias on the weights. Using Custom Orthogonal Weights instead of sWeights will require a minimum change to the analysis strategy and, therefore, should be the first thing to try. However, it needs further studies and the unbiased nature of Custom Orthogonal Weights should be confirmed by comparing the results of fitting

the signal-only PDF to the background subtracted dataset with fitting a model, which includes correlations, to the unmodified dataset.

If in the future it is impossible to factorize mass from the final PDF, the non-factorized PDF should be fitted for both the signal and background.

Other considerations

It is important to mention, that this thesis and [63] measures not just the $\mathcal{CP}$-violating parameters with the world's best precision, but also the B_s^0 lifetimes: $\Gamma_s - \Gamma_d$ and $\Delta\Gamma_s$. The best measurement of lifetimes is not the main goal of this analysis, but rather a welcome side-effect to the measurement ϕ_s and $|\lambda|$. Unfortunately, though, often the analysis decisions that benefit ϕ_s, like isolating vertices to reduce the decay time resolution dilution of the sample, are not optimal for the lifetime measurements. Controversially, it seems that in the future it will be increasingly challenging to keep the precision level for both ϕ_s and lifetimes with the same measurement. Therefore, it is possible that the measurement will have to be split based on ϕ_s-optimized and lifetimes-optimized selections.

Measuring ϕ_s is one of the main tasks for flavour physics of this decade. And with every measurement, the observation of a non-zero value for ϕ_s will be closer and closer within our reach. And hopefully, a deviation from the SM prediction will be observed. So the best of luck to those, who will measure ϕ_s in the future, and huge gratitude to all of those, who measured it before and laid the ground for this measurement.

A Appendix

A.1 Two dimensional mass fit

A.1.1 Signal mass uncertainty shape

The signal mass uncertainty shape is taken from the simulated $B_s^0 \to J/\psi\, K^+K^-$ sample. The mass uncertainty shape in the simulated $B_s^0 \to J/\psi\, K^+K^-$ sample and *sPlot* weighted data sample for 2018 taking is shown in Fig. A.1. The simulated and measured mass uncertainty differ by a scale, which is seen in Fig. A.1 as a shift between two distributions.

The signal mass uncertainty distribution is modeled as a Bukin probability distribution function, defined in [126]. Bukin distribution function is based on the convolution of Gaussian and exponential distributions and can be used to describe asymmetric peaks. It is chosen as it has a relatively small number of parameters but accommodates asymmetric shape. In Fig. A.2 Bukin distribution function is fitted to the mass uncertainty distribution in $B_s^0 \to J/\psi\, K^+K^-$ simulation for 2018 in six bins of K^+K^- mass.

A.1.2 Background mass uncertainty shape

Deriving combinatorial background mass uncertainty shape is a non-trivial task. If $B_s^0 \to J/\psi\, K^+K^-$ is treated as a two-body decay of B_s^0 to J/ψ and Φ, then B_s^0 mass uncertainty is derived as:

$$\sigma_m = \sqrt{\left(\frac{\Delta p_{J/\psi}}{p_{J/\psi}}\right)^2 + \left(\frac{\Delta p_{\phi(1020)}}{p_{\phi(1020)}}\right)^2 + \frac{\Delta\theta}{(1 - cos(\theta_{J/\psi\,,\phi(1020)}))}} \cdot \frac{m^2_{B_s^0} - m^2_{J/\psi} - m_{\phi(1020)^2}}{2m_{B_s^0}} \tag{A.1}$$

where p is a momentum of a particle; Δp is the momentum uncertainty; θ is an angle between two particles; $\Delta\theta$ is angle's uncertainty; m is a particle

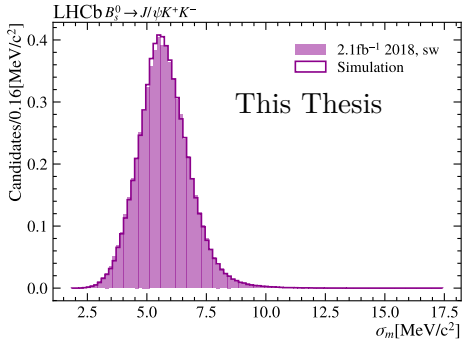


Figure A.1: Per event mass uncertainty distribution for the simulated $B_s^0 \to J/\psi K^+K^-$ sample compared to the *sPlot* weighted data distribution for 2018. Ghost tracks are excluded from the simulation. Simulation is kinematically reweighted to the data sample with weights for BDT training.

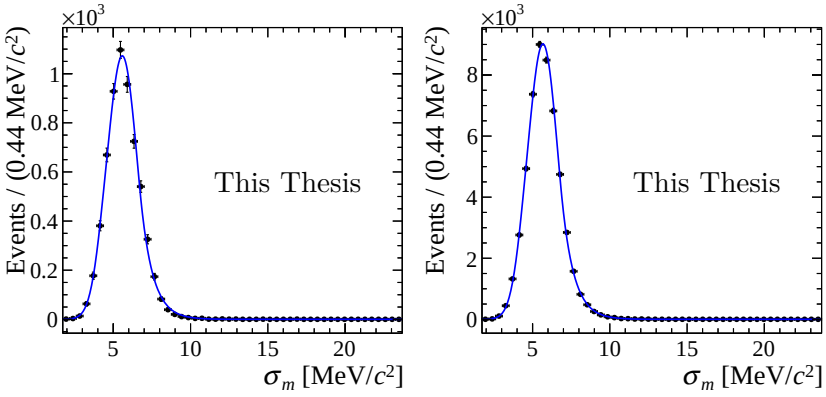
mass.

To derive the combinatorial background mass uncertainty shape, similar to [76], the transverse momentum p_T and pseudorapidity η distributions as randomly generated for J/ψ and $\phi(1020)$. The transverse momentum follows an exponential distribution with expected p_T set to the average p_T data. The pseudorapidity is uniformly distributed between $(2, 5)$. The angle between J/ψ and $\phi(1020)$ is generated uniformly between $(-\pi, \pi)$. Both J/ψ and $\phi(1020)$ masses are taken from [70]: $m_{J/\psi} = 3096.9$ MeV/c^2 and $m_{\phi(1020)} = 1019.461\,\text{MeV}/c^2$. mass of B_s^0 is then computed as,

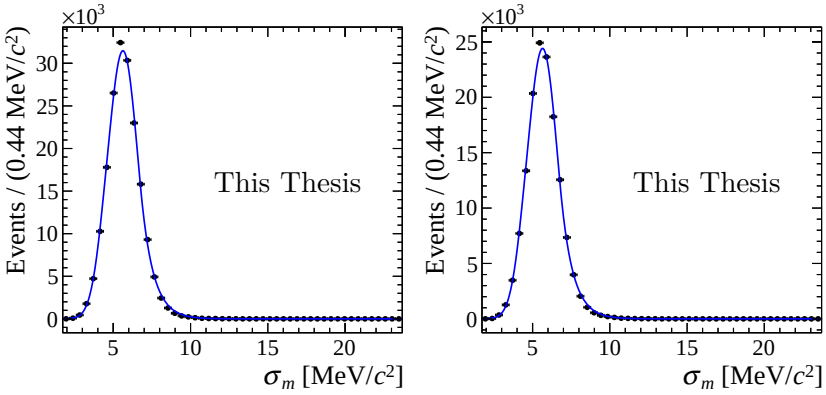
$$m_{B_s^0} = \sqrt{\left(\sqrt{p^2_{J/\psi} + m^2_{J/\psi}} + \sqrt{p^2_{\phi(1020)} + m^2_{\phi(1020)}}\right)^2 - \left(p^2_{J/\psi} + p^2_{\phi(1020)} + 2p^2_{J/\psi}\cdot^2_{\phi(1020)} \cdot cos(\theta_{J/\psi\,,\phi(1020)})\right)} \tag{A.2}$$

where $p = \eta \cosh p_T$. In Fig. A.3, the generated distributions are shown. The mass uncertainty shape in Fig. A.3 is fitted with the Bukin distribution. In data, the sum of the Bukin distribution and Gaussian distribution is used as a combinatorial background model.

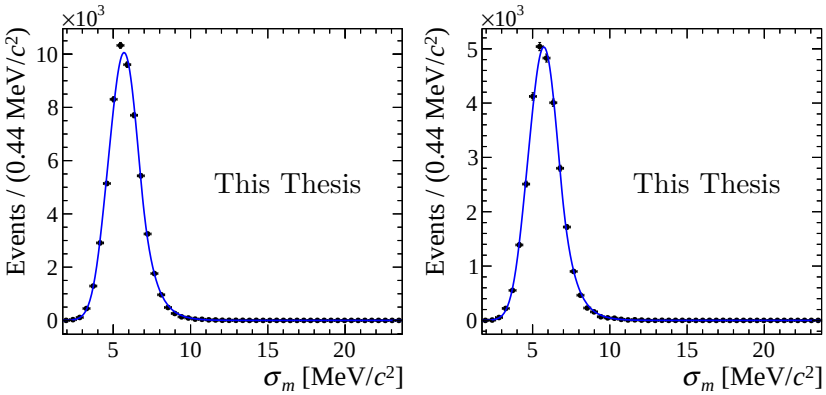
The $B^0 \to J/\psi K^+K^-$ background is expected to have the same shape as $B_s^0 \to J/\psi K^+K^-$ with a slightly different mean, as the mass of B^0 is lower than of B_s^0. However, to improve fit stability, the mean of the Bukin distribution for $B^0 \to J/\psi K^+K^-$ is fixed to be shifted with respect to the mean of the Bukin distribution for $B_s^0 \to J/\psi K^+K^-$ signal by a difference computed using Eq. A.1.



(a) $m(K^+K^-) \in [990, 1008)\,\mathrm{MeV}/c^2$ (b) $m(K^+K^-) \in [1008, 1016)\,\mathrm{MeV}/c^2$

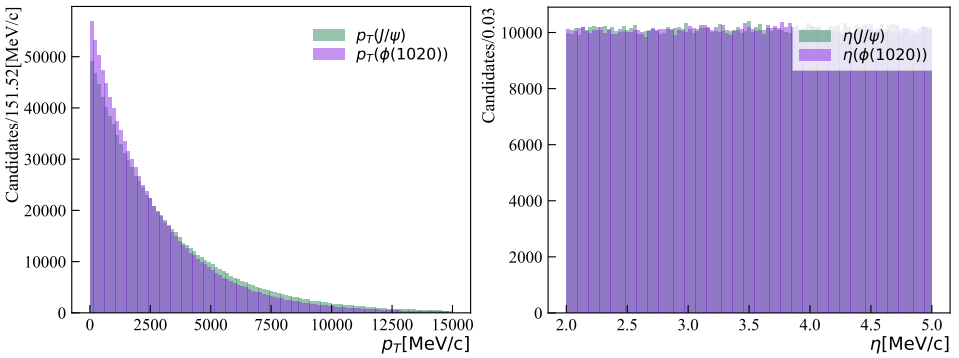


(c) $m(K^+K^-) \in [1016, 1020)\,\mathrm{MeV}/c^2$ (d) $m(K^+K^-) \in [1020, 1024)\,\mathrm{MeV}/c^2$



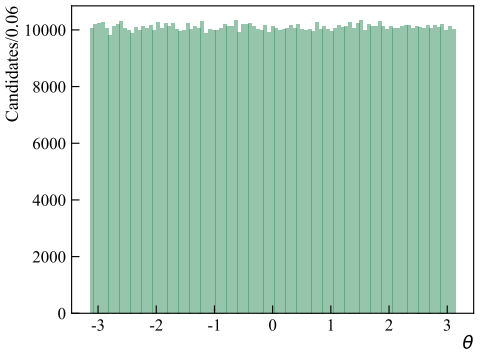
(e) $m(K^+K^-) \in [1024, 1032)\,\mathrm{MeV}/c^2$ (f) $m(K^+K^-) \in (1032, 1050)\,\mathrm{MeV}/c^2$

Figure A.2: 2018 $B_s^0 \to J/\psi\, K^+K^-$ simulation mass uncertainty distribution fitted with Bukin particle distribution function. Ghost tracks are excluded.

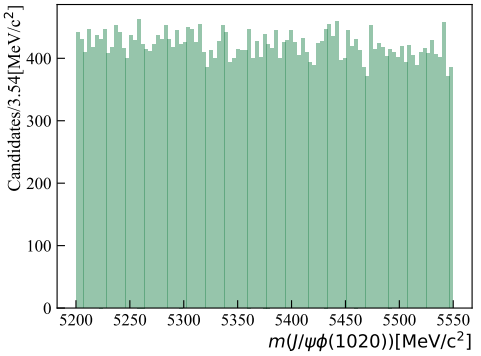


(a) p_T

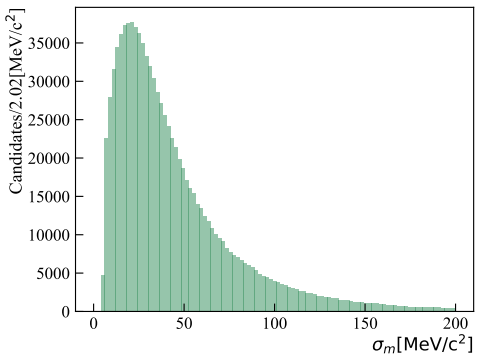
(b) η



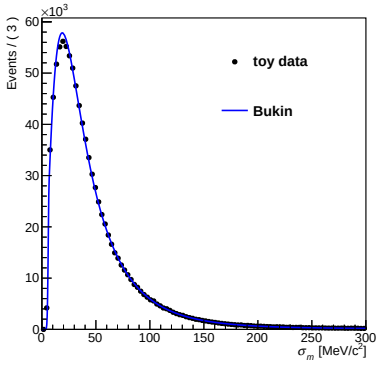
(c) θ



(d) $m_{B_s^0}$



(e) σ_m



(f) σ_m fit

Figure A.3: Generated distributions for J/ψ and $\phi(1020)$ transverse momentum p_T; pseudorapidity η and angles between them θ. B_s^0 mass $m_{B_s^0}$ and it's uncertainty σ_m computed from the generated samples. The last plot (f) shows the fit of the distribution obtained in (e). In analysis, events with a mass uncertainty above 50 MeV/c^2 are excluded.

A.1.3 Two-dimensional mass fit procedure

The two-dimensional probability distribution function is defined as,

$$\begin{aligned} p(m,\sigma_m) = & N_{sig} f(m;\mu,s_1,s_2,\alpha_L,n_L,\alpha_R,n_R|\sigma_m) f_{Bukin}(\sigma_m;\mu_\sigma,\sigma_\sigma,\xi,\rho_L,\rho_R) \\ & + N_{comb} e^{-\gamma m}\left[f_{Bukin}(\sigma_m;\mu_\sigma^{comb,1},\sigma_\sigma^{comb,1},\xi^{comb},\rho_L^{comb},\rho_R^{comb}) + N(\sigma_m;\mu_\sigma^{comb,2},\sigma_\sigma^{comb,2})\right] \\ & + N_{B^0} f(m;\mu-\Delta\mu_{B^0},s_1,s_2,\alpha_L,n_L,\alpha_R,n_R|\sigma_m) f_{Bukin}(\sigma_m;\mu_\sigma-\Delta\mu_{\sigma,B^0},\sigma_\sigma,\xi,\rho_L,\rho_R) \end{aligned} \quad \text{(A.3)}$$

The two-dimensional fit procedure is divided into four steps:

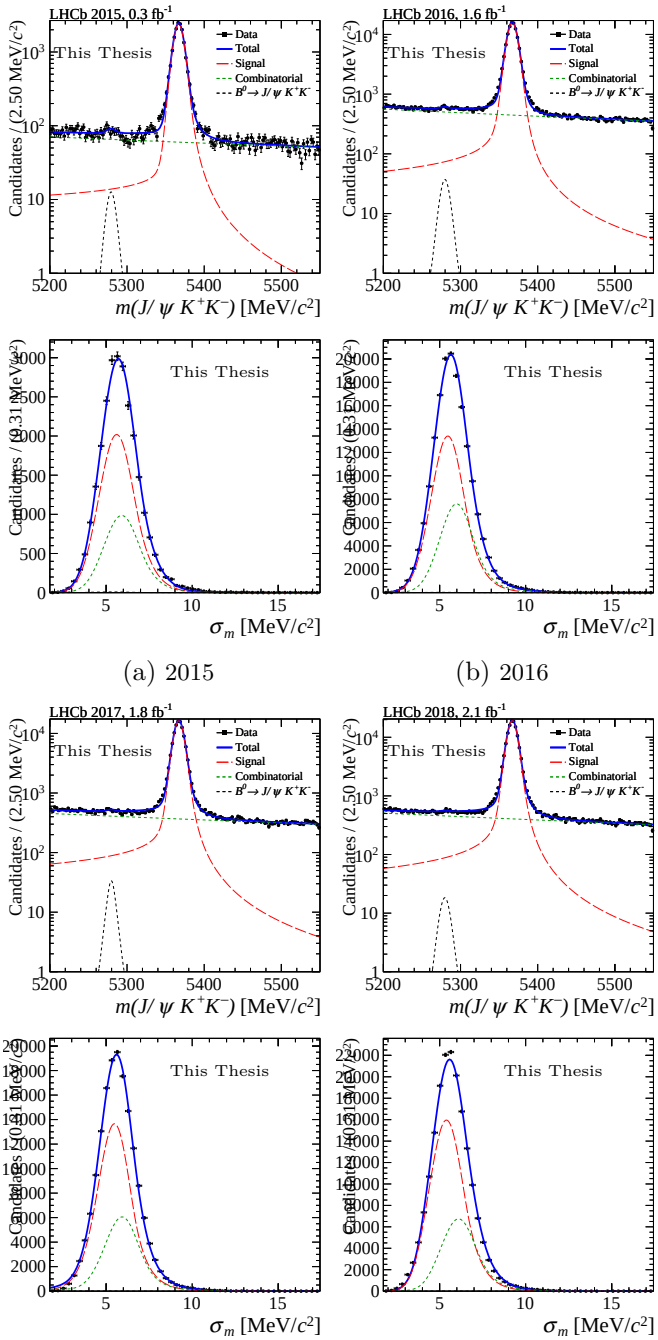
1. One-dimensional fits of mass and mass uncertainty distributions of the simulated $B_s^0 \to J/\psi K^+K^-$ sample with ghost tracks removed.

2. Two-dimensional fit of the simulated $B_s^0 \to J/\psi K^+K^-$ sample with ghost tracks removed. The tail in signal mass distribution and mass uncertainty parameters ξ, ρ_L, ρ_R are fixed from the one-dimensional fits in step 1.

3. One-dimensional fits of mass and mass uncertainty distributions of the collected $B_s^0 \to J/\psi K^+K^-$ data. The signal mass shape is fixed from the two-dimensional simulation fit, except for the μ and s_1, s_2 of DCB. The mass uncertainty signal shape is also fixed to the results of the two-dimensional simulation fit, except for the μ_σ. In the one-dimensional mass uncertainty fit, the species yields are fixed to the results of the mass one-dimensional fit.

4. Two-dimensional fit of the $B_s^0 \to J/\psi K^+K^-$ data. All the background shapes are fixed from the one-dimensional fits of step 3. The signal shape is fixed in the same way, except for the parameter of the DCB width s_1 that is left floating.

The fit is performed in six bins of K^+K^- mass and each year separately. In Fig. A.4, combined results of the two-dimensional fit are shown for all the data taking years. The main obstacle in using this fit as a baseline lies in the uncertainty of the true combinatorial mass uncertainty shape. In particular, it is very correlated to the signal shape and vice versa. To disentangle those two shapes, the one-dimensional mass uncertainty fit is constrained in yields from the one-dimensional mass fit, which is not a satisfactory solution. Extra studies are necessary to understand the combinatorial mass uncertainty shape distribution.



(a) 2015 (b) 2016

(c) 2017 (d) 2018

Figure A.4: Projections of two-dimensional mass fits on mass and mass uncertainty in bins for $B_s^0 \to J/\psi K^+K^-$ data sample collected in 2015-2018.

A.2 Parameters used to generate the example decay time and angular dependent decay rate

Table A.1 shows the parameter value used to generate an example pdf for Fig. 4.19-4.22. Fig. A.5 and Fig. A.6 show the applied decay time resolution and decay time acceptance, respectively. In Sec. A.2.1 angular acceptance is described.

Parameter	Value
ϕ_s	-0.039 rad
λ	1
Γ_s	0.65139
$\Delta\Gamma_s$	0.268
Δm_s	17.74 ps^{-1}
$\lvert A_0\rvert^2$	0.5172
$\lvert A_\perp\rvert^2$	0.247
$\lvert A_S\rvert^2$	0.
δ_0	0.
$\delta_\perp$	2.9041
$\delta_\parallel$	3.1422
δ_S	0.

Table A.1: Parameters an example decay time and angular dependent decay rate. The following normalisation is used $|A_0|^2 + |A_\perp|^2 + |A_\parallel|^2 = 1$.

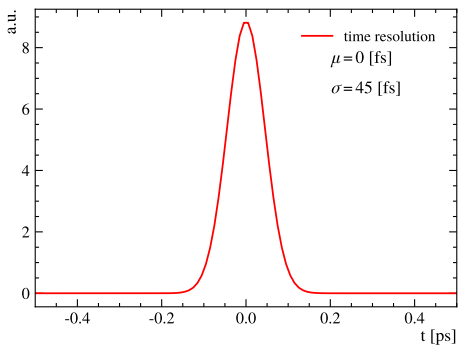


Figure A.5: Decay time resolution used for an example decay time and angular dependent decay rate.

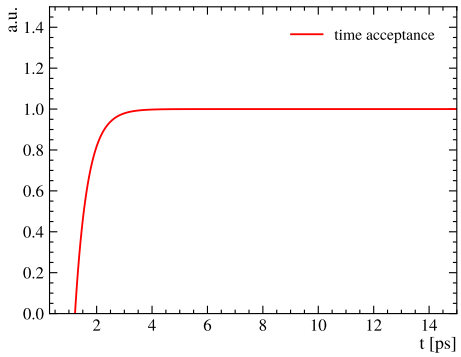


Figure A.6: Decay time acceptance used for an example decay time and angular dependent decay rate. It is generated from the function $f(t) = (1 - e^{-\beta t})$ with $\beta = 2$ ps^{-1}.

A.2.1 Description of the toy angular acceptance

Angular acceptance is implemented as a modification to the relevant angular terms in the distribution. The example acceptance is described as a combination of Legendre polynomials $P_i(\cos\theta_K)$ and real-values of spherical harmonics $Y_{ml}(\theta_\mu, \phi)$ as

$$\varepsilon(\Omega) = \sum_{i,l,m} c^i_{lm} P_i(\cos\theta_K) Y_{ml}(\theta_\mu, \phi) \tag{A.4}$$

. The description is simplified. For the $\cos\theta_\mu$, $\cos\theta_K$ and ϕ distributions the important terms in the description are $(0,0,0)$, $(2,0,0)$, $(0,2,0)$, $(2,2,0)$, $(2,2,0)$. For $\cos\theta_K$ additionally $(1,0,0)$, $(1,2,0)$, $(1,2,2)$ are needed. For ϕ additionally $(0,2,-1)$ and $(2,2,-1)$ are added.

The set of c^i_{lm} constant is chosen randomly for illustrative purposes as $c^0_{00} = 1$, $c^2_{00} = 1$, $c^0_{20} = 0.7$, $c^2_{20} = 1$, $c^0_{22} = 0.5$, $c^2_{22} = 0.5$, $c^1_{00} = 1$, $c^1_{20} = 1$, $c^1_{22} = 1$, $c^0_{2,-1} = 1$, $c^2_{2,-1} = 1$.

A.3 Bias in maximum likelihood estimator from limited decay time range

An untagged decay time distribution of $B^0_s \to J/\psi K^+K^-$ without any detector effects and assuming $\Delta\Gamma_s = 0$ ps is a simple exponent,

$$\frac{d\Gamma}{dt} \propto \frac{e^{-\Gamma t}}{\int_a^b dt e^{-\Gamma t}}, \tag{A.5}$$

where normalization is done over the analysis decay time range $t \in [a, b)$ ps. The maximum likelihood estimator for a lifetime is then,

$$\frac{1}{\Gamma} = \frac{1}{N}\sum_c t_c - a + \frac{(b-a)e^{-(b-a)\Gamma}}{1-e^{-(b-a)\Gamma}}, \tag{A.6}$$

where c denotes a $B_s^0 \to J/\psi K^+K^-$ candidate index. After substituting $a = 0.3\,\text{ps}$, $b = 15\,\text{ps}$ and $1/\Gamma_{B_s^0} = 1.52\,\text{ps}$, Eq. A.6 becomes,

$$\frac{1}{\Gamma} = \frac{1}{N}\sum_c t_c - 0.3\,\text{ps} + 0.93\,\text{fs}, \tag{A.7}$$

where the first is the lower side of decay time range and the second is a correction due to the finite decay time range. Both are included in the PDF normalization.

A.4 Correlation matrix of the baseline result.

	$\phi_s{}^0$	$\lvert\lambda^0\rvert$	$\Gamma_s - \Gamma_d$	$\Delta\Gamma_s$	Δm_s	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel - \delta_0$	$\delta_\perp - \delta_0$
$\phi_s{}^0$	1.00	0.10	-0.01	-0.02	0.03	0.00	-0.00	0.01	0.02
$\lvert\lambda^0\rvert$		1.00	0.01	-0.02	-0.12	0.01	-0.01	-0.02	-0.12
$\Gamma_s - \Gamma_d$			1.00	-0.46	-0.00	0.40	-0.31	-0.02	-0.01
$\Delta\Gamma_s$				1.00	0.01	-0.70	0.63	0.01	0.01
Δm_s					1.00	-0.01	0.01	0.01	0.74
$\lvert A_\perp\rvert^2$						1.00	-0.61	-0.06	-0.03
$\lvert A_0\rvert^2$							1.00	-0.00	0.00
$\delta_\parallel - \delta_0$								1.00	0.33
$\delta_\perp - \delta_0$									1.00

Table A.2: Correlation matrix of the results in Tab. 4.21.

A.5 Measured asymmetry

The decay time-dependent CP-asymmetry of $B_s^0 \to J/\psi K^+K^-$ and $\overline{B}_s^0 \to J/\psi K^+K^-$ is defined as,

$$A_{CP}(t) = \frac{\Gamma(\overline{B}_s^0 \to J/\psi K^+K^-) - \Gamma(B_s^0 \to J/\psi K^+K^-)}{\Gamma(\overline{B}_s^0 \to J/\psi K^+K^-) + \Gamma(B_s^0 \to J/\psi K^+K^-)}, \tag{A.8}$$

where $\Gamma(\overline{B}_s^0 \to J/\psi K^+K^-)$ and $\Gamma(B_s^0 \to J/\psi K^+K^-)$ are the decay rates of $\overline{B}_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi K^+K^-$ respectively. Using the $B_s^0 \to$

$J/\psi K^+K^-$ decay rate, as defined in Eq. 2.44 and its counterpart for $\overline{B}^0_s \to J/\psi K^+K^-$, the time-dependent asymmetry becomes,

$$A_{CP}(t) = \frac{-C_f \cos(\Delta m_s t) + S_f \sin(\Delta m_s)}{\cosh(\Delta\Gamma_s t/2) + D_f \sinh(\Delta\Gamma_s t/2)}, \tag{A.9}$$

where $C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}$ is the amount of direct CP-asymmetry, whose value depends on the CP-eigenstate f (or the de-facto polarization); $S_f = \frac{2\Im\lambda_f}{1+|\lambda_f|^2}$ is the amount of mixing induced $\mathcal{CP}$-violation and $\Im\lambda_f = \sin(\phi_s)$; $D_f = -\frac{2\Re\lambda_f}{1+|\lambda_f|^2}$ is not a $\mathcal{CP}$-violating part, where $\Re\lambda_f = \cos(\phi_s)$. Assuming no polarization dependence, $\Delta\Gamma_s = 0$ and $|\lambda| = 1$, the time-dependent CP-asymmetry is simply $\sin(\phi_s)\sin(\Delta m_s t)$. As ϕ_s are rather small, see Tab. 4.21, $\sin(\phi_s)$ can be approximated by ϕ_s, which means that ϕ_s is the amplitude of $A_{CP}(t)$.

Equation A.8 describes the analytical asymmetry. To plot the observed asymmetry, an ML estimator of the asymmetry given a time bin is defined as,

$$\hat{A}_{CP} = \frac{\sum_i w_i \mathcal{D}_i}{\sum_i w_i^2 \mathcal{D}_i^2}, \tag{A.10}$$

where w_i is per-candidate sWeight and D_i is it's dilution. The uncertainty of $\hat{A}_{CP}$ is,

$$\sigma = \sqrt{\frac{\sum_i w_i^2 \mathcal{D}_i^2}{(\sum_i w_i \mathcal{D}_i^2)^2}}. \tag{A.11}$$

The dilution $\mathcal{D}_i$ can be broken down into the individual components,

$$\mathcal{D}_i = q \times D_{tag} \times D_{tr} \times D_{CP}, \tag{A.12}$$

where q is flavour tagging tag; D_{tag} is a flavour tagging dilution factor, described in Sec. 4.4; D_{tr} is the decay time resolution dilution factor, described in Sec. 4.5.1 and D_{CP} describes the dilution due to the combination of amplitudes with different $\mathcal{CP}$-eigenstate.

The flavour tagging dilution comes from the non-zero mistag rate ω and, for a single tagger, is $1 - 2\omega$. However, there are two taggers used in the analysis: `OS` tagger (combined) and `SS` tagger (`SSKaon`); see Sec. 4.4. For some candidates, both taggers are present, and therefore a combined dilution is computed. For a candidate with two tag decisions q^{OS} and q^{SS},

the combined flavour tagging dilution D_{tag} is,

$$q \cdot D_{B_s^0} = \frac{q^{OS} \cdot D_{tag}^{OS} + q^{SS} \cdot D_{tag}^{SS}}{1 + q^{OS} \cdot D_{tag}^{OS} \cdot q^{SS} \cdot D_{tag}^{SS}}, \tag{A.13}$$

$$q \cdot \overline{D}_{\overline{B}_s^0} = \frac{q^{OS} \cdot \overline{D}_{tag}^{OS} + q^{SS} \cdot \overline{D}_{tag}^{SS}}{1 + q^{OS} \cdot \overline{D}_{tag}^{OS} \cdot q^{SS} \cdot \overline{D}_{tag}^{SS}}, \tag{A.14}$$

$$q = \begin{cases} \frac{q \cdot D_{B_s^0}}{|q \cdot D_{B_s^0}|} & |q \cdot D_{B_s^0}| > |q \cdot \overline{D}_{\overline{B}_s^0}| \\ \frac{q \cdot \overline{D}_{\overline{B}_s^0}}{|q \cdot \overline{D}_{\overline{B}_s^0}|} & \text{otherwise,} \end{cases} \tag{A.15}$$

$$D_{tag} = \frac{|q \cdot D_{B_s^0}| + |q \cdot \overline{D}_{\overline{B}_s^0}|}{2 + q \cdot (|q \cdot D_{B_s^0}| - |q \cdot \overline{D}_{\overline{B}_s^0}|)}, \tag{A.16}$$

where D_{tag}^X is dilution for the B_s^0 initial flavour and $\overline{D}_{tag}^X$ is dilution for the $\overline{B}_s^0$ initial flavour; $X \in \{\text{OS, SS}\}$.

The second diluting effect comes from the decay time resolution D_{tr}, which is described in Sec. 4.5.1.

Lastly, $A_{CP}(t)$ to a good approximation is just $\sin(\phi_s)\sin(\Delta m_s)$. The sole term containing the $\sin(\phi_s)$ is D_{CP} and thus it is simply a derivative of the likelihood with respect to $\sin(\phi_s)$,

$$D_{CP}^{\sin(\phi_s)} = \frac{\sum_{k=1}^{10} N_k f_k(\Omega)}{\sum_{k=1}^{10} \left(a_k f_k(\Omega) \cosh(\frac{1}{2}\Delta\Gamma_s/t) + b_k f_k(\Omega) \sinh(\frac{1}{2}\Delta\Gamma_s/t)\right)} \frac{dd_k}{d\sin(\phi_s)}\Big|_{\phi_s=\hat{\phi_s}} \tag{A.17}$$

where $\sum_{k=1}^{10}$ is a summation over polarization amplitudes; N_k, a_k, b_k and d_k[1] are the coefficients from Tab. 2.1; f_k are the angular functions from Tab. 2.1; $\Delta\Gamma_s$ is the B_s^0 decay width difference. For clarity, this D_{CP} is referred to as $D_{CP}^{\sin(\phi_s)}$. Unlike the D_{tr} and D_{tag}, the D_{CP} depends on the observables of interest: t and Ω. The value of ϕ_s is rather small, which means that $\sin(\Delta m_s t)$ will be barely seen on the plot. Therefore, it is interesting to look at the likelihood derivative with respect to $\cos(\phi_s)$ instead of $\sin(\phi_s)$ and build an "asymmetry"-like variables from it. To plot the MLE

[1]Note, that d_k is a coefficient in front of $\sin(\Delta m_s t)$

of $\cos(\phi_s)\sin(\Delta m_s t)$ asymmetry the $D_{CP}^{\cos(\phi_s)}$ is defined as,

$$D_{CP}^{\cos(\phi_s)} = \frac{\sum_{k=1}^{10} N_k f_k(\Omega)}{\sum_{k=1}^{10} \left(a_k f_k(\Omega)\cosh(\frac{1}{2}\Delta\Gamma_s/t) + b_k f_k(\Omega)\sinh(\frac{1}{2}\Delta\Gamma_s/t)\right)} \frac{dd_k}{d\cos(\phi_s)}\Big|_{\phi_s=\hat{\phi}_s}. \tag{A.18}$$

In this case, one effectively plots MLE of $\cos(\phi_s)\sin(\Delta m_s)$, where $\cos(\phi_s) \approx 1$ and the B_s^0-$\overline{B}_s^0$ oscillations are clearly visible. Note that here the acceptance effects are ignored in both Eq. A.17-A.18, but in principle, these would affect the asymmetry MLE.

Figure A.7 shows the MLE of $\sin(\phi_s)$ CP-asymmetry in bins of decay time. Figure A.8 shows $\cos(\phi_s)$-terms of the "asymmetry".

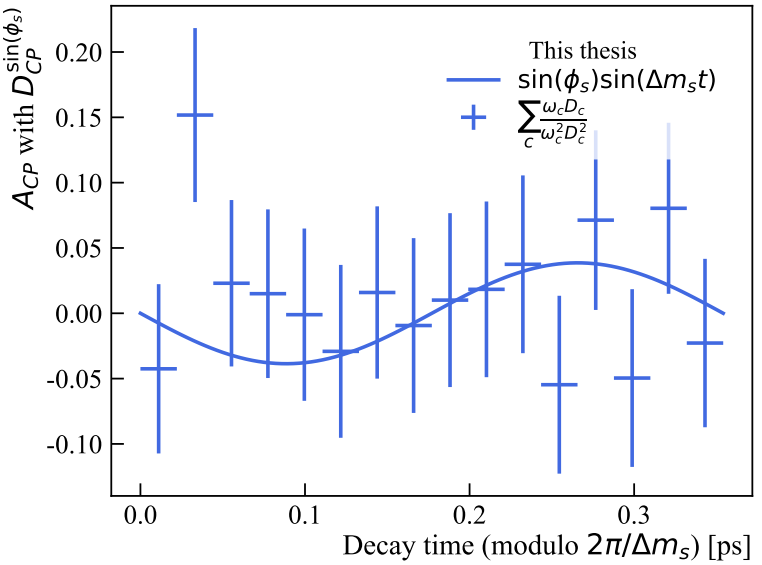


Figure A.7: Decay-time asymmetry $\sin(\phi_s)\sin(\Delta m_s t)$ in the $B_s^0 \to J/\psi K^+K^-$ decays. See the text for the derivation of the points.

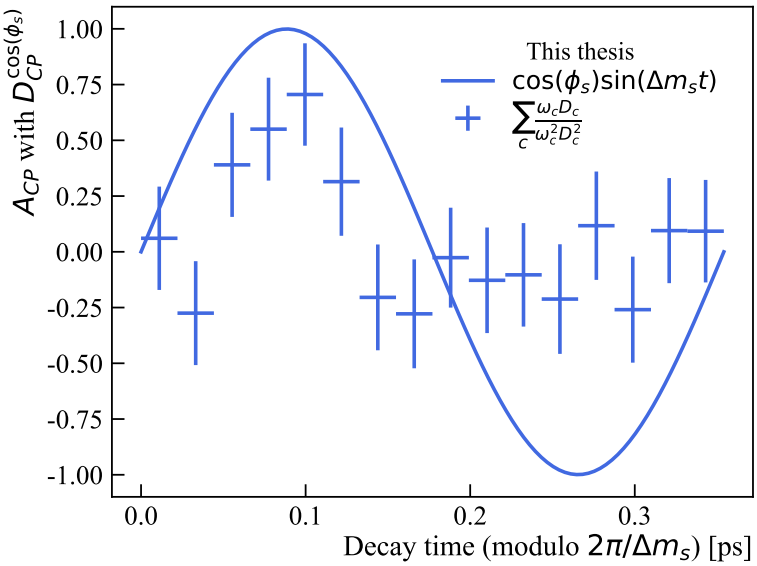


Figure A.8: Decay-time asymmetry $\cos(\phi_s)\sin(\Delta m_s t)$ in the $B_s^0 \to J/\psi K^+K^-$ decays. See the text for the derivation of the points.

Bibliography

[1] W. Hulsbergen, M. Merk, and I. van Vulpen, *Particle Physics 1: Lecture notes for the first year master course on electroweak part of the Standard Model*, Lecture notes, 2018.

[2] N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, Astronomy & Astrophysics **641** (2020) A6.

[3] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[4] G. 't Hooft, *Symmetry Breaking Through Bell-Jackiw Anomalies*, Phys. Rev. Lett. **37** (1976) 8.

[5] M. Laine and M. Meyer, *Standard model thermodynamics across the electroweak crossover*, Journal of Cosmology and Astroparticle Physics **2015** (2015) 035.

[6] M. Z. Barel, K. D. Bruyn, R. Fleischer, and E. Malami, *In pursuit of new physics with $B_d^0 \to J/\psi K^0$ and $B_s^0 \to J/\psi\phi$ decays at the high-precision Frontier*, Journal of Physics G: Nuclear and Particle Physics **48** (2021) 065002.

[7] R. Fleischer, *Flavour Physics and CP Violation: Expecting the LHC*, `arXiv:0802.2882`.

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

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

[10] J. P. Ellis, *Ti k z-feynman: Feynman diagrams with ti k z*, Computer Physics Communications **210** (2017) 103.

[11] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[12] M. Kobayashi and T. Maskawa, *$\mathcal{CP}$-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[13] P. Kooijman and N. Tuning, *Particle Physics 2: Lectures on CP violation*, Lecture notes, 2015.

[14] CKMfitter group, J. Charles *et al.*, *$\mathcal{CP}$ violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`, updated results and plots available at `http://ckmfitter.in2p3.fr/`.

[15] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[16] BABAR Collaboration, B. Aubert *et al.*, *Measurement of CP-violating asymmetries in b^0 decays to CP eigenstates*, Phys. Rev. Lett. **86** (2001) 2515.

[17] Belle Collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802.

[18] LHCb Collaboration, R. Aaij *et al.*, *Observation of CP Violation in Charm Decays*, Phys. Rev. Lett. **122** (2019) 211803.

[19] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal* CP *nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

[20] Particle Data Group, R. L. Workman *et al.*, *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2022** (2022) , 083C01.

[21] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics **28** (1962) 870, `arXiv:https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf`.

[22] T. Asaka and M. Shaposhnikov, *The νMSM, dark matter and baryon asymmetry of the universe*, Physics Letters B **620** (2005) 17.

[23] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[24] M. Artuso, G. Borissov, and A. Lenz, *CP Violation in the B_s^0 system*, arXiv:1511.09466.

[25] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[26] Y. Xie, P. Clarke, G. Cowan, and F. Muheim, *Determination of $2\beta_s$ $B_s^0 \to J/\psi K^+K^-$ decays in the presence of a K^+K^- S-wave contribution*, Journal of High Energy Physics **2009** (2009) 074–074.

[27] LHCStudyGroup, T. S. Pettersson and P. Lefèvre, *The Large Hadron Collider: conceptual design* , CERN, 1995.

[28] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s}$ =7 and 13 TeV*, JHEP **12** (2017) 026, arXiv:1710.04921.

[29] H. G. Moser and A. Roussarie, *Mathematical methods for B^0 anti-B^0 oscillation analyses*, Nucl. Instrum. Meth. **A384** (1997) 491.

[30] K. Govorkova, *Measuring asymmetry in strange beauty*, PhD thesis, Nikhef, 2020, (thesis) Presented 01 Jul 2020.

[31] D. Fazzini, *Flavour Tagging in the LHCb experiment*, PoS **LHCP2018** (2018) 230.

[32] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.

[33] LHCb collaboration, R. Aaij *et al.*, *B flavour tagging using charm decays at the LHCb experiment*, JINST **10** (2015) P10005, arXiv:1507.07892.

[34] LHCb collaboration, *Physics case for an LHCb Upgrade II — Opportunities in flavour physics, and beyond, in the HL-LHC era*, arXiv:1808.08865.

[35] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, 2001.

[36] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[37] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002.

[38] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001.

[39] D. van Eijk *et al.*, *Radiation hardness of the LHCb Outer Tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **685** (2012) 62.

[40] J. V. Tilburg, *Track simulation and reconstruction in LHCb*, PhD thesis, Vrije U. Amsterdam, 2005, (thesis) Presented on 01 Sep 2005.

[41] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, 2003.

[42] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[43] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031, 2003.

[44] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[45] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000.

[46] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001.

[47] R. Aaij *et al.*, *A comprehensive real-time analysis model at the LHCb experiment*, Journal of Instrumentation **14** (2019) P04006.

[48] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of $\mathcal{CP}$ violation in $B_s^0 \to J/\psi K^+K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801, `arXiv:1411.3104`.

[49] LHCb collaboration, R. Aaij *et al.*, *Updated measurement of time-dependent $\mathcal{CP}$-violating observables in $B_s^0 \to J/\psi K^+K^-$ decays*, Eur. Phys. J. **C79** (2019) 706, Erratum ibid. **C80** (2020) 601, `arXiv:1906.08356`.

[50] LHCb collaboration, R. Aaij *et al.*, *Resonances and $\mathcal{CP}$-violation in B_s^0 and $\overline{B}_s^0 \to J/\psi K^+K^-$ decays in the mass region above the $\phi(1020)$*, JHEP **08** (2017) 037, arXiv:1704.08217.

[51] LHCb collaboration, R. Aaij *et al.*, *First measurement of the $\mathcal{CP}$-violating phase in $B_s^0 \to J/\psi(e^+e^-)\phi$ decays*, The European Physical Journal C **81** (2021) , arXiv:2105.14738.

[52] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\mathcal{CP}$-violating phase ϕ_s in $\overline{B}_s^0 \to J/\psi\pi^+\pi^-$ decays*, Phys. Lett. **B736** (2014) 186, arXiv:1405.4140.

[53] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\mathcal{CP}$ violating phase and decay-width difference in $B_s^0 \to \psi(2S)\phi$ decays*, Phys. Lett. **B762** (2016) 253, arXiv:1608.04855.

[54] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\mathcal{CP}$-violating phase ϕ_s in $\overline{B}_s^0 \to D_s^+D_s^-$ decays*, Phys. Rev. Lett. **113** (2014) 211801, arXiv:1409.4619.

[55] CDF Collaboration, T. Aaltonen *et al.*, *Measurement of lifetime and decay-width difference in $B_s^0 \to J/\psi\phi$ decays*, Phys. Rev. Lett. **100** (2008) 121803, arXiv:0712.2348.

[56] D0 Collaboration, V. M. Abazov *et al.*, *Measurement of B_s^0 mixing parameters from the flavor-tagged decay $B_s^0 \to J/\psi\phi$*, Phys. Rev. Lett. **101** (2008) 241801, arXiv:0802.2255.

[57] ATLAS Collaboration, G. Aad *et al.*, *Flavor tagged time-dependent angular analysis of the $B_s \to J/\psi\phi$ decay and extraction of $\Delta\Gamma s$ and the weak phase ϕ_s in ATLAS*, Phys. Rev. D **90** (2014) 052007, arXiv:1407.1796.

[58] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the CP-violating phase ϕ_s and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi\phi$ decays in ATLAS*, JHEP **08** (2016) 147, arXiv:1601.03297.

[59] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the CP-violating phase ϕ_s in $B_s^0 \to J/\psi\phi$ decays in ATLAS at 13 TeV*, Eur. Phys. J. C **81** (2021) 342, arXiv:2001.07115.

[60] CMS Collaboration, V. Khachatryan *et al.*, *Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the*

$B_s^0 \to J/\psi\phi(1020)$ *decay channel in pp collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Lett. B **757** (2016) 97, arXiv:1507.07527.

[61] CMS Collaboration, A. M. Sirunyan *et al.*, *Measurement of the CP-violating phase* ϕ_s *in the* $B_s^0 \to J/\psi\phi(1020) \to \mu^+\mu^- K^+K^-$ *channel in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Let. B **816** (2021) 136188.

[62] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and* τ*-lepton properties as of 2021*, Physical Review D **107** (2023) .

[63] LHCb Collaboration, R. Aaij *et al.*, *Improved measurement of cp violation parameters in* $B_s^0 \to j/\psi K^+K^-$ *decays in the vicinity of the* $\phi(1020)$ *resonance*, Phys. Rev. Lett. **132** (2024) 051802, arXiv:2308.01468.

[64] G. Cowan, *Statistical data analysis*, "Oxford Science Publications", 1998.

[65] T. du Pree, *Search for a Strange Phase in Beautiful Oscillations*, PhD thesis, Vrije U., Amsterdam, 2010, Presented on 22 Oct 2010.

[66] S. Stone and L. Zhang, *S-waves and the Measurement of CP Violating Phases in* B_s *Decays*, Phys. Rev. D **79** (2009) 074024, arXiv:0812.2832.

[67] P. Li *et al.*, *Measurement of* $B_s^0 \to J/\psi K^+K^-$ *with the full Run 2 data*, LHCb-ANA-2019-032 ; CERN-LHCb-ANA-2019-032, 2019.

[68] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[69] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, arXiv:0905.0724.

[70] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001, and 2019 update.

[71] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **552** (2005) 566.

[72] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2018) 1, arXiv:1803.00824.

[73] A. Poluektov, *Correction of simulated particle identification response in LHCb using kernel density estimation.* , CERN, Geneva, 2017.

[74] LHCb collaboration, R. Aaij *et al.*, *Global PID performance for charged particles*, LHCB-FIGURE-2020-012, 2020.

[75] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[76] K. Govorkova, *Shape of the combinatorial background in $B_s \to J/\psi\phi$ decay* , CERN, Geneva, 2019.

[77] H. Dembinski, M. Kenzie, C. Langenbruch, and M. Schmelling, *Custom orthogonal weight functions (COWs) for event classification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1040** (2022) 167270.

[78] LHCb Collaboration, *Material for presentation*, https://lhcb.web.cern.ch/speakersbureau/html/Material_for_Presentations.html. Accessed: 2023-08-18.

[79] J. Wishahi, *Tikz version of flavour tagging schematics*, https://gitlab.cern.ch/lhcb-ft/TikZPics/-/tree/master/. Accessed: 2023-08-18.

[80] LHCb collaboration, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, arXiv:1602.07252.

[81] LHCb collaboration, R. Aaij *et al.*, *Precise determination of the B_s^0 - $\bar{B}_s^0$ oscillation frequency*, Nature Physics **18** (2022) 1, arXiv:2104.04421.

[82] C. R. Harris *et al.*, *Array programming with NumPy*, Nature **585** (2020) 357.

[83] R. Aaij, *Triggering on CP Violation Real-Time Selection and Reconstruction of $B_s^0 \to J/\psi K^+K^-$ Decays*, PhD thesis, Nikhef, 2015, (thesis) Presented 07 May 2015.

[84] W. Verkerke and D. Kirkby, *The roofit toolkit for data modeling*, arXiv:physics/0306116.

[85] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, arXiv:1307.2782.

[86] LHCb collaboration, R. Aaij *et al.*, *Measurements of the B^+, B^0, B_s^0 meson and Λ_b^0 baryon lifetimes*, JHEP **04** (2014) 114, arXiv:1402.2554.

[87] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, to be published in Prog. Theor. Exp. Phys. **6** (2020) 083C01.

[88] C. Runge *et al.*, *Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten*, Zeitschrift für Mathematik und Physik (1901) 224–243.

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

[90] C. de Boor, *Subroutine package for calculating with B-splines* , Los Alamos Scientific Laboratory of the University of California, 1971. doi: 10.2172/4740859.

[91] S. Benson *et al.*, *Measurement of the CP-violating phase ϕ_s using $B_s^0 \to J/\psi K^+K^-$ decays in Run 2*, LHCb-ANA-2017-028 ; CERN-LHCb-ANA-2017-028, 2017.

[92] M. Romero Lamas, *Measurement of the bottom-strange meson oscillations at LHCb*, PhD thesis, Universidade de Santiago de Compostela, 2022, Presented 20 December 2022.

[93] J. Wallis, *A Treatise of Algebra, both Historical and Practical*, Oxford: Richard Davis, 1685.

[94] J. Raphson, *Analysis Æequationum Universalis*, London: Thomas Bradyll, 1697.

[95] J. F. Steffensen, *Remarks on iteration*, Scandinavian Actuarial Journal **1933** (1933) 64, arXiv:https://doi.org/10.1080/03461238.1933.10419209.

[96] J. F. Steffensen, *Further remarks on iteration*, Scandinavian Actuarial Journal **1945** (1945) 44, arXiv:https://doi.org/10.1080/03461238.1945.10404918.

[97] A. C. Aitken, *XXV.—On Bernoulli's Numerical Solution of Algebraic Equations*, Proceedings of the Royal Society of Edinburgh **46** (1927) 289–305.

[98] R. Barlow, *Systematic errors: facts and fictions*, arXiv:0207026.

[99] A. C. Davison and D. V. Hinkley, *Bootstrap methods and their application*, Cambridge university press, 1997.

[100] C. Langenbruch, *Parameter uncertainties in weighted unbinned maximum likelihood fits*, The European Physical Journal C **82** (2022) .

[101] P. Koppenburg, *Statistical biases in measurements with multiple candidates*, arXiv:1703.01128.

[102] M. Needham, *Momentum scale calibration using resonances*, .

[103] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the B_s^0–$\overline{B}_s^0$ oscillation frequency in the decay $B_s^0 \to D_s^- \pi^+$*, New J. Phys. **15** (2013) 053021, arXiv:1304.4741.

[104] D. M. Santos, *Study of the very rare decay $B_s \to \mu^+\mu^-$ in LHCb*, PhD thesis, Santiago de Compostela, Universidade de Santiago de Compostela, 2010, (thesis) Presented on 05 May 2010.

[105] Y. Xie, *Some physics and technical issues in analysis of the decay $B_s \to J/\psi K^+ K^-$* , CERN, Geneva, 2012.

[106] R. Aaij *et al.*, *P2VV*, https://gitlab.cern.ch/lhcb-b2cc/p2vv, 2008-2023.

[107] A. Lenz and G. Tetlalmatzi-Xolocotzi, *Model-independent bounds on new physics effects in non-leptonic tree-level decays of b-mesons*, Journal of High Energy Physics **2020** (2020) .

[108] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[109] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007, 2012.

[110] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021, 2013.

[111] T. Poikela *et al.*, *VeloPix: the pixel ASIC for the LHCb upgrade*, Journal of Instrumentation **10** (2015) C01057.

[112] M. van Beuzekom *et al.*, *VeloPix ASIC development for LHCb VELO upgrade*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **731** (2013) 92, PIXEL 2012.

[113] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016, 2014.

[114] LHCb collaboration, *Computing Model of the Upgrade LHCb experiment*, CERN-LHCC-2018-014, 2018.

[115] LHCb Collaboration, R. Aaij *et al.*, *RTA and DPA dataflow diagrams for Run 1, Run 2, and the upgraded LHCb detector* , LHCB-FIGURE-2020-016, 2020.

[116] LHCb collaboration, R. Aaij *et al.*, *A comparison of CPU and GPU implementations for the LHCb experiment run 3 trigger*, Computing and Software for Big Science **6** (2021) .

[117] LHCb Collaboration, V. Lukashenko, *`Trigger Objects Compatibility Checker`*, `https://gitlab.cern.ch/lhcb/Rec/-/blob/master/Tr/TrackCheckers/src/TriggerObjectsCompatibilityChecker.cpp`, 2022.

[118] Q. Fuhring, *Flavour Tagging at the LHCb Experiment*, CERN-Poster-2020-1031, 2020.

[119] ATLAS Collaboration, M. Aaboud *et al.*, *CP-violation measurement prospects in the $B_s^0 \to J/\psi\phi$ channel with the upgraded ATLAS detector at the HL-LHC*, .

[120] CMS Collaboration, A. Sirunyan *et al.*, *CP-violation studies at the HL-LHC with CMS using* $\mathrm{B_s^0}$ *decays to* $\mathrm{J}/\psi\phi(1020)$, .

[121] CMS Collaboration, V. Khachatryan *et al.*, *Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B_s^0 \to J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8\ TeV$*, Physics Letters B **757** (2016) 97.

[122] CMS Collaboration, A. Sirunyan *et al.*, *Recording and reconstructing 10 billion unbiased b hadron decays in CMS*, .

[123] K. D. Bruyn and R. Fleischer, *A roadmap to control penguin effects in $B_d^0 \to J/\psi K_s^0$ and $B_s^0 \to J/\psi\phi$*, Journal of High Energy Physics **2015** (2015) .

[124] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\mathcal{CP}$-violating phase β in $\overline{B}^0 \to J/\psi\pi^+\pi^-$ decays and limits on penguin effects*, Phys. Lett. **B742** (2015) 38, arXiv:1411.1634.

[125] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\mathcal{CP}$ violation parameters and polarisation fractions in $B_s^0 \to J/\psi\,\overline{K}^{*0}$ decays*, JHEP **11** (2015) 082, arXiv:1509.00400.

[126] A. D. Bukin, *Fitting function for asymmetric peaks*, arXiv:0711.4449.

Summary

Beep-beep-beep.

This is your alarm clock ringing at 8 am. You open your eyes and slowly get up. Dragging your feet, you walk to the kitchen and take a kettle. The morning is chilly, and the steam from the kettle rushes up as the water is boiled. Half-asleep, you put some coffee in a cup and pour in some hot water. You can not wait for this first cup of the day! But all of a sudden, a spray of all kinds of fundamental particles: photons, electrons, neutrinos, and quarks beam into your eyes. Your coffee is gone, annihilated, and transformed into a ray of quantum particles! How did this happen? You look at the coffee jar. "Oh, no!" - you are disappointed in yourself. The label on the coffee jar says in big warning letters: "Anti-coffee. Please do not mix with matter water!". Bad luck, my friend. You accidentally mixed antimatter with matter. What a rough, rough morning.

It sounds like a science fiction movie, but this is the world according to the Standard Model - the theory that describes the quantum world. "Well, it must describe it wrong then," you say. Evidently, one does not come across antimatter so often to spoil their morning coffee! Indeed, there is something wrong with the theory that particle physics has developed in the last seventy years: it predicts that the Universe should consist of almost equal amounts of matter and antimatter. Luckily, this is not the case! If the prediction were true, all the matter and antimatter would have annihilated, leaving nothing behind but an empty Universe without any stars, planets, or life. Our daily observation shows that we exist and that almost everything around us is made of matter: protons, neutrons, and electrons. So where did all the antimatter go? Or why is there so much matter? And why did the Universe survive? These questions are at the core of this thesis.

What can explain the difference? Let's look closer at the Standard Model. The Standard Model describes three of four fundamental forces: electromagnetic, strong, and weak, omitting the gravitational force. Two of them, electromagnetic and strong, are the same for matter and antimatter.

However, the last one - the weak force, which is responsible for nuclear decay, distinguishes between matter and antimatter. The weak force changes the probability of matter and antimatter interactions, preferring one over another.

So, could not this explain the disappearance of antimatter? Unfortunately, the amount of difference created by the weak force is insufficient to explain the massive difference between matter and antimatter in the Universe. And I mean *really* not enough. The weak force of the Standard Model predicts the matter-antimatter difference to be 10 billion times smaller than what we observe in the Universe. Therefore, physicists look for new sources that increase the difference between matter and antimatter.

In this thesis, I studied the difference in decay rates of B_s^0 particles and anti-B_s^0 particles, which is sensitive to the new sources of the matter-antimatter difference. The B_s^0 particles contain two types of quarks, also known as flavors of quarks: beautiful and strange. The poetic names of quark types might seem funny to the outsider because they kind of are. But the names simply label the quantum charges each quark carries. The B_s^0 particles are produced in the collisions of protons at the Large Hadron Collider, located on the border of Switzerland and France, near Geneva. The quantum world is defined by probabilities, so some of B_s^0 particles decay to other two particles: J/ψ and $\phi(1020)$. However, both J/ψ and $\phi(1020)$ immediately decay into muons and kaons. Both muons and kaons are charged and will leave electrical signals when flying through the LHCb detector. This allows me to reconstruct them, just like you do with your phone camera and photons, which are particles that light consists of.

The anti-particle of B_s^0 is $\overline{B}_s^0$, which can decay into the same particles J/ψ and $\phi(1020)$. Naively, both the $B_s^0 \to J/\psi\,\phi(1020)$ and $\overline{B}_s^0 \to J/\psi\,\phi(1020)$ decays should happen with the same probability. However, because of the weak force, there is a difference between the probabilities of these two processes. The amount of this difference in the B_s^0 particle system is quantified by a parameter called ϕ_s. The Standard Model predicts a tiny, almost zero, ϕ_s. However, if there are new effects that modify the difference between matter and antimatter in the B_s^0 system, the value of ϕ_s should be bigger than one predicted in the Standard Model. The enhanced value of ϕ_s would mean the discovery of new unknown effects that increase the difference between matter and antimatter in the Universe, therefore contributing to the solution of the puzzle of why we are all made of matter and not antimatter.

In this thesis, I used the full dataset of $B_s^0 \to J/\psi\,\phi(1020)$ and $\overline{B}_s^0 \to J/\psi\,\phi(1020)$ decays, collected between 2015 and 2018 by LHCb detector to measure the difference between matter and antimatter. The measured value

is,

$$\phi_s = -2.212 \pm 1.232 \pm 0.34\ ^\circ,$$

where the first value is the measured value, the second is the uncertainty from the limited statistics of the sample, followed by the uncertainty from systematical effects in the measurement. This is the world's most precise measurement of ϕ_s, and it is in agreement with the Standard Model. However, this does not conclude that the new effects do not contribute. Moreover, the measured value is also compatible with zero and, therefore, can be just a fluctuation from the case when no difference between matter and antimatter exists! More data is needed to determine whether this is a fluctuation or a real observation. To achieve this, I worked on preparing for the new data-taking period at LHCb experiment. I had the pleasure of helping configure and calibrate the new pixel VELO tracking detector located in the heart of the LHCb detector. I also worked on the development of new, faster reconstruction software, which allows physicists to collect much more data.

Samenvatting

Piep-piep-piep.

Dit is je wekker om acht uur 's ochtends. Je opent je ogen en staat op. Je slentert naar de keuken en vult de waterketel. De ochtend is koud, en er dampt stoom uit de waterketel terwijl het water kookt. Nog half slapend schep je wat koffie in je cafetière en giet er heet water bij. Je kunt niet wachten op de eerste kop van de dag! Maar ineens worden je ogen geraakt door een straal fundamentele deeltjes: fotonen, elektronen, neutrinos en kwarks raken je netvlies. Je koffie is weg, vernietigd, gentransformeerd in een straal kwantumdeeltjes. Hoe kon dit gebeuren? Je kijkt op de koffiepot. "Oh, nee!" - je bent teleurgesteld in jezelf. Op het etiket staat in grote waarschuwingsletters: "Anti-koffie. Niet mixen met water!". Helaas, mijn vriend. Je hebt per ongeluk antimaterie met materie gemixt. Wat een zware, zware ochtend.

Het klinkt als een sciencefictionfilm, maar dit is de wereld volgens het Standaardmodel - de theorie die de kwantumwereld beschrijft. "Nou, die beschrijving moet dan verkeerd zijn," zeg je bij jezelf. Zo vaak komt het namelijk niet voor dat je in aanraking komt met anti-materie en dat je koffie daardoor annihileert. Er is inderdaad iets mis net de theorie die de deeltjesfysica heeft ontwikkeld in de afgelopen zeventig jaar: het voorspelt namelijk dat het universum in bijna gelijke mate uit materie en antimaterie bestaat. Gelukkig is dit niet het geval! Als deze beschijving juist, zou alle materie en antimaterie vernietigd zijn, en zou de oerknal een leeg universum zou hebben achtergelaten, zonder sterren, planeten, of leven. Onze alledaagse ervanigen vertellen ons dat we bestaan, en dat bijna alles om ons heen gemaakt is van materie: protonen, neutronen en elektronen. Dus waar is dan alle antimaterie gebleven? Of, anders gesteld, waarom is er zoveel materie? Hoe komt het dat het universum heelal nog steeds bestaat? Deze vragen liggen in het hart van dit proefschrift.

Wat kan het verschil in materie en antimaterie verklaren? Laten we het standaardmodel eens beter bestuderen. Het standaardmodel beschrijft drie

van de vier fundamentele natuurkrachten: de elektromagnetische kracht, de zwakke kernkracht en de sterke kernkracht, maar niet de zwaartekracht. Twee van deze, de elektromagnetische en sterke kracht, zijn hetzelfde voor zowel materie als antimaterie. De zwakke kernkracht, echter, die verantwoordelijk is voor nucleair verval, maakt onderscheid tussen materie en antimaterie. De waarschijnlijkheid dat bepaalde vervallen voorkomen verschilt tussen materie en antimaterie, waarbij de een waarschijnlijker is dan de ander.

Is dit verschil genoeg om het probleem op te lossen? Helaas is het verschil in materie en antimaterie dat voorspeld wordt door de zwakke kernkracht niet voldoende om het enorme verschil tussen materie en antimaterie in het universum te verklaren. En dan bedoel ik echt *bij lange na* niet genoeg. De zwakke kernkracht uit het standaardmodel voorspelt een verschil tussen materie en antimaterie dat 10 miljard maal kleiner is dan wat we observeren in het universum. Om die reden zoeken natuurkundigen naar een andere uitleg voor het verschil tussen materie en antimaterie.

In dit proefschrift heb ik het verschil in vervalsnelheid van B_s^0-deeltjes en anti-B_s^0-deeltjes bestudeerd, wat gevoelig is voor de nieuwe bronnen van het verschil tussen materie en antimaterie. De B_s^0 deeltjes bevatten twee typen kwarks, deze typen worden ook wel de *smaak* van de kwark genoemd. De twee smaken dragen de Engelse namen beauty en strange. Deze poëtische namen klinken voor de buitenstaander misschien grappig – en zijn dat ook wel – maar ze refereren naar een kwantumlading die elke kwark draagt. Deze kwantumlading stelt natuurkundigen in staat om de deeltjes te onderscheiden, en om uit te vinden hoe de kwarks zich gedragen onder invloed van de fundamentele krachten. De B_s^0 deeltjes worden geproduceerd bij de botsingen tussen protonen in de Large Hadron Collider, gelegen op de grens tussen Frankrijk en Zwitserland, vlakbij Geneve. De kwantumwereld wordt geregeerd door waarschijnlijkheid, dus sommige B_s^0 deeltjes vervallen naar twee andere deeltjes: J/ψ en $\phi(1020)$. Echter, zowel J/ψ als $\phi(1020)$ vervallen onmiddelijk in muonen en kaonen. Zowel muonen als kaonen zijn geladen, en geven elekrische signalen af wanneer ze door de LHCb detector vliegen. Dit stelt mij in staat om ze te reconstrueren, net zoals je telefooncamera doet met fotonen, de deeltjes waar licht uit bestaat.

Het antideeltje van de B_s^0 is de $\overline{B}_s^0$ en kan vervallen in precies dezelfde deeltjes J/ψ en $\phi(1020)$. Normaal gespideen zou je verwachten dat de $B_s^0 \to J/\psi\,\phi(1020)$ en $\overline{B}_s^0 \to J/\psi\,\phi(1020)$ vervalprocessen in gelijke mate voor komen. Echter, vanwege de sterke kernkracht is er een verschil in kansen dat deze processen optreden. Dit verschil wordt gekwantificeerd door een parameter die ϕ_s wordt genoemd. Het standaardmodel voorspelt een zeer kleine ϕ_s, bijna gelijk aan nul. Echter, als er nieuwe effecten zijn

die het verschil tussen materie en antimaterie veranderen, zou de waarde van ϕ_s groter moeten zijn dan die voorspeld door het standaardmodel. De vergrote waarde van ϕ_s zou de ontdekking aanduiden van nieuwe, onbekende effecten die het verschil tussen materie en antimaterie vergroten, en daarmee bijdragen aan de oplossing van de puzzel waarom we allemaal van materie en niet van antimaterie zijn gemaakt.

In dit proefschrif heb ik de volledige dataset van $B_s^0 \to J/\psi\,\phi(1020)$ en $\overline{B}_s^0 \to J/\psi\,\phi(1020)$ vervalprocessen gebruikt, die zijn gemeten van 2015 - 2018 in de LHCb detector, om het verschil tussen materie en antimaterie te bestudeien. De gemeten waarde is

$$\phi_s = -2.212 \pm 1.232 \pm 0.34\ ^\circ,$$

waar het eerste getal de gemeten waarde aanduidt, het tweede de onzekerheid van de beperkte grotte van de dataset, en het laatste de onzekerheid van systematische effecten in de meting. Dit is 's werelds meest precies meting van ϕ_s, en het is in overeenstemming met het Standaardmodel. Dit betekent echter niet dat de nieuwe effecten niet bijdragen. Bovendien is de gemeten waarde ook compatibel met nul en kan daarom slechts een fluctuatie zijn ten opzichte van het geval waarin er geen verschil bestaat tussen materie en antimaterie! Er is meer data nodig om uit te sluiten of dit een uitzondering is of een echte waarneming van een grotere waarde van ϕ_s. Om dit te bewerkstelligen, heb ik meegewerkt aan de voorbereiding voor de nieuwe data-verzamel periode van het LHCb experiment. Ik had het genoegen om te helpen bij het calibreren van de nieuwe pixel VELO tracking detector, die zich in het hart van de LHCb detector bevindt. Ik heb ook gewerkt aan de ontwikkeling van nieuwe, snellere reconstructie software, die natuurkundigen in staat moet stellen om veel meer data te vergaren.

Короткий зміст

Біп-біп-біп.

Це звук твого будильника о восьмій ранку. Ти продираєш очі та вислизуєш зі свого ліжка. Волочачи ноги до кухні, ти ставиш чайник. Ти дивишся як від чайника біжить струмок пари. З напіввідкритими очима, ти кладеш трохи кави у чашку і заливаєш її кип'ятком. Нарешті, перша кава цього дня! Аж раптом, потік найрізноманітніших фундаментальних частинок: фотонів, електронів, нейтрино та кварків, засліплює тобі очі. Кава зникла, анігілювала і перетворилася у промені з квантових частинок! Як так сталося? Приголомшено ти дивишся на банку з-під кави. "Ох, ні! розчаровано вигукуєш ти. "Антикава. Будь ласка, не змішуйте з водою з матерії"виведено на банці. Це точно не твій день. Випадково змішати антиматерію з матерією! Який невдалий ранок.

Звучить як науково-фантастичне кіно, але саме так виглядає наш Всесвіт відповідно до Стандартної Моделі - теорії, яка описує квантовий світ. "Тоді вона описує його невірно кажеш ти. Очевидно, що треба сильно постаратися аби знайти антиматерію для того, щоб зіпсувати ранкову каву! Дійсно, щось не так з теорією, яку фізика частинок побудувала за останні сімдесят років. Ця теорія передбачає що Всесвіт складається з майже однакової кількості матерії та антиматерії. Але на щастя це не так! Тому що якби це передбачення справдилось, матерія та антиматерія давно б анігілювали, а наш Всесвіт залишився б порожнім без ніяких зірок, планет та життя. Щоденні спостереження вказують на те що ми існуємо і що майже все навколо нас зроблено з матерії: протонів, нейтронів та електронів. Тож куди поділась уся антиматерія? Або ж чому у світі так багато матерії? І чому Всесвіт існує? Ці питання в центрі цієї дисертації.

Що може пояснити різницю між кількістю матерії та антиматерії? Розглянемо детальніше Стандартну Модель. Стандартна Модель описує три з чотирьох фундаментальних взаємодій: електромагнітну, сильні та слабку. Дві з них, електромагнітна та сильна, однаково впливають

що на матерію, що на антиматерію, не розрізняючи між ними. Однак, слабка взаємодія, яка відповідає за ядерний розпад, розрізняє матерію та антиматерію. Вона змінює вірогідність взаємодій матерії чи антиматерії, віддаючи перевагу одній чи іншій.

То чи не розв'язує це нашу проблему? Через слабку взаємодію вся антиматерія могла розпастись до зустрічі з матерією. На жаль, слабкої взаємодії недостатньо щоб пояснити ту величезну різницю між кількістю матерії та антиматерії у Всесвіті. Ду-у-у-у-уже недостатньо. Слабка взаємодія передбачає що різниця між матерією та антиматерією мала бути в 10 000 000 000 разів меншою, ніж різниця присутня у Всесвіті. Тому фізики шукають нові джерела, які могли б пояснити цю різницю.

У цій дисертації, я шукала такі нові джерела у розпадах частинок, відомих як B_s^0 частинки. B_s^0 частинки складаються з двох типів кварків, які звуться ароматами: красивого та дивного. Ці поетичні назви здаються кумедними сторонній людині (і це дійсно так!), але ці назви просто позначають квантові заряди носіями яких є кожен кварк. B_s^0 частинки народжуються у зіткненнях протонів у Великому Адронному Колайдері, який знаходиться біля Женеви, на кордоні між Швейцарією та Францією. Квантовий світ визначений вірогідностями, тому деякі зі створених B_s^0 частинок розпадуться у дві інші частинки: J/ψ та $\phi(1020)$. А J/ψ та $\phi(1020)$ негайно розпадаються на мюони та каони. І мюони, і каони - заряджені частинки, які залишають електричні сигнали пролітаючи крізь об'єм детектора. Ці сигнали дозволяють мені реконструювати ці частинки, так само як камера твого телефону відображає сигнали залишені фотонами - частинками світла.

Античастинки B_s^0 частинок називаються $\overline{B}_s^0$. Вони розпадаються на такі самі частинки J/ψ та $\phi(1020)$. Здавалося б імовірність розпадів $B_s^0 \to J/\psi\,\phi(1020)$ та $\overline{B}_s^0 \to J/\psi\,\phi(1020)$ має бути однаковою. Однак через слабку взаємодію існує невелика різниця в ймовірностях цих двох процесів. Кількісно ця різниця описується параметром, який називається ϕ_s. Стандартна Модель передбачає що ϕ_s приблизно дорівнює нулю. Але якщо існують ефекти які можуть впливати на різницю між матерією та антиматерією, значення ϕ_s буде більшим ніж це передбачає Стандартна Модель. Якщо експериментально виміряне значення буде набагато більшим за передбачення Стандартної Моделі, це означатиме відкриття невідомих ефектів які збільшують різницю між матерією та антиматерією.

Для того, щоб поміряти цю різницю, я використала дані експерименту LHCb, зібрані між 2015 та 2018 роками. Виміряне значення різниці

дорівнює:

$$\phi_s = -2.212 \pm 1.232 \pm 0.34\ ^\circ,$$

де перше число - це виміряне значення, друге - це похибка зумовлена обмеженим розміром зібраних даних, а остання - це систематична похибка. Це найточніше вимірювання ϕ_s у світі.

Тож про що нам каже цей результат? Якщо насправді ϕ_s має нульове значення, то у приблизно 2.5% випадків ми спостерігатимемо такий результат. Для того, щоб з'ясувати чи це флуктуація, чи це справжнє значення нам потрібно більше даних. Саме через це, я також працювала над підготовкою детектора до збору нових даних. Я допомогла відкалібрувати новий піксельний VELO детектор для трекінгу, що знаходиться в самому центрі LHCb детектора. Я також працювала над новим програмним забезпеченням для швидкого реконструювання частинок з електричних сигналів.

Acknowledgements

If there is anything I can bet on,
it is that all of you will read this page.

First and foremost, I would like to thank my supervisor *Wouter Hulsbergen* and my promotor *Gerhard Raven* for giving me an opportunity to do a PhD with Nikhef and LHCb. Thank you for supporting every crazy endeavor I had during these years and for the freedom to do whatever I found interesting! Thank you for always being there and finding 10 minutes to explain anything I wanted to know about. *Wouter*, thank you for being strict and fair through these years and not letting a single thing slide, even if I was stupidly stubborn in most cases. Your curiosity and kindness exemplify what a scientist should be like. *Gerhard*, thank you for being there every time I needed to complain about everything and everyone, for the best road trip to Geneva *ever*, and for teaching me the most important - one should always be paranoid!

I thank my paranymphs, *Suzanne* and *Alice* for helping with this defense. *Suzanne*, thank you for being there for me during these years, especially at the end of it, when I had to do the worst thing imaginable - write! Thank you for all the talk and our clay adventures. *Alice*, thank you for being my opera&ballet buddy. Although we have been perfectly time-misaligned with our CERN/Nikhef years, you have always been a good friend to me. I am happy my plants found a new home!

Dear *post-apocalyptic Bfys group*, also known as "the best social experiment of Nikhef", you are the best colleagues one could dream of. I thank our gracious ex-leader *Marcel* for making this social experiment successful and ensuring we all arrive home safe after the LHCb BBQs. *Antonio*, thank you for being my best buddy during coronavirus times, feeding me Italian food, and always being chill. I will be forever grateful to you for teaching Pepijn how to cook parmigiana. *Mara*, thank you for caring for me and all other PhDs during our time at Nikhef. Thank you for helping me navigate the postdoc application. *Kazu*, thanks for being open to the terrible jokes I make, for letting me catsitting your beautiful cats, and all the chocolates I

got from you when I was hungry. *Jacco*, thank you for being the best Dutch male officemate I ever had! Finally, you can enjoy peaceful days at Nikhef! :) *Keri*, thank you for being badass and saying yes to being on the committee of this thesis. *Patrick*, thank you for being my CERN officemate for almost seven months! Thanks to all other Bfys seniors *Niels, Kristof, Chris, Greco, Tjeerd, Ann-Kathrin, Suzan*, whom I did not have a chance to work closely with, but hopefully, next time!

Maarten, thank you for listening to my rants during coffees for all these years, giving excellent suggestions, and keeping your door open for discussion at any point in time. I hope to see you at the next Ukrainian music concert! *Andrii and Olya*, thank you for being good friends for Pepijn and me during our PhD times and for providing your apartment to be an art studio. *Aleksandra*, thank you for always organizing something for us and for all the cultural references I learned from you. Also, thank you for catsitting Amadeya. I must come to Wrocław to pick up the best cherry liquor! *Miriam*, thank you for backing me up when necessary and helping me navigate the beginning of my PhD. *Swedish Daniel*, thank you for pinging Wouter and Gerhard at Nikhef when I was absent. Use wisely the power of Kägi/Maltesers/Korivka. *Lex*, thank you for keeping the group's mental state in check and reminding me about a work-life balance. *Brian*, thank you for forcing me to go out sometimes, introducing me to rugby, and throwing the best Halloween party in my apartment! *Elena*, thank you for being so kind to me during my master's and taking me to caves ouvertes. *Robbert*, thank you for being the best BND buddy and explaining how VeloPix works. *Sevda*, thank you for helping me out when I started in ϕ_s analysis. *Jordy*, thank you for all the beers and all the cool times we had together! *Silvia*, thank you for being a great officemate and being fun to be around! *Igor*, thank you for being there for me through my master's and the first years of my PhD. *Mauricio*, thank you for being the best shift leader to have sleepless nights with. *Cristina*, thank you for sharing Stan's office (and privileges) with me and lending me your ski suit for Moriond! *Efren and Gangrong*, thank you for a year of fun! I thank all others of Bfys group who were part of this journey: *Spanish Daniel, Scottish Daniel, Emmy, Maxime, Roel, Mick, Andrea, Laurent, Carlos, Jan, MD, Andrej, Carolina, Davide, Xenofan, Mark, Evridiki, Wouter, Michele, Katya, Maurice, Olaf, Bas, Anna, Alex, Yuchan, Max and Melika.*

I thank all the non-Nikhef ϕ_s team: *Marcos, Ramon, Peilian, Diego, Veronika, Maurice, Kechen, Pete, and others* for finishing ϕ_s! *Marcos* and *Ramon*, thank you for being there, not giving up even if we wanted to, and for all the beers and wines that we have shared. Thank both of you, *Diego*

and Veronika, for welcoming me to Santiago de Compostela, one of the best cities in the world, and showing me all that tasty food. *Peilian*, thank you for always being everywhere and doing everything. What is your secret?

Thank you, *Victor, Paula, Edgar, Claudia, Karol, Nathan, Pawel, and all LHCb VELO team*, for welcoming me at LHCb during commissioning and making the VELO experience one of the best experiences I had during my PhD! See you in Upgrade II?

I thank all the people that I met during my time at CERN. *Lukas*, thank you for being probably the oldest colleague I had. Since Summer Studentship till PhD! Thank you and *Alessandro* for making Starterkit Run III possible despite all the circumstances. Thanks to both of you, *Christina* and *Luana*, for making your house the hottest spot in the entire SGP. *Kevin*, thank you for all the memes, your support of Ukraine, and just being there for me.

I would also like to thank my *Kyiv group*: Vasya, VM, Sasha, and others for standing strong.

I thank my *mom* for supporting my application to the master's program in Amsterdam, which ultimately led to this PhD. Thank you for all the sacrifices made in my name along this way. I also thank my *grandma* for her support of my physics adventures.

I thank *Amadeyechka* for being the sweetest cat-potato.

I thank *Pepijn* for all his love, kindness, and support through the hard times and the good.

Whoever you are, thank you for making this PhD the most incredible experience one could have.

Cover by Margarita Yakimenko