technische universität
dortmund

Dissertation zur Erlangung des akademischen Grades
Dr. rer. nat.

Measurements of Beauty Hadron to Charmonium Decays

–

Branching Fraction of $\Lambda_b^0 \to \psi(2S)\Lambda$ and CP Violation in $B^0 \to \psi K_S^0$

Patrick Mackowiak
geboren in Datteln

2021

Lehrstuhl für Experimentelle Physik V
Fakultät Physik
Technische Universität Dortmund

Der Fakultät Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften vorgelegte Dissertation.

Erstgutachter:	Prof. Dr. Bernhard Spaan
Zweitgutachter:	Prof. Dr. Kevin Kröninger
Vorsitzender der Prüfungskommission:	Prof. Dr. Christoph Lange
Vertreter der wissenschaftlichen Mitarbeiter:	Dr. Ulf Berges
Datum des Einreichens der Arbeit:	22. März 2021
Datum der mündlichen Prüfung:	27. Mai 2021

Abstract

This thesis presents two measurements involving a beauty hadron decaying to a charmonium meson and a strange hadron. First, a measurement of the ratio of branching fractions for $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$ decays is presented. This analysis uses a data set corresponding to $3\,\text{fb}^{-1}$, collected by the LHCb experiment at a centre-of-mass energy of 7 and 8 TeV. The ratio of branching fractions is determined to be

$$\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)/\mathcal{B}(\Lambda_b^0 \to J/\psi\Lambda) = 0.513 \pm 0.023\,(\text{stat}) \pm 0.016\,(\text{syst}) \pm 0.011(\mathcal{B}).$$

The second analysis is a on-going measurement of *CP* violation in $B^0 \to J/\psi K_S^0$, $B^0 \to \psi(2S)K_S^0$ and $B_s^0 \to J/\psi K_S^0$ decays. The analysed LHCb data sets correspond to $6\,\text{fb}^{-1}$, at a centre-of-mass energy of 13 TeV. The measured sensitivities of the *CP* parameters S and C in the analysed B_d^0 decays are

$$\sigma(S) = 0.0133\,(\text{stat}) \pm 0.0063\,(\text{syst}), \quad \sigma(C) = 0.0125\,(\text{stat}) \pm 0.0042\,(\text{syst}).$$

This single measurement will have a better sensitivity than the current world average.

Kurzfassung

In dieser Arbeit werden zwei Messungen in Zerfällen von beauty Hadronen in ein charmonium Meson und einem strange Hadron vorgestellt. In der ersten Messung wird das Verhältnis von Verzweigungsverhältnissen von $\Lambda_b^0 \to \psi(2S)\Lambda$ und $\Lambda_b^0 \to J/\psi\Lambda$ Zerfällen bestimmt. Diese Analyse benutzt einen Datensatz der $3\,\text{fb}^{-1}$ entspricht und mit dem LHCb-Experiment bei einer Schwerpunktsenergie von 7 und 8 TeV aufgezeichnet wurde. Das Verhältnis von Verzweigungsverhältnissen ist bestimmt zu

$$\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)/\mathcal{B}(\Lambda_b^0 \to J/\psi\Lambda) = 0.513 \pm 0.023\,(\text{stat}) \pm 0.016\,(\text{syst}) \pm 0.011(\mathcal{B}).$$

In der zweiten Messung wird *CP* Verletzung in $B^0 \to J/\psi K_S^0$, $B^0 \to \psi(2S)K_S^0$ und $B_s^0 \to J/\psi K_S^0$ Zerfällen bestimmt. Der analysierte LHCb Datensatz entspricht $6\,\text{fb}^{-1}$, aufgenommen bei einer Schwerpunktsenergie von 13 TeV. Die Analyse ist noch nicht abgeschlossen und noch blind. Die gemessenen Präzessionen der *CP* Parameter S und C in den analysierten B_d^0 Zerfällen sind

$$\sigma(S) = 0.0133\,(\text{stat}) \pm 0.0063\,(\text{syst}), \quad \sigma(C) = 0.0125\,(\text{stat}) \pm 0.0042\,(\text{syst}).$$

Diese Einzelmessung wird eine bessere Präzision als der aktuelle Weltmittelwert haben.

Contents

1 Introduction

The interactions on the smallest scales are described by the Standard Model of particle physics (SM) [1, 2, 3]. Although the SM is very successful in describing these interactions and no effects contradicting it have been experimentally found, there are still open questions. What is dark matter? How can gravity be incorporated in the SM? Why are we living in a world dominated by matter, although matter and antimatter were produced in equal amounts in the Big Bang?

These questions cannot be answered by the SM in its current form. Further, there are about 19 free parameters in the SM that cannot be predicted by the theory. Therefore, it is necessary to determine these parameters experimentally. With these, the SM can predict many more observables, which can be measured precisely and compared to the theory.

Modern particle physics experiments involve accelerator experiments, where light particles, mainly electrons and/or protons, are collided in a well-defined environment. In these collisions, heavier particles are produced, whose decays are detected in large detectors. The largest particle accelerator is the Large Hadron Collider (LHC) at CERN, where protons are collided with a centre-of-mass energies of up to 13 TeV. One of the four large particle detectors at the LHC is the LHC beauty (LHCb) experiment, whose data is analysed in this thesis.

This thesis describes two measurements using very similar decays of a b hadron (Λ_b^0 or B^0) to a ψ meson and a longer-lived neutral s hadron (Λ or K_S^0) to measure different parameters in the SM. The first is a measurement of the ratio of branching fractions of $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$ decays, which confirms a tension seen by the ATLAS experiment [4] to a theory prediction [5, 6] and helps to constrain form factor models. Experimentally, the knowledge of the branching fraction of a decay is important to estimate the background contributions from the decay in other measurements. This analysis is published in the Journal of High Energy Physics as [7]

> R. Aaij et al. "Measurement of the ratio of branching fractions of the decays $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$". In: JHEP 03 (2019), p. 126. doi: 10.1007/JHEP03(2019)126. arXiv: 1902.02092 [hep-ex].

The second analysis involves a time-dependent CP asymmetry measurement in $B^0 \to \psi K_S^0$ decays. CP is the combination of charge and parity inversion, that corresponds to changing a particle to its antiparticle. Hence, the difference between particle and antiparticle in this decay are measured. CP violation is described

within the SM, but its amount is too little to explain the matter-antimatter asymmetry in the universe. The amount of CP violation in the SM is included in the Cabibbo–Kobayashi–Maskawa (CKM) matrix [8, 9], whose elements are free parameters in the SM. The CKM matrix has to be unitary, which can be represented in unitarity triangles. By measuring the angles as well as the length of the sides these triangles, they can be over-constrained and a fundamental property of the SM can be tested. In $B^0 \to \psi K^0_S$ decays one of these angles, β, or rather the derived quantity $\sin(2\beta)$, is measured. The decay $B^0 \to J/\psi K^0_S$ is the golden mode to measure β as it is experimentally and theoretically very clean. Historically, the discovery of CP violation in B^0 mesons with the measurement of $\sin(2\beta)$ by BaBar and Belle [10, 11], which confirmed the CKM mechanism, led to the Nobel Prizes for Kobayashi and Maskawa in 2008. Today, β is the most precisely known CKM angle. The current world average is $\sin(2\beta)_{\text{world average}} = 0.699 \pm 0.017$ [12], and the ongoing analysis presented in this thesis is expected to result in a better precision than the current world average. This measurement is done in close cooperation with Vukan Jevtić and Gerwin Meier. To give a full picture of the analysis, all parts are described, while an emphasis is placed on the parts, which are mainly done by the author. The analysis is at an advanced stage but still ongoing and blinded.

The thesis is structured by introducing the SM and the quantities measured in the two analyses in Chap. 2. This is followed by a short description of the LHCb detector and common analysis methods in Chaps. 3 and 4. Then the analysis of the ratio of branching fractions in $\Lambda^0_b \to \psi(2S)\Lambda$ and $\Lambda^0_b \to J/\psi\Lambda$ decays is presented in detail in Chap. 5. The measurement of CPV in $B^0 \to \psi K^0_S$ decays is described in Chap. 6. The thesis closes with a conclusion and an outlook for the ongoing measurement (Chap. 7).

2 Standard Model of Particle Physics

The Standard Model of particle physics (SM) [1, 2, 3] describes all fundamental particles and known forces, besides gravity. Although the SM is well tested and only some hints to physics beyond the SM are found [13, 14], it does not describe every process in the universe. Aside from the large known matter-antimatter asymmetry in the universe, there is also no dark matter or dark energy in the SM, which accounts for 95% of the energy in the universe [15]. Thus, the SM only describes around 5% of our universe.

2.1 Fundamental Particles and Forces

The particles in the SM can be split into fermions (half-integer spin) and bosons (integer spin). The observed matter in the universe consists of fermions, while bosons are the carriers of the fundamental forces. The fermions can be divided into leptons and quarks. Each of these groups can further be divided depending on their electric charge: charged leptons are the electron (e^-), muon (μ^-) and tauon (τ^-) and neutral leptons are the corresponding neutrinos (ν_e,ν_μ, ν_τ) for each of the charged leptons. In the quark sector, all quarks are charged. Two different charges are possible: the up-type quarks up (u), charm (c), top (t) with charge +2/3e, and the down-type quarks, down (d), strange (s), bottom (b) with charge equal to -1/3e. Leptons and quarks can be grouped in three families with a very distinct mass hierarchy, ranging from some eV/c^2 for the up and down quark to 180 GeV/c^2 for the top quark [16].

Bosons are the mediators of the forces in the SM. The neutral and massless photon is the mediator of the electromagnetic force, which couples to electric charge. At high energies, the electromagnetic force is unified with the weak force, which has three massive mediators in the neutral Z and the two charged $W^\pm$ bosons. The strong force couples at the so-called colour charge and is mediated through eight massless gluons. There are three kinds of colour charge and each colour has an corresponding anti-colour. Colour charge is only carried by quarks and gluons. Each quark carries one colour charge, while an anti-quark carries an anti-colour. In contrast gluons contain both a colour and an anti-colour. An important effect is confinement, which describes that only colour neutral bounded states can exists, therefore no single quark can be measured. This colour neutral bonded state can be a combination of a quark and an anti-quark, which is called a meson or a combination of three quarks or anti-quarks, which is called a baryon. There are recently measurements of more

exotic tetraquarks [17, 18] and pentaquarks [19], which consists of a combination of four and five quarks and anti-quarks, respectively. The last predicted elementary boson in the SM, is the Higgs Boson, which was discovered in 2012 at the LHC [20, 21]. The Higgs mechanism [22, 23] gives masses to the fermions. In Fig. 2.1 a list of all particles in the SM, with its basic properties are given.

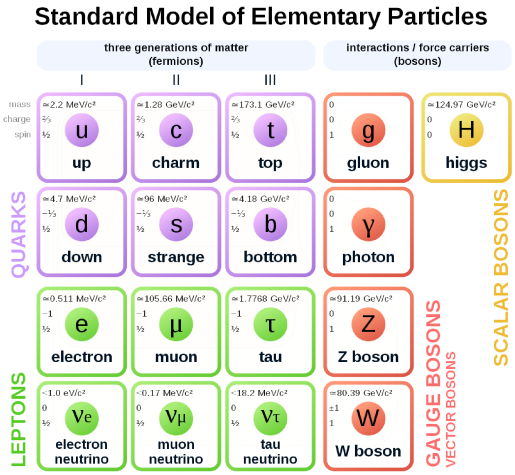

Figure 2.1: Table of the elementary particles in the SM, with their mass, charge and spins taken from [16].

2.2 Feynman Diagrams and Form Factors

The transition rate of an initial state $|i\rangle$ to a final state $|f\rangle$ is given by Fermi's golden rule [24] through

$$\text{transition rate} \propto |A|^2 \times \text{phase space}, \tag{2.1}$$

where A is the transition amplitude, that contains the dynamic of the transition, while all kinematic properties are described by the phase space. The amplitude can be determined for particle transitions using calculations defined by Feynman diagrams. In Fig. 2.2 the leading Feynman diagrams of the different decays studied in this thesis are given. All decays are $b \to ccs$ transitions through the weak interaction.

Besides this leading diagrams so-called penguin diagrams [25, 26] contribute to the transitions amplitudes. Penguin diagrams can include flavour changing neutral currents, where in a loop a quark emits and absorbs a $W^{\pm}$, leading to a neutral

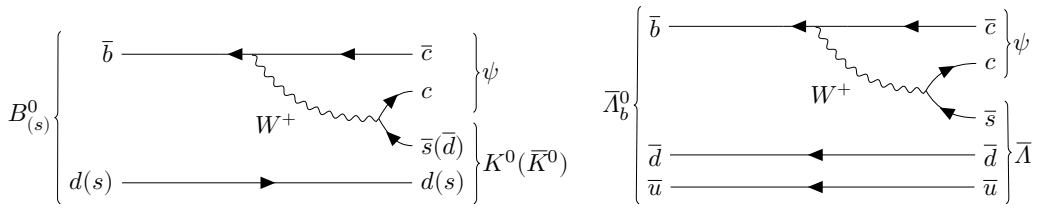

Figure 2.2: Tree diagrams for $B^0 \to \psi K^0_S$ ($B^0_s \to \psi K^0_S$) (left) and $\Lambda^0_b \to \psi\Lambda$ (right).

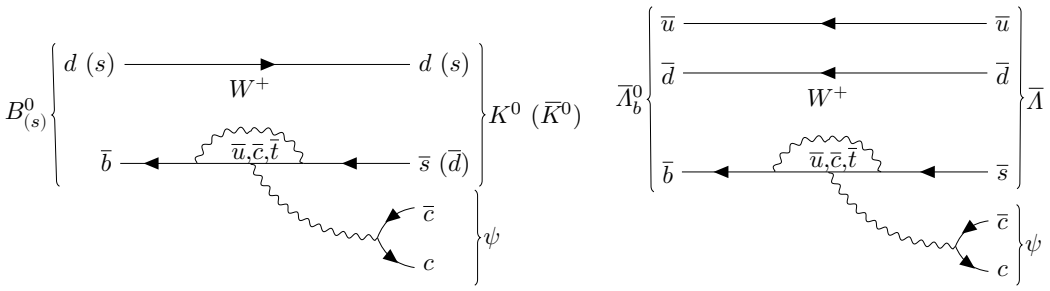

Figure 2.3: Penguin diagrams for $B^0 \to \psi K^0_S$ ($B^0_s \to \psi K^0_S$) (left) and $\Lambda^0_b \to \psi\Lambda$ (right).

flavour change. The corresponding penguin diagrams are shown in Fig. 2.3. Theses penguin diagrams are doubly cabbibo-suppressed in the analysed decays. Thus, the penguins have a smaller effect on the decay rates. However, since different phases occur in these penguin decays, they can affect the measured *CP* violation, which is further discussed in Sec. 2.3.2. The involved vertices are the very same in the analysed B^0_d and Λ^0_b decays. Nevertheless, since hadrons are measured and not single quarks, the transition amplitudes differ between these decays. This is included in the theory by introducing form factors. There are different approaches to determine form factors [27, 28, 29, 30]. To validate these different approaches, experimental input is need. This is done experimentally by measuring branching fractions. The branching fraction, $\mathcal{B}(H \to f)$, is defined as the rate of the hadron, H, to the measured final state, f,

$$\mathcal{B}(H \to f) = \frac{N_{H\to f}}{N_H}, \tag{2.2}$$

where $N_{H\to f}$ are the number of decays of H to f and N_H is the total number of H. In this thesis H is a ground state beauty hadron (B^0 or Λ^0_b). In an accelerator experiment the total number of produced H is given by the integrated luminosity, $\mathcal{L}$ and the production cross section of the hadron. In proton-proton collisions at the

LHC, b quarks are produced in pairs. Therefore, the production cross section divides further in the production cross section of a $b\bar{b}$ pair, $\sigma_{b\bar{b}}$, and the hadronisation fraction to H, f_H, as

$$N_H = \mathcal{L} 2 f_H \sigma_{b\bar{b}}. \tag{2.3}$$

The factor two is included as in the presented measurement, the decay $H \to f$ as well as its charge conjugated decay are used. In an experiment, not all decays can be reconstructed and measured, which leads to an efficiency, ϵ, being needed. With this efficiency, the branching fraction can be measured as

$$\mathcal{B}(H \to f) = \frac{N_{\text{measured}}}{\epsilon \mathcal{L} 2 f_H \sigma_{b\bar{b}}}. \tag{2.4}$$

To eliminate uncertainties from the total number of H, ratios of branching fractions are measured. Thus, the ratio of branching fractions of $H \to f_1$ over $H \to f_2$ is given as

$$\frac{\mathcal{B}(H \to f_1)}{\mathcal{B}(H \to f_2)} = \frac{N_{\text{measured},H\to f_1} \epsilon_{H \to f_2}}{N_{\text{measured},H\to f_2} \epsilon_{H \to f_1}}. \tag{2.5}$$

When the normalisation channel ($H \to f_2$) is a channel, where a large number of decays can be measured, the added uncertainty from the number of measured candidates is negligible. The efficiency is usually taken from simulation. Systematic uncertainties from mismatch between simulation and data, cancel largely in the ratio of efficiencies as long as both decay channels are similar.

2.3 CKM Matrix and *CP* Violation

The change of quark-flavour is possible in the SM through weak charged currents. The process is described in the CKM matrix, which describes the connection between the mass eigenstates q' and the flavour eigenstate q as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.6}$$

Since the matrix element V_{ij} enters in the transition amplitude of quark i to quark j, the transition probability is proportional to $|V_{ij}|^2$ as described in Eq. (2.1). Thus, the CKM matrix has to be unitary to fulfil probability conservation. A general $N \times N$ matrix has $2N^2$ parameters. Further, a unitary matrix has to fulfil the

condition

$$V_{\mathrm{CKM}}^{\dagger} V_{\mathrm{CKM}} = I, \tag{2.7}$$

which constrains N^2 parameters. $2N-1$ of these parameters are relative phases and can be absorbed in the quark fields. This leads to four independent parameters in case of the 3×3 CKM matrix. These parameters are 3 angles and one complex phase. The CKM matrix has a strong hierarchical structure, with the diagonal elements being close to unity. This structure can be shown in the Wolfenstein parametrization [31], where the CKM elements are developed in terms of $\lambda = V_{us} = 0.224837^{+0.000251}_{-0.000060}$ [32] as

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - \mathrm{i}\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-\mathrm{i}\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4), \tag{2.8}$$

with $A = 0.8235^{+0.0056}_{-0.0145}$, $\rho = 0.1569^{+0.0102}_{-0.0061}$ and $\eta = 0.3499^{+0.0079}_{-0.0065}$ [32]. Thus, on the diagonal the elements are 1 with small corrections of $\mathrm{O}(\lambda^2)$, while the off-diagonal elements are of different order of λ.

The unitarity condition of the CKM Matrix (Eq. (2.7)) leads to six orthogonality relations from the off-diagonal elements. Since these orthogonality relations are a sum of three complex numbers equal to zero, they can be interpreted as triangles in a complex plane. Two of these triangles have side length in the same order of λ. One of these is the relation

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

For convenience, all sides are normalised to $V_{cd}V_{cb}^{*}$. The CKM triangle can be over-constrained by measuring the side length and the angles to confirm the unitarity of the CKM matrix. The three angles are defined as

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^{*}}{V_{ud}V_{ub}^{*}}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^{*}}{V_{td}V_{tb}^{*}}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}}\right). \tag{2.10}$$

The current situation of all measurements constraining the CKM triangle combined by the CKMfitter group [32] is shown in Fig. 2.4. All measurements are compatible and the uncertainty of the apex of the triangle is small.

2.3.1 B^0 Meson System

CP violation (*CPV*) is one necessary condition to explain the matter-antimatter

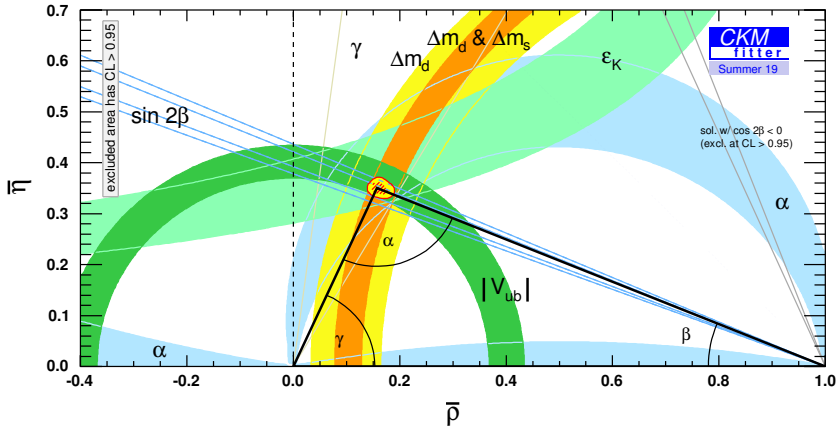

Figure 2.4: Experimental status of the CKM triangle combined by the CKMfitter group [32].

asymmetry in the universe [33]. *CPV* can occur in the decay of a particle to a final state. This thesis presents a measurement of *CPV* in neutral b meson decay, where the neutral is either a B^0 or B_s^0 meson. The B^0 (B_s^0) meson contains a $\overline{b}$ quark and a d (s) quark, while its *CP* conjugated anti-particle the $\overline{B}^0$ ($\overline{B}_s^0$) contains a b quark and a $\overline{d}$ ($\overline{s}$) quark. To simplify the notation in this section both mesons are commonly noted as B^0, while the explicit notation (B_d^0 or B_s^0) are used if needed. The theory is valid for every neutral meson. A B^0 meson can oscillate into $\overline{B}^0$ meson and vice versa, which in the SM is possible in quantum loops of charged currents. The leading Feynman diagrams are box diagrams shown in Fig. 2.5.

The top quark dominates these loops compared to the up and charm quark because of its large mass and the involved Inami-Lim functions [34]. The time evolution of a initial B^0 ($\overline{B}^0$) state can be derived by solving the effective Schrödinger equation [35]

$$i\frac{\mathrm{d}}{\mathrm{dt}}\begin{pmatrix}|B^0\rangle\\|\overline{B}^0\rangle\end{pmatrix} = \mathbf{H}\begin{pmatrix}|B^0\rangle\\|\overline{B}^0\rangle\end{pmatrix} = \left(\mathbf{M} - \mathrm{i}\frac{\Gamma}{2}\right)\begin{pmatrix}|B^0\rangle\\|\overline{B}^0\rangle\end{pmatrix}, \tag{2.11}$$

where $\mathbf{H}$ has to be non-hermitian because the mesons will decay in time. $\mathbf{M}$ and Γ are hermitian. *CPT* invariance leads to B^0 and $\overline{B}^0$ having the same mass and lifetime. Thus, the diagonal elements in $\mathbf{H}$ have to be the same and $\mathbf{H}$ can be

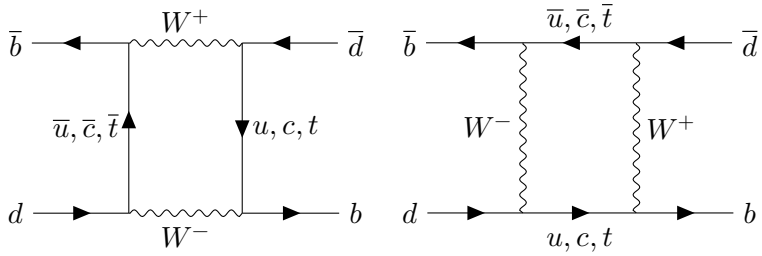

Figure 2.5: Box diagrams of the B^0-$\overline{B}^0$ mixing.

defined as

$$\mathbf{H} = \begin{pmatrix} m - \frac{i\Gamma}{2} & M_{12} - \frac{i\Gamma_{12}}{2} \\ M_{12}^* - \frac{i\Gamma_{12}^*}{2} & m - \frac{i\Gamma}{2} \end{pmatrix}. \tag{2.12}$$

The off-diagonal elements describe the transition of $B^0 \leftrightarrow \overline{B}^0$, where M_{12} describes the short distance off-shell transitions and Γ_{12} describes the virtual intermediate decay to final states, that are accessible from B^0 and $\overline{B}^0$. The eigenvectors of $\mathbf{H}$ are defined as the light and heavy mass eigenstates B^0_{H} and B^0_{L}

$$\begin{aligned} |B^0_{\mathrm{H}}\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle, \\ |B^0_{\mathrm{L}}\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle, \end{aligned} \tag{2.13}$$

where p and q are complex numbers that fulfil the normalisation $|p^2| + |q^2| = 1$. The mass eigenstates have defined masses m_L, m_H and decay width Γ_L, Γ_H. Thus, it is useful to define the differences as

$$\Delta m = m_H - m_L \qquad\qquad \Delta\Gamma = \Gamma_L - \Gamma_H, \tag{2.14}$$

and the average mass, decay width as

$$m = \frac{m_H + m_L}{2} \qquad\qquad \Gamma = \frac{\Gamma_L + \Gamma_H}{2}. \tag{2.15}$$

With these quantities, the ratio of the mixing parameters p/q can be written as

$$\frac{p}{q} = \frac{\Delta m + \mathrm{i}/2\Delta\Gamma}{2(M_{12} - \mathrm{i}/2\Gamma_{12})}. \tag{2.16}$$

The time evolution for the mass eigenstate follows as

$$|B^0_{\mathrm{H}}(t)\rangle = \exp^{-\mathrm{i}m_H t}\exp^{-\mathrm{i}\mathrm{i}/2\Gamma_H t}\,|B^0_{\mathrm{H}}\rangle, \tag{2.17}$$

$$|B^0_{\mathrm{L}}(t)\rangle = \exp^{-\mathrm{i}m_L t}\exp^{-\mathrm{i}\mathrm{i}/2\Gamma_L t}\,|B^0_{\mathrm{L}}\rangle. \tag{2.18}$$

Using this time evolution in Eq. (2.13), the time evolution for a pure B^0 or $\overline{B}^0$ state at $t = 0$ is given as

$$|\overline{B}^0(t)\rangle = g_+(t)|B^0\rangle - \frac{q}{p}g_-(t)|\overline{B}^0\rangle, \tag{2.19}$$

$$|B^0(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\overline{B}^0\rangle, \tag{2.20}$$

with $g_\pm(t)$ defined as

$$g_\pm(t) = \frac{1}{2}\left(\exp^{\mathrm{i}m_H t}\exp^{-\frac{\Gamma_H}{2}t} \pm \exp^{\mathrm{i}m_L t}\exp^{-\frac{\Gamma_L}{2}t}\right). \tag{2.21}$$

With this the time-dependent decay rates of an initial B^0 or $\overline{B}^0$ decaying into a common final state f can be calculated as the square of the corresponding amplitudes as

$$\Gamma(B^0(t) \to f) = |\langle f|H|B^0\rangle|^2, \quad \Gamma(\overline{B}^0(t) \to f) = |\langle \bar{f}|H|B^0\rangle|^2. \tag{2.22}$$

A useful quantity for the following calculation is λ_f defined as

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}, \tag{2.23}$$

where A_f ($\overline{A}_f$) is the amplitude of a B^0 ($\overline{B}^0$) decaying to a final state f. With this the decay rates are given by

$$\Gamma(B^0(t) \to f) = |A_f|^2 \qquad \left(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 - 2\mathrm{Re}\left(\lambda_f g^*_+(t)g_-(t)\right)\right), \tag{2.24}$$

$$\Gamma(\overline{B}^0(t) \to f) = |A_f|^2\left|\frac{p}{q}\right|^2\left(|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 + 2\mathrm{Re}\left(\lambda_f g_+(t)g^*_-(t)\right)\right). \tag{2.25}$$

Expressing $|g_\pm(t)|^2$ and $g_+(t)g^*_-(t)$ with trigonometric and hyperbolic functions, the

time-dependent decay rates can be written as

$$\begin{aligned}\Gamma(B^0(t)\to f) = \frac{1}{2}|A_f|^2\left(1+|\lambda_f|^2\right) \qquad & \exp^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right)+A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right)\right.\\ & \left.+C_f\cos(\Delta m t)-S_f\sin(\Delta m t)\right),\\ \Gamma(\overline{B}^0(t)\to f) = \frac{1}{2}|A_f|^2\left(1+|\lambda_f|^2\right)\left|\frac{p}{q}\right|^2 & \exp^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right)+A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right)\right.\\ & \left.-C_f\cos(\Delta m t)+S_f\sin(\Delta m t)\right),\end{aligned} \tag{2.26}$$

with the *CP* observables defined as

$$A_f^{\Delta\Gamma} = -\frac{2\mathrm{Re}\lambda_f}{1+|\lambda_f|^2}, \qquad C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \qquad S_f = \frac{2\mathrm{Im}\lambda_f}{1+|\lambda_f|^2}, \tag{2.27}$$

which fulfil the relation

$$\left(A_f^{\Delta\Gamma}\right)^2+\left(C_f\right)^2+\left(S_f\right)^2 = 1. \tag{2.28}$$

The cos and cosh terms correspond to decays with and without oscillation of the B^0, in contrast the sin and sinh terms correspond to interference of the decay and decay after mixing.

2.3.2 *CPV* in neutral $B^0 \to \psi K_S^0$ Decays

The following discussion focuses on the *CPV* in $B^0 \to \psi K_S^0$ decays, which is a *CP* eigenstate in the final state ($f = \bar{f}$). Three different types of *CPV* can occur in neutral B^0 decays, which all arise from the interference of different phases in the SM: direct *CP* violation, indirect *CP* violation and *CP* violation in the interference between decay and decay after mixing. There are three different types of these phases: so-called weak phases that changes its sign under *CP* transformation (*CP*-odd) strong phases keep their sign under *CP* transformation (*CP*-even) and spurious global phases, which are convention-dependent and thus cannot be measured. The complex phase in the CKM matrix is a weak phase, while the strong phases usually enter from final-state-interaction scatterings of on-shell states, which can come from electromagnetic or strong interactions. Direct *CPV* describes *CPV* in the decay, which means that the absolute value of the decay amplitude is different between a

$B^0 \to f$ transition compared to the *CP* conjugated $\overline{B}^0 \to \bar{f}$ transition

$$\left|\frac{A_f}{\overline{A}_{\bar{f}}}\right| \neq 1. \tag{2.29}$$

At least a weak and a strong phase has to be present for direct *CPV* to occur otherwise the total phase does not change under *CP* transformation. As this kind of *CPV* does not involve meson mixing, it is the only *CPV* that can be measured in the decay of charged mesons [36]. Direct *CPV* is negligible in $B^0 \to \psi K^0_S$ decays as the involved weak phase, ϕ_{weak} is much larger than the involved strong phase ϕ_{strong}. This results in

$$CP(\phi_{\text{weak}} + \phi_{\text{strong}}) \approx -\phi_{\text{weak}} \tag{2.30}$$

and thus $A_f \approx -\overline{A}_{\bar{f}}$ or rather $A_f \approx -\overline{A}_f$.

Indirect *CPV* occur when the mixing rates between $B^0 \to \overline{B}^0$ and $B^0 \to \overline{B}^0$ differ. As shown in Eq. (2.12), these transitions are given by the off-diagonal elements of $\mathbf{H}$, it can be shown that this implies

$$\left|\frac{q}{p}\right| \neq 1. \tag{2.31}$$

In the SM indirect *CPV* in B^0 mesons is expected to be small [37], which is confirmed by experimental measurements [38, 36]. The measurement are performed using flavour specific decays like semi-leptonic decays, where the lepton-flavour of the final state defines the decay flavour of the B^0 [39].

The *CPV* in the interference of decay and decay after mixing is the dominant kind of *CPV* in $B^0 \to \psi K^0_S$ decays. This occurs when a final state f can be reached by B^0 as well as $\overline{B}^0$. Thus, the amplitudes of the direct decay and the decay after mixing interfere. The size of the interference term depends on the imaginary part of λ ($\text{Im}\lambda \neq 0$) and not on the absolute value of λ. Therefore, *CPV* in the interference can occur even if neither direct nor indirect *CPV* is present. The following equations simplify largely for $B^0 \to \psi K^0_S$ decays as the final state f is a *CP* eigenstate and therefore $f = \bar{f} = f_{CP}$, with the *CP* eigenvalue η_f. Here, the final state ψK^0_S has an eigenvalue $\eta_f = 1$ (when only considering its final state). To measure *CPV* in the interference, the time-dependent *CP* asymmetry can be defined as

$$\mathcal{A}_{CP} = \frac{\Gamma(\overline{B}^0 \to f_{CP}) - \Gamma(B^0 \to f_{CP})}{\Gamma(\overline{B}^0 \to f_{CP}) + \Gamma(B^0 \to f_{CP})}. \tag{2.32}$$

Here, the B^0, $\overline{B}^0$ corresponds to an initial ($t = 0$) pure B^0, $\overline{B}^0$ meson state. Using

Eq. (2.26), $\mathcal{A}_{CP}$ can be written as

$$\mathcal{A}_{CP} = \frac{S \cdot \sin(\Delta m t) - C \cdot \cos(\Delta m t)}{\mathcal{A}_{\Delta\Gamma} \cdot \sinh\left(\frac{\Delta\Gamma}{2}t\right) + \cosh\left(\frac{\Delta\Gamma}{2}t\right)}. \tag{2.33}$$

Comparing the definitions of the *CP* parameters S, C and $\mathcal{A}_{\Delta\Gamma}$ (Eq. (2.27)) with the conditions for the three kinds of *CPV*, C vanishes if no direct or mixing *CPV* is present and the same is true for S if no interference *CPV* is present.

As aforementioned direct *CP* and indirect *CP* are negligible for $B^0 \to \psi K^0_S$ decays. Further, in the B^0_d system $\Delta\Gamma_d$ is negligible [39], which simplifies $\mathcal{A}_{CP}$ further to

$$\mathcal{A}_{CP} = S \cdot \sin(\Delta m t) - C \cdot \cos(\Delta m t). \tag{2.34}$$

To link the *CP* parameter to the *CP* violating phase in the CKM matrix, the decay $B^0 \to \psi K^0_S$ is further described in the following. The decay $B^0 \to \psi K^0_S$ is dominated by a Cabibbo-favoured $b \to c\bar{c}s$ transition. The corresponding Feynmann diagrams (Fig. 2.2) show that the flavour eigenstates K^0, $\overline{K}^0$ are involved instead of the mass eigenstate K^0_S. Nevertheless, the K^0_S is a superposition of K^0 and $\overline{K}^0$ similar to B^0_H, B^0_L in the B^0 system (Eq. (2.13)). Thus, the ratio of decay amplitudes can be written using the factors p_{K^0}, q_{K^0}, which are defined analogously to the B^0 system, as

$$\frac{\overline{A}_{J/\psi K^0_S}}{A_{J/\psi K^0_S}} = -\frac{p_{K^0}}{q_{K^0}} \frac{\overline{A}_{J/\psi \overline{K}^0}}{A_{J/\psi K^0}}. \tag{2.35}$$

The ratio of p_{K^0} and q_{K^0} is determined in a similar way as in the B^0 system but in the K^0 system the leading contribution is from a charm quark in the loop instead of a top quark. This calculation leads to

$$\frac{p_{K^0}}{q_{K^0}} = -\frac{V_{cs}V^*_{cd}}{V^*_{cs}V_{cd}}. \tag{2.36}$$

Neglecting non-tree contributions to the decay, the ratio of decay amplitudes is given by the involved CKM elements

$$\frac{\overline{A}_{J/\psi \overline{K}^0}}{A_{J/\psi K^0}} = \frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}}. \tag{2.37}$$

Using these values and taking the CKM elements from the B^0 mixing into account,

$\lambda_{J/\psi K^0_S}$ is given by

$$\lambda_{J/\psi K^0_S} = -\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\frac{V_{cs}V^*_{cd}}{V^*_{cs}V_{cd}}\frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}} = -\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\frac{V^*_{cd}V_{cb}}{V_{cd}V^*_{cb}}. \tag{2.38}$$

Due to neglecting each contribution besides the tree, the absolute value of $\lambda_{J/\psi K^0_S}$ is unity and as mentioned before, direct *CP* is not present ($C_{J/\psi K^0_S} = 0$). Thus, S can be computed from Eq. (2.27) as

$$S_{J/\psi K^0_S} = \frac{2\mathrm{Im}(\lambda_{J/\psi K^0_S})}{2} = \sin\left(\arg\left(-\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\frac{V^*_{cd}V_{cb}}{V_{cd}V^*_{cb}}\right)\right) = \sin\left(\arg\left(-\left(\frac{V^*_{cd}V_{cb}}{V_{tb}V^*_{td}}\right)^2\right)\right), \tag{2.39}$$

which is $\sin(2\beta)$ as defined in the CKM triangle (Fig. 2.4). Therefore, $B^0 \to J/\psi K^0_S$ is the golden mode to measure the CKM angle $\sin(2\beta)$. The calculation is the very same for higher charmonium resonances as in the decay $B^0 \to \psi(2S)K^0_S$.

For $B^0_s \to J/\psi K^0_S$ decays, the same calculation can be made by switching each d to an s, which would result in measuring ϕ_s the mixing angle in the B^0_s system. However, the assumption that only the tree contribution is relevant for the decay does not hold, since the $B^0_s \to J/\psi K^0_S$ decay is Cabibbo-suppressed. Thus, higher order contributions from penguin decays contribute significantly to the decay. These contributions result in a shift of the measured phases. Thus experimentally an effective phase, ϕ_{eff}, defined as

$$\phi_{\mathrm{eff}} = \phi + \Delta\phi. \tag{2.40}$$

As the involved Feynman diagrams of $B^0_s \to J/\psi K^0_S$ and $B^0_d \to \psi K^0_S$ decays are connected via U-spin symmetry, the relative large measured shift in B^0_s decays can be used to constrain the effects of the penguin contribution in $B^0 \to J/\psi K^0_S$ decays [40, 41]. In case of $B^0_d \to \psi K^0_S$ decays, where the penguin contribution is suppressed, the shift is estimated to be at the order of 1° depending on the used method [42, 43, 41]. As the experimental sensitivity gets better with the coming LHCb measurement and with future Belle II results, these penguin contributions will become significant in $\sin(2\beta)$ measurements and the measurement of *CPV* in $B^0_s \to J/\psi K^0_S$ decays will become more important.

3 The LHC and the LHCb Detector

The LHC at CERN near Geneva is the world largest collider with a circumference of 27.3 km. Protons or heavy ions, mainly lead, are accelerated up to a centre-of-mass energy of 13 TeV and are collided inside four large experiments ATLAS [44], CMS [45], ALICE [46] and LHCb [47]. ATLAS and CMS are multi-purpose experiments, while ALICE is specialised for the measurement of quark-gluon plasma, which is created in heavy ion collisions. Since the LHC is not able to accelerate the hadrons from rest to their final energy, multiple pre-accelerators are used. The whole LHC accelerator complex is shown in Fig. 3.1.

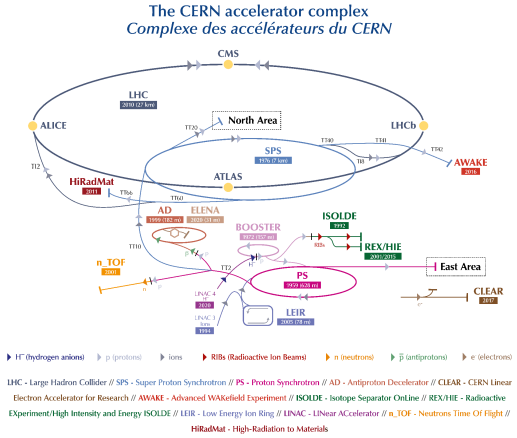

Figure 3.1: The CERN accelerator complex, with all pre-accelerators that end up in the LHC [48].

The LHCb experiment focuses on precision flavour physics measurements, ranging from the search for very rare decays to the measurement of CP violation. These measurements are performed analysing the decays of c or b hadrons. Since c and b hadrons are light (2-6 GeV/c^2) compared to the large centre-of-mass energy of the LHC, most produced c and b hadrons are highly boosted and thus fly and decay near the beam pipe. Therefore, LHCb is constructed as a one-arm forward spectrometer covering a pseudo-rapidity, η, from 2 to 5.

3.1 The LHCb Detector

The LHCb detector consists of multiple sub-detectors, which can be divided into detectors for particle tracking and for particle identification. A sketch of the LHCb detector [47, 49] in the LHC Run 1 (2011-2012) and Run 2 (2015-2018) is shown in Fig. 3.2 and discussed in the following.

The tracking detectors measure the points where a charged particles traversed them. This information is used to reconstruct the particle's trajectory. Using the information that the charged particle track will bend in the known magnetic field of LHCb, the particle's momentum can be determined. The tracking system is described in the following starting from the interaction point downwards the detector.

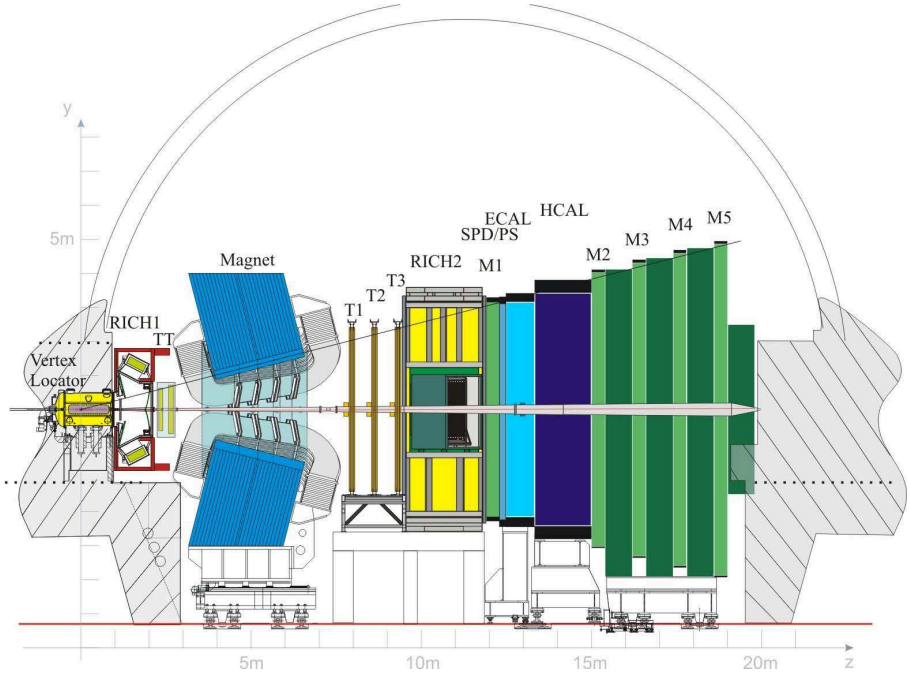

Figure 3.2: The LHCb detector with its subdetectors [47]. The tracking system consists of the Vertex Detector (VELO), the Tracker Turicensis (TT) and the tracking stations (T1-T3) with the magnet and the muons stations (M1-M5). The particle identification consists of two the Ring Imaging Cherenkov detectors (RICH I/II), the calorimeter system (SPD/PS, ECAL, HCAL).

The Vertex Locator (VELO) [50] is the sub-detector that is the nearest to the beam pipe. Its main purpose is to detect displaced secondary vertices (SV) from the decay of b and c hadrons, which have a mean flight distance of a few millimetres.

It is a silicon stripe detector consisting of 42 modules divided on both sides of the beam pipe. The detector modules are vertically placed along the beam pipe with increasing distances further from the collision point. To secure the VELO from unstable beam conditions, which can happen during the injecting phase of the LHC, it is able to be moved away from the beam pipe. When the beam is stable, the modules can be moved as near as 7 mm to the beam pipe. The VELO helps to distinguish the primary vertex (PV), which is the production vertex of these hadrons, from the SV, which is its decay vertex. Further, the VELO helps to assign the correct PV to the candidate [51]. In the beam crossing multiple PVs (inelastic collisions) can be present with most collisions having one or two PVs. A wrongly assigned PV leads can lead to a wrong reconstructed flight distance.

Downstream of the VELO is the Tracker Turicensis (TT) [52], which is a silicon strip detector right in front of the magnet. It is 1.5 m wide and 1.3 m in height.

The TT is located right in front of the magnet. The magnet provides a magnetic power of 2Tm and bends charged particles traversing the magnet field.

Following the magnet, there is the Inner Tracker (IT) and the Outer Tracker (OT) [53, 54]. The IT very similar to the TT. Is is a silicon strip detector around the beam pipe with dimensions of $1.2 \times 0.4\,\mathrm{m}^2$. The OT is located next to the IT in the region with lower occupation further from the beam pipe. It consists of gas-filled straw-tube modules and covers a much larger area of $6 \times 5\,\mathrm{m}^2$. IT and OT are grouped in three stations each consisting of four detector layers. The two middle layers are tilted by $\pm 5°$ to ensure two dimensional resolution of the particles.

The last tracking detectors are the muon chambers [55] before and after the calorimeter system. The muons chambers are multi-wires-proportional-chambers, that are used to determine the momentum of muons for the hardware trigger [56]. Further, it is also used as particle identification detector as mainly muons are able to traverse the calorimeter systems and leave a signature in the muons chambers. Therefore, muons are the best particles to use for triggering the events, which is one reason that most decay channel analysed in this thesis contain a pair of muons in the final state.

Other particle identification sub-detectors are the Ring Imaging Cherenkov detectors (RICH) [57], which use the Cherenkov effect [58] to measure the velocity of a charged particle. The Cherenkov effect occurs when a charged particle traverses material with a higher velocity than the speed of light in that material, then light is emitted in a cone, where the opening angle is proportional to the velocity. These light crones are reflected to hybrid photon detectors. The resulting rings are fitted and the velocity information is determined. Using this velocity information and the momentum measurement from the tracking system leads in principle to a determination of the

mass, which is unique for a particle. There is one RICH detector (RICH I) in front of the TT, optimised for lower momentum particles, that are possibly bent out of the detector acceptance by the magnet downstream. A second RICH detector (RICH II) is located directly in front of the first muon chamber and the calorimeter system. With the RICH system, LHCb is able to distinguish light hadrons (mainly pions from kaons and protons), which is a unique ability at the LHC. In practice, the circles measured by the photon detectors are fitted with the track information of the other detectors with different particle hypothesises. From these fits a commonly used variable are PID variables like difference of the log likelihood (DLL) between a particular particle hypothesis (*e.g.* proton or kaon) compared to a fit with a pion hypothesis. The pion hypothesis is used because a pion is the most commonly particle in the detector. A more advanced method is to use neural networks, which use addition information from the detector to discriminate between stable light particles. The different performance for the different particles and momentum is shown in Fig. 3.3.

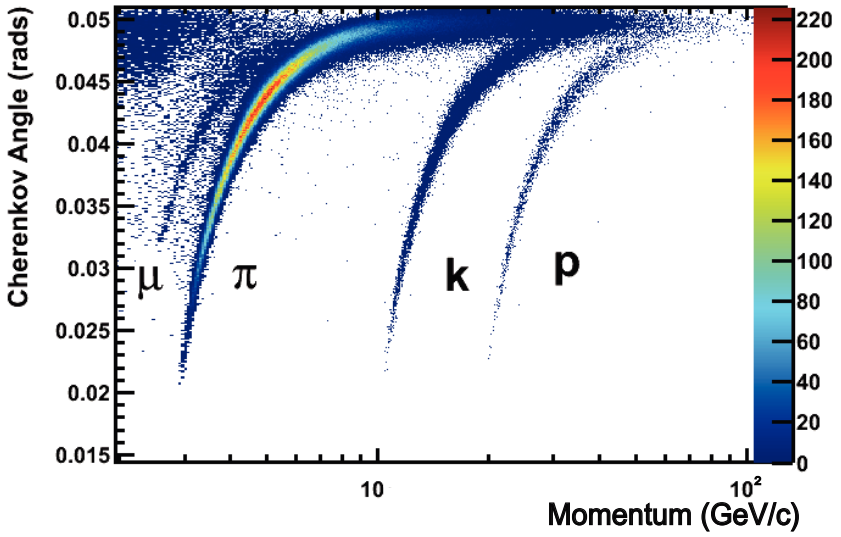

Figure 3.3: Reconstructed Cherenkov angle as a function of track momentum in the C_4F_{10} radiator for different light charged particles [57].

The last part of the particle identification system is the calorimeter system, consisting of the Silicon Pad Detector (SPD), preshower detector (PS), the electromagnetical calorimeter (ECAL), and the hadronic calorimeter (HCAL). Each of these detectors use scintillation to measure the energy deposition. The calorimeter system is mainly

used in the trigger and for the election identification as well as measurements including neutral particles in the final state. In the SPD only charged particles are detected, which helps to distinguish electrons from photons. The following PS improves the identification of electrons and helps to distinguish them from other light charged particles (e.g charged pions). The ECAL and the HCAL both consist of alternating layers of iron and scintillating tiles, which measure the energy deposition from the particle showers inside. Different particles types produce different shapes of the showers, which can help to identify the particles that have been reconstructed.

3.2 Track Reconstruction and Track Types

The reconstruction of the decay chain is done in LHCb with a bottom to top approach. Starting with the stable final state particles, intermediate particles are combined till the head of the decay chain is reach. Here, stable particles refers to particles that do not decay within the LHCb detector, this includes *e.g.* muons, charged kaons and pions. As example, in $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ decays, the muon and pions are measured by the tracking detector, from which their trajection and their momentum is determined. From the two tracks the common vertex is determined and the momentum of the intermediate resonance (here the J/ψ and the K^0_S meson) is calculated. This step is repeated to form the B^0 meson from the intermediate resonances. Mostly so-called long tracks are used in LHCb analyses. A long track is a track that is measured by all tracking detectors. Both analysed decays in this theses involve a long living strange hadron, namely a K^0_S meson or a Λ baryon. Their lifetime is $89.54 \pm 0.04\,\text{ps}$ and $263.2 \pm 2.0\,\text{ps}$, respectively, which is about a factor 100 larger then the lifetime of B and D mesons. Thus, these two hadrons are unique in the LHCb Detector as their lifetime is not long enough to be stable within the detector but they often decay outside the VELO. In these cases their decay products can only be measured by the TT and the tracking stations. Tracks reconstructed in this way are called downstream track. Thus, the main reconstruction combinations from K^0_S and Λ are the combination of two long track, which is called long or LL, and the combination of two downstream tracks, which is called downstream or DD. In the *CPV* measurement in $B^0 \to J/\psi K^0_S$ decays, two additional combinations are used to maximise the reconstruction efficiency. The combination of a long and a downstream reconstructed pion is used,which is abbreviated as LD. Such combinations can occur when the K^0_S meson decays at the edges of the VELO or very near to the beam pipe and one of the pions does not leave enough hits in the VELO, to be reconstructed. Another possibility is that a true downstream pion is falsely upgraded to a long track pion in the reconstruction, due to unrelated matching hits or noise in the VELO. The relevant properties of such combinations (*e.g.* mass

resolution and vertex resolution) are very similar to the common used combinations. An additional used method of reconstructing the K^0_S meson is the combination of a long track pion and an upstream pion, which is abbreviated as UL. Upstream pions are reconstructed from tracks that have hits in the VELO and the TT but not enough in the other tracking stations. Mostly upstream tracks occur when charged low momentum particles are bent out of the detector acceptance by the magnet. These upstream tracks are rarely used because only a very small magnet field is present between the VELO and the TT. Thus, the bending of charged tracks is small in this region, which leads to a much worse momentum resolution. Nevertheless, the resulting resolution is sufficient to distinguish real K^0_S mesons from random combinations of pions. The poor vertex resolution of upstream-long track candidates is a minor effect, because the secondary vertex, which is the B^0 decay vertex, can be well-reconstructed through the decay of the J/ψ meson. These additional types of K^0_S reconstruction add about 15% candidates. These combinations are not used in the $\Lambda^0_b \to \psi(2S)\Lambda$ branching fraction measurement, because there the reconstruction and selection efficiency directly affects the result. Since these types of reconstruction are not well studied, disagreement between simulation and data is expected. In the *CPV* measurement such disagreement, leads to a not optimal selection but does not bias the result.

3.3 Trigger

It is not possible to store every collision on tape. The data rate is reduce by the trigger system from the 40 MHz bunch crossing rate to 5 and 12.5 kHz in Run I and Run II, respectively. The LHCb trigger system [56] consists of a hardware stage followed by two software stages. In the hardware trigger (L0), basic information from the calorimeter and muon system is used to define a trigger decision. These decisions need to be simple and fast, *e.g.* a L0 muon trigger requires a muon with high momentum present in the collision. These two detector components are used, because they can be readout very fast. The L0 reduces the data rate to 1 MHz. This reduced data rate is further reduced by a two staged software trigger (HLT1 and HLT2), to the final rate that is written to tape. In the software trigger a full event reconstruction is performed. This reconstruction is very similar to the offline reconstruction. This allows to apply efficient trigger lines for example a detached J/ψ line requires two muons with higher momentum from a common vertex, which has to be displaced from the PV and the combined mass of the muons has to be near the known J/ψ mass. One of the main upgrades for LHCb in Run III is the change to a fully software-based trigger, which will lead to more efficient trigger decisions [59].

3.4 Simulation

Simulation is used at multiple different steps in the presented analysis. Signal simulation is used to optimise the signal selection and to determine its signal efficiency. Further, simulation of background decays, which are reconstructed as signal decays are used to extract fit templates or to estimate the need for harder selections. In the simulation, PYTHIA [60] is used to generate pp collisions with a specific LHCb configuration [61]. The decay of non-final state particles is calculated by EVTGEN [62], in which final-state radiation is taken into account by PHOTOS [63]. Simulation without the following steps is called generator level simulation. The GEANT4 toolkit [64, 65] describes the interaction of the generated particles with the detector, and its response. This step is further described in Ref. [66]. After this step, the simulation is treated in the vary same way as recorded data but the information from the generation are kept (*e.g.* the truth particle information, momenta).

4 Common Methods

In the following important techniques that are commonly used in both presented analysis are described. The maximum likelihood fit as method to estimate parameters is described inSec. 4.1. The sPlot method to extract signal weights is discussed inSec. 4.2.

4.1 Maximum Likelihood Fit

Estimating parameters from a data set is an important task in a analysis. Parameters that are estimated in the presented measurements are for example the signal yields or the *CP* parameters. The main method to extract parameters in this thesis is the maximum likelihood fit [67]. For the fit of the observable, x, a probability density function, $\mathcal{P}(x;p)$, with parameter, p, needs to be defined. The PDF can have multiple parameters and a multidimensional fit to multiple observables can be done. Thus, p and x can be vectors. A likelihood function, $\mathcal{L}$, is defined as

$$\mathcal{L}(x|p) = \prod_{i}^{N} P(x_i|p), \tag{4.1}$$

where x_i is the value of the observable for candidate i and N is the total number of data points in the fit. This likelihood function can be modified by a Poisson term $e^n \cdot n^N/N!$ to include the probability to measure N candidates when n are expected. A maximum likelihood fit with this term is called extended maximum likelihood fit. In both cases the likelihood function is maximised by varying the parameters. For technical reasons, usually the negative logarithm of the likelihood function is minimised. The logarithm leads to smaller numbers and to a sum instead of a product, which is computational beneficial, also there are better algorithms to find a minimum than a maximum, hence the negative logarithm is minimised.

Uncertainties from external parameters, y, can be incorporated in the likelihood function by multiplying a Gaussian function $\mathcal{G}(y|\mu,\sigma)$, where μ is the expected value of y and σ its uncertainty. This term gives a penalty to the likelihood function, when the fitted value of y deviates from the expected value.

4.2 sPlot Method

The recorded data sets are never pure signal decays but a mixture of signal and background decays. Thus, if fitting a distribution, *e.g.* the b hadron life time, a description for each component is needed. A model to describe the signal decays is often given by theoretical functions, which have to be modified by experimental effect. However, defining a model for various background components can lead to large systematic uncertainties. The sPlot method is a method that defines signal weights that projects out non-signal components in an observable of interest. To define these weights, a extended fit to an uncorrelated observable, x, is needed, where the distributions of the two components to unfold are known. In practice, a fit to the reconstructed b hadron mass is used, as in this observable the components can be modelled well. From the fit model, $\mathcal{P}(x) = \sum_k \mathcal{P}_k(x)$, signal weights [68, 69], $sw(x_i)$, also called sWeights, are calculated as

$$sw(x_i) = \frac{\sum_j = 1 V_{nj} \mathcal{P}_j(x_i)}{\sum_k = 1 N_k \mathcal{P}_k(x_i)}, \tag{4.2}$$

where the different components are indexed by j and k and V_{nj} is the covariance between the yields of the components and N_k is yield of the component. The distribution of the weighted data sets in the uncorrelated observable of interest corresponds to the signal distribution. This distribution can then be used in a fit to extract parameters or to compare recorded data with simulation. When sweights are used in a fit, the likelihood function has to be modified, otherwise the uncertainties of the fitted parameters are underestimated. Thus, a correction factor, α, defined as

$$\alpha = \frac{\sum_i^N sw(x_i)}{\sum_i^N sw(x_i)^2}. \tag{4.3}$$

This correction factor leads to better coverage but is not necessary correct [70]. Nevertheless, the coverage in this case can be tested using pseudo-experiments.

5 Measurement of $\Lambda_b^0 \to \psi(2S)\Lambda^0$ Decays

The branching fraction of $\Lambda_b^0 \to \psi(2S)\Lambda$ decays is measured relative to branching fraction of well known $\Lambda_b^0 \to J/\psi\,\Lambda$ decays. The ratio of branching fractions is defined as

$$\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)} = \frac{N_{\text{measured},\Lambda_b^0\to\psi(2S)\Lambda} \cdot \epsilon_{\Lambda_b^0\to J/\psi\,\Lambda}}{N_{\text{measured},\Lambda_b^0\to J/\psi\,\Lambda} \cdot \epsilon_{\Lambda_b^0\to\psi(2S)\Lambda}}. \tag{5.1}$$

which is discussed in detail in Sec. 2.2. The relative measurement helps to cancel most experimental systematic uncertainties. Both decays are reconstructed from the same final state particles. The J/ψ or $\psi(2S)$ meson, further collectively named ψ, is reconstructed in its decay to two muons and the Λ baryon is reconstructed from a proton and a pion. Through this thesis a particle implies always its charged conjugated particle.

First, the used data sets and different types of reconstruction are discussed in Sec. 5.1, followed by the signal selection (Sec. 5.2) and the evaluation of the selection efficiencies (Sec. 5.3). From the selected data sets, the measured signal yields are determined, which is discussed in Sec. 5.4. With the yields and the selection efficiencies, the result is determined in Sec. 5.5. Finally, the systematic uncertainties are evaluated in Sec. 5.6 and the result is discussed in Sec. 5.7.

5.1 Data Sets

The analysed data is the Run I data set of LHCb, which corresponds to an integrated luminosity of 3 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 7$ and 8 TeV. Since the difference between the two centre-of-mass energies does not lead to significant differences in the data, both years of data taking are handled in the very same way. Long and downstream reconstructed Λ baryons are used. The different track types are described in detail in Sec. 3.2. The additional or missing hits between long and downstream tracks have significant effects on the momentum resolution and different background occur. Therefore, long and downstream track reconstructed Λ baryons are treated differently. The combinations of both decay channels and both track types lead to four distinct data sets.

Beside these recorded data sets, signal and background simulation are used in this analysis.

5.2 Signal Selection

The following selection steps are based on Ref. [71]. The data sets have to fulfil multiple steps of selection requirements, starting with an online trigger selection at data taking, followed by a general collaboration-wide preselection and an offline selection.

As the trigger system must select an event on a very short time scale only information of certain sub-detectors can be used. Both decay channels studied contain muons in the final state, which are easy and fast to identify in the muon chambers. Therefore, only trigger requirements on the muons are used. In the three staged trigger, requirements are set with increasing complexity, which ensure that a muon pair with high momentum is in the event. This muons are required to come from a common vertex, which is not the PV. Both decay channels have to fulfil the very same selection. Therefore, no dedicated J/ψ or $\psi(2S)$ trigger lines are used but general high mass dimuon trigger. All used trigger lines and their requirements are listed in Table A.1.

Two often used quantities are the impact parameter and the direction angle. The impact parameter, IP, describes the shortest distance of a track to a point, which is the PV. Often requirements on the significant of the IP are used to ensure that a particle is not coming from the PV. The direction angle, DIRA, is defined as the cosine of the angle between the reconstructed momentum and the origin and decay vertex. For a perfectly reconstructed particle this angle is zero and DIRA equal to one.

In the following step reconstructed tracks are assigned to the final-state particles. Then the intermediate particles are built by combining these final state particles. This bottom-up approach is done until the head of the decay chain is reached. This means the ψ is reconstructed in its decay to two muons, which must be responsible for the positive trigger response. These muons must be well reconstructed and must be unlikely prompt muons from the PV, which are most of the muons in the detector. In particular, these muons are required to not point to any PV (min IP χ^2 to any PV > 9) and must have a well reconstructed track (track $\chi^2/\text{ndf} < 5$). Furthermore, the muons must be likely be a muon ($\text{DLL}_{\mu\pi} > -3$) and unlikely a ghost, which means that random detector noise fakes a particle track (track ghost probability < 0.4). The two muons must originate from a common vertex (vertex $\chi^2/\text{ndf} < 12$) and the combined ψ is required to be displaced from the PV (χ^2 separation to PV > 9) to suppress ψ, which are not originating from a b hadron. Also slight sanity cuts are applied to veto reconstructed ψ, that do not point in the direction of the detector (DIRA > -0.9).

The Λ baryon candidate is reconstructed by combining a proton and a pion. As aforementioned, due to its large lifetime, the Λ can be reconstructed from long or downstream track protons and pions. Since the particle density is higher near to the interaction point, long track protons and pions must fulfil stronger requirements. The proton and pion candidates must have a large p ($p > 2\,\text{GeV}/c$) and p_T (p_T (p)>650(550) MeV/c and p_T (π)>130(250) MeV/c for downstream (long track) candidates). Both final state hadrons must be incompatible with coming from the PV (minIP χ^2 (PV) > 4 or 9 for downstream and long track candidates). Additionally, long track protons have be identified as protons ($\text{DLL}_{p\pi} > -5$). The combined mass of the two hadrons has to be in a small range around the known Λ mass ($m_{p\pi}$ in $[1105, 1125]\,\text{MeV}/c^2$ or $[1110, 1125]\,\text{MeV}/c^2$ for downstream and long track candidates) and they have to originate from a common, well reconstructed vertex (vertex $\chi^2/\text{ndf} < 23.25$ or < 30 for downstream and long track candidates). Further, the reconstructed Λ baryon is required to have a large decay time ($t > 2$ and 9 ps for long track and downstream candidates), to separate real Λ baryons from random combinations of proton and pions. This requirement further suppresses non-signal intermediate resonances (*e.g.* K^*), which have much lower lifetimes.

The Λ_b^0 candidate is then reconstructed from the ψ and the Λ candidate. The reconstructed Λ_b^0 mass has to be in a broad range around the known Λ_b^0 mass [72] ($4900 < m_{\pi p\mu\mu} < 7000\,\text{MeV}/c^2$). Furthermore, the decay vertex of the Λ_b^0 baryon must be well reconstructed (vertex $\chi^2 < 23.75(16.5)$ for downstream (long track) candidates) and the Λ_b^0 must compatible with coming from the assigned PV (IP to PV $\chi^2 < 16$ and DIRA > 0.9999). A further requirement is that the Λ_b^0 baryon decay vertex is significantly displaced from the PV (χ^2 distance from related PV > 121). Finally, a requirement for the whole decay chain is placed by requiring that a decay tree fit [73], where the Λ and ψ masses are constrained to their known values [72], describes the data well ($\text{DTF}_{\psi\Lambda} < 26(36)$ for downstream (long track) candidates). On top of the selection fiducial cuts on the transverse momentum of the Λ_b^0 ($< 20000\,\text{MeV}/c$) and the pseudo rapidity η (between 2 and 4.5) are placed to be able to extract the selection efficiency reliably from simulation, which is further explained in Sec. 5.3.4. After the selection multiple candidates are present in about 1% of the events. As two real signal decays are very unlikely and often these multiple candidates share tracks or particles. One of these candidates is randomly retained to avoid any bias.

The J/ψ and $\psi(2S)$ decays are disentangled by dividing the data in the reconstructed ψ mass ($m_{\mu^+\mu^-} \lessgtr 3400$). Since this cut is far away from both ψ mass peaks, no signal is lost and no cross feeds between both data sets persists. In Fig. 5.1 the relevant reconstructed masses for $\Lambda_b^0 \to J/\psi\,\Lambda$ decays are shown for both track types separately.

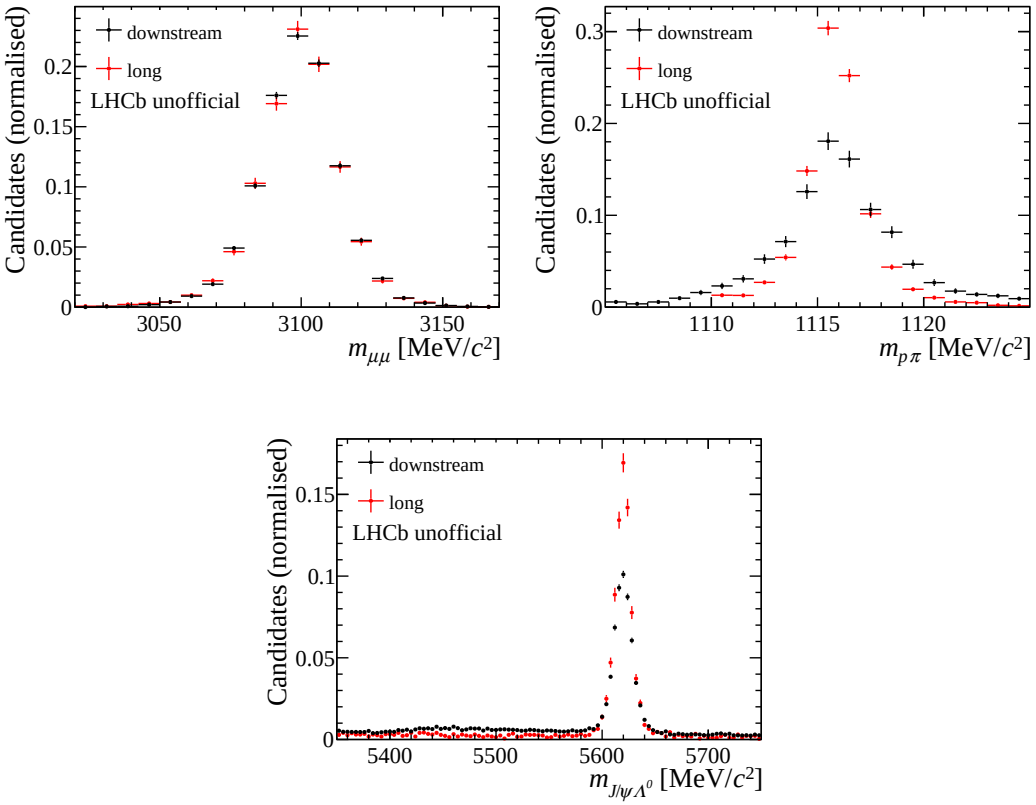

Figure 5.1: The normalised distributions of the reconstructed (upper-left) J/ψ, (upper-right) Λ and (lower) Λ_b^0 mass for downstream reconstructed candidates (black) and for long track reconstructed candidates (red) after all selections applied.

The daughter (ψ and Λ) masses show a Gaussian-like peak with nearly no contribution from combinatorial background, which is random combinations of muons and pions and protons, respectively. As expected, the mass resolution of long track reconstructed Λ baryons is better compared to downstream reconstructed candidates. In the reconstructed Λ_b^0 mass one can see a clear signal peak at the nominal Λ_b^0 mass with a small contribution from combinatorial background as flat distribution. Furthermore, a peaking structure can be seen in the downstream track reconstructed sample at lower masses, the origin of this structure is discussed in Sec. 5.4.1.

5.3 Signal Efficiency Determination

A crucial part of a branching fraction measurement is the correct determination of all efficiencies involved in the measurement of the signal candidates as the absolute efficiency enters directly in the calculation (see Eq. (2.5)). Inefficiencies occur at several points ranging from the detector acceptance, the track reconstruction to the selection requirements. All efficiencies are evaluated using simulation and cross-checked with data, where possible. It is known that the simulation does not describe the data perfectly. Especially, in the b hadron production kinematic differences are expected. Therefore, efficiencies are divided into ones that depend on the production kinematic of the Λ_b^0 baryon ($p_{\mathrm{T}}(\Lambda_b^0)$ and $\eta(\Lambda_b^0)$), and ones that are independent. The geometrical efficiency ϵ_{geo}, the efficiency for hadronic interaction $\epsilon_{\Lambda\ \mathrm{int}}$ and fiducial cuts $\epsilon_{\mathrm{fiducial}}$ are considered as independent. In contrast, the preselection, reconstruction, offline and trigger efficiencies $\epsilon_{\mathrm{rec,\ off,\ trig}}$ are considered as dependent. The total efficiency, $\epsilon_{\mathrm{total}}$, is then given as

$$\epsilon_{\mathrm{total}}(p_{\mathrm{T}},\eta) = \epsilon_{\mathrm{geo}} \cdot \epsilon_{\Lambda\ \mathrm{int}} \cdot \epsilon_{\mathrm{fiducial}} \cdot \epsilon_{\mathrm{rec}}(p_{\mathrm{T}},\eta) \cdot \epsilon_{\mathrm{off}}(p_{\mathrm{T}},\eta) \cdot \epsilon_{\mathrm{trig}}(p_{\mathrm{T}},\eta). \tag{5.2}$$

All these steps are discussed in the following.

5.3.1 Simulation Corrections

The studied decays are the decay of a spin 1/2 particle (Λ_b^0) to a spin 1 (ψ) and a spin 1/2 particle (Λ). This introduces a helicity structure in decay. The available simulation does not include this helicity structure and is flat in phase space. To describe this structure, one can define three characteristic angles: The angle θ_1 is defined as azimuthal angle of the proton in the Λ rest-frame, θ_2 is defined as azimuthal angle of the positive muon in the ψ rest-frame, while θ is defined as polar angle of the Λ momentum in the Λ_b^0 rest-frame with respect to a unit vector perpendicular to the production plane. This is illustrated in Fig. 5.2.

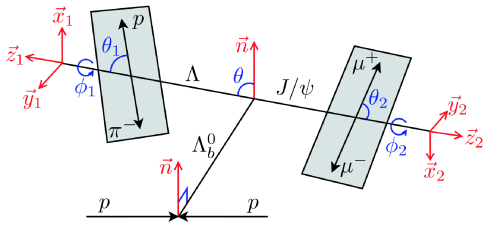

Figure 5.2: The different angles involved in the decays. Taken from [74].

It is convenient to use the cosine of these angles as it is a flat distribution for phase space decays. The angle θ gives a measure of the production polarisation of the Λ_b^0. This angle is measured compatible with being unpolarised (flat in $\cos(\theta)$) [74]. Therefore, no correction for is needed in the simulation. For the other two helicity angles corrections are needed. In Fig. 5.3 the used simulation on generator level, which does not include effects from the detection and reconstruction at this stage, and the expected functions from theory predictions [5] are shown.

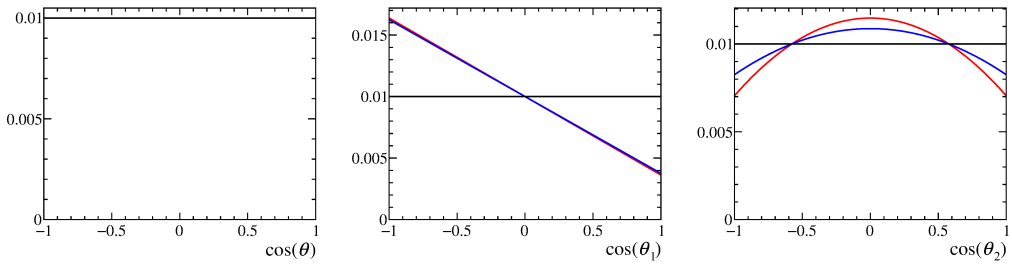

Figure 5.3: Theoretical distributions for $\cos(\theta)$, $\cos(\theta_1)$, $\cos(\theta_2)$ with the values from [5] for $\Lambda_b^0 \to J/\psi\Lambda$ (red) and $\Lambda_b^0 \to \psi(2S)\Lambda$ (blue).

In particular, $\cos(\theta_1)$ shows a large difference between prediction and phase space simulation, since the parity between pion and proton is largely violated [16], which is not included in the simulation. This leads to a underestimated momentum of the pion in the uncorrected simulation, which has large effects on the reconstruction and selection efficiencies. The difference in $\cos(\theta_2)$ is clearly visible but has a smaller effect on the estimated efficiencies. These differences are corrected by reweighting the flat phase space simulation to the predicted distribution from theory. Since the correlation between $\cos(\theta_1)$ and $\cos(\theta_2)$ is small, the reweighting is done independently for both observables and then the final weights are the product of the two individual weights. The reweighted distributions with the theory predictions as well as seen distributions on signal weighted data from long track $\Lambda_b^0 \to J/\psi\Lambda$

decays are shown in Fig. 5.4, while the distributions of the other data sets are given in Fig. A.1.

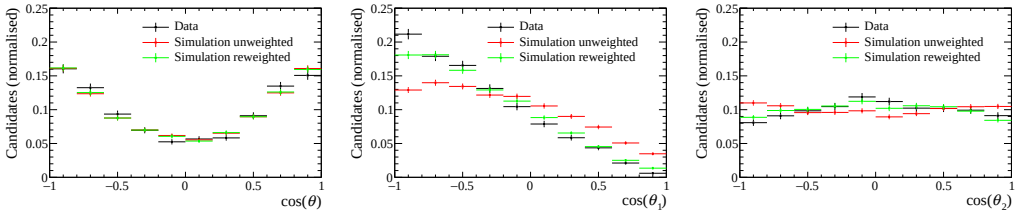

Figure 5.4: Distributions for (left) $\cos(\theta)$, (middle) $\cos(\theta_1)$, (right) $\cos(\theta_2)$ for sweighted Data (blue), unweighted MC (red) and reweighted MC (green) for long track $\Lambda_b^0 \to J/\psi\,\Lambda$ decays.

5.3.2 Geometrical Efficiency of the Detector

Since the LHCb detector is not covering all solid angles, reconstruction inefficiencies occur when particles are produced in all directions. To save time in the simulation, events where the signal particles are obviously outside of the LHCb acceptance are rejected in a very early stage of the generation. For the signal simulation this results in an efficiency of

$$\begin{aligned} \epsilon_{\text{geo}}^{J/\psi} &= (19.17 \pm 0.04)\%, \\ \epsilon_{\text{geo}}^{\psi(2S)} &= (18.80 \pm 0.03)\%, \end{aligned}$$

which is very similar between both decay channels.

5.3.3 Hadronic Detector Interaction Efficiency

Due to the long lifetime of the Λ baryon, it is possible that the Λ interacts with detector material. In such an interaction the Λ does not decay to a proton and a pion and is not reconstructed. This interaction is simulated with Geant4 [65, 64] in the detector simulation. This leads to a difference between the number of generated events and events in the generated data set. The efficiencies are determined to be

$$\begin{aligned} \epsilon_{\Lambda\ \text{int}}^{J/\psi} &= (89.11 \pm 0.03)\%, \\ \epsilon_{\Lambda\ \text{int}}^{\psi(2S)} &= (92.09 \pm 0.02)\%. \end{aligned}$$

The slight difference between both decay channels can be explained by the larger momentum of the Λ baryon in $\Lambda_b^0 \to J/\psi\Lambda$ decays compared to $\Lambda_b^0 \to \psi(2S)\Lambda$ decays, which leads to a larger flight distance on average. Thus, the Λ baryon travels further through the detector material and has a higher probability to interact.

5.3.4 Momentum and Pseudo-rapidity Dependent Efficiencies

It is known that the available simulation does not describe the measured data perfectly. This leads especially to different results for the two track types as the pseudo-rapidity distribution is different for long track compared to downstream reconstructed candidates. The fact that the Λ decay vertex has to be inside the VELO places constraints on the possible flight direction leading to higher pseudo-rapidity values for long track reconstructed candidates. Thus, the mismodelling has a non-trivial effect on the efficiency determination. To account for this effect, efficiencies are determined dependent on the Λ_b^0 phase space (p_T, η) and the measured Λ_b^0 phase space on data is used. The dependence is characterized by evaluating the efficiency of the reconstruction, trigger and selection steps in bins of $p_\mathrm{T}(\Lambda_b^0)$ and $\eta(\Lambda_b^0)$. To ensure that each of these bins contains a sufficient amount of simulated events, fiducial cuts are introduced, $p_\mathrm{T}(\Lambda_b^0) < 20\,\mathrm{GeV}/c$ and $2 < \eta(\Lambda_b^0) < 4.5$. The efficiency of these fiducial cuts is evaluated using simulation and determined to be

$$\epsilon_\mathrm{fiducial}^{J/\psi} = (78.30 \pm 0.03)\%,$$
$$\epsilon_\mathrm{fiducial}^{\psi(2S)} = (78.17 \pm 0.03)\%.$$

The binning scheme is the same for both decay channels because the distributions are similar and binning effects cancel in first order. However, different schemes are used for the two track types. To ensure low fluctuations due to the number of simulated events in a certain bin, the binning is constructed in a way that each bin contains the same number of simulated $\Lambda_b^0 \to J/\psi\Lambda$ events. This sample is used because it is the smallest of the available data sets. The total number of bins is set to 7 bins for p_T (Λ_b^0) and $\eta(\Lambda_b^0)$ each resulting in 49 bins. This ensures that each bin contains more than 100 simulated events after reconstruction for each track type. The construction of the bins starts with binning $\eta(\Lambda_b^0)$. This variable is more flat leading to bins containing more similar areas. Then the data is binned in $p_\mathrm{T}(\Lambda_b^0)$. To obtain readable bin boundaries, the resulting bin ranges are rounded to three significant numbers in each dimension. The resulting two dimensional binned efficiencies for each data set are shown in Fig. 5.5.

The efficiency differences between two neighbouring bins can be large especially for large p_T and large η, where the bin size increases. Given that the efficiency is

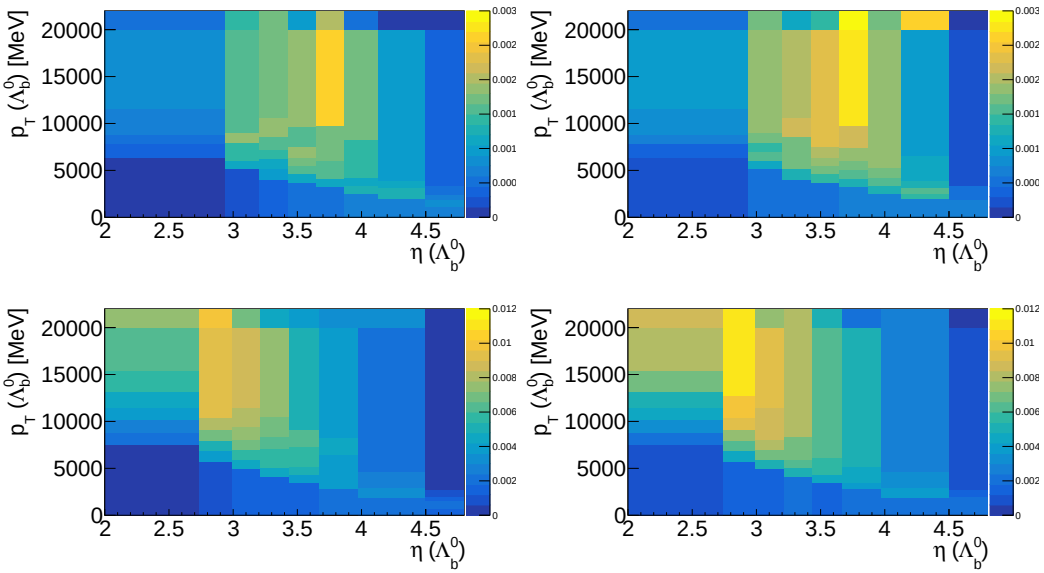

Figure 5.5: Total efficiency for each p_{T} and η bin for (top) long track and (down) downstream (left) $\Lambda_b^0 \to J/\psi\,\Lambda$ and (right) $\Lambda_b^0 \to \psi(2S)\Lambda$ candidates evaluated on simulation.

expected to be a smooth function in these two dimensions, the binned efficiencies are interpolated. This reduces the dependence of the efficiencies arising from the specific binning scheme used. As the bin size varies largely, a simple and robust linear interpolation procedure between the bins is applied. To ensure a good interpolation at the limits of the two dimensions, additional bins are added for $20\,\mathrm{GeV}/c < p_{\mathrm{T}}(\Lambda_b^0) < 30\,\mathrm{GeV}/c$ and $4.5 < \eta(\Lambda_b^0) < 8$ and the efficiency is set to zero for $p_{\mathrm{T}}(\Lambda_b^0) = 0$ and for $\eta(\Lambda_b^0) = 2$. This reflects that $\eta(\Lambda_b^0) = 2$ is outside of the detector acceptance and $p_{\mathrm{T}}(\Lambda_b^0) = 0$ means that the Λ_b^0 baryon flies along the beam pipe and is unlikely to be detected. The points for the interpolation are defined as the mean of the p_{T} of all simulated candidates in one bin, while for the η dimension the centre of the bin is used. Both dimensions are treated differently because the distribution in $\eta(\Lambda_b^0)$ is much more flat and this simplifies the method without impacting the resulting interpolation. Starting with the p_{T} dimension the point between two bin centres are interpolated, in a second step the η dimension is interpolated between these interpolated points. The interpolation procedure is depicted in Fig. 5.6, while the resulting interpolated efficiency maps are given in Fig. 5.7. Some fluctuations can be seen at the limits of the range, due to fluctuations in the additional bins, which do not have necessary enough simulation inside. Nevertheless, these fluctuations have no effect on the recorded data, as it is

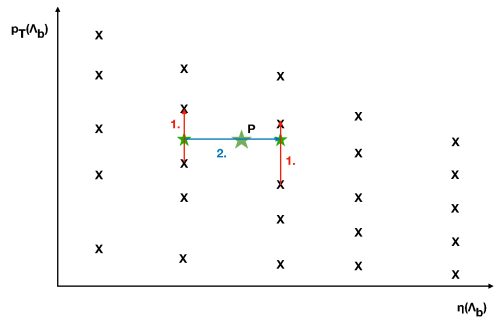

Figure 5.6: Scheme of the linear interpolation, crosses depict the mean values of the bin, with the mean efficiency. The first step is the interpolation in $p_{\mathrm{T}}(\Lambda_b^0)$ (red), followed by the interpolation in $\eta(\Lambda_b^0)$ to the point P, which is the point in phase space, where the efficiency needs to be determined.

very unlikely to produce Λ_b^0 with high p_{T} and high pseudo-rapidity, which means very high total momentum.

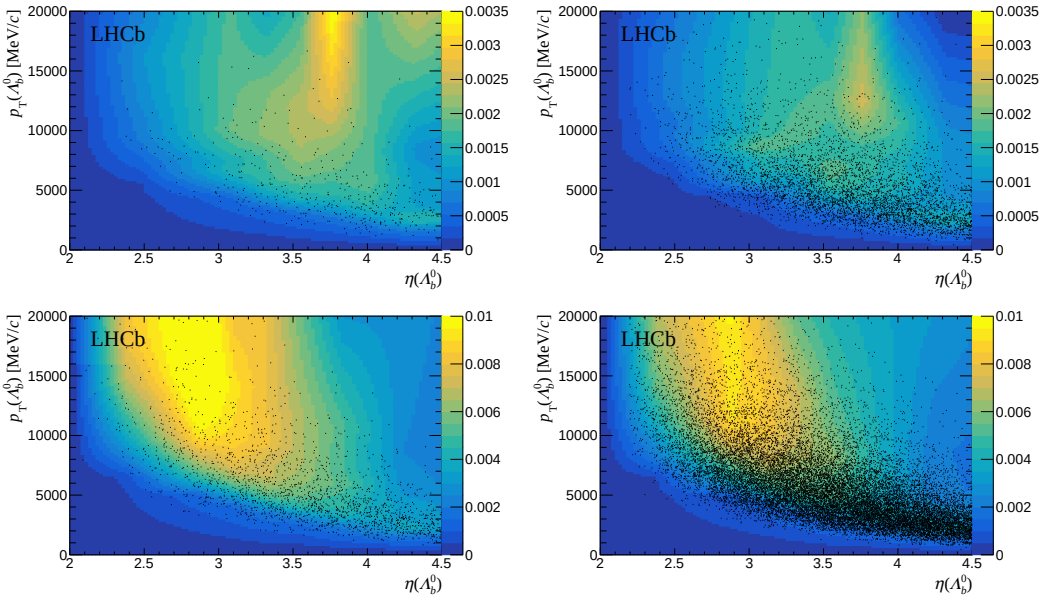

Figure 5.7: Interpolated efficiency function for long-track candidates for (upper-left) $\Lambda_b^0 \to \psi(2S)\Lambda$ and (upper-right) $\Lambda_b^0 \to J/\psi\,\Lambda$ and for downstream-track candidates for (lower-left) $\Lambda_b^0 \to \psi(2S)\Lambda$ and (lower-right) $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates. The distribution of the candidates on data is shown with black dots.

5.4 Signal Yield Extraction

To measure the ratio of branching fractions, the signal yields for both decay channels are needed. These yields are determined by modelling every contribution to the measured Λ_b^0 mass spectrum using an extended unbinned maximum likelihood fit (Sec. 4.1). In Sec. 5.4.1 the remaining background component are discussed, followed by the description of the fit model in Sec. 5.4.2. The fit and its results are presented in Sec. 5.4.2.

5.4.1 Resonant Backgrounds

Looking at the reconstructed Λ_b^0 mass distribution (Fig. 5.1), an additional structure can be seen besides the signal and the combinatorial background slightly below the signal peak. This is especially prominent in the downstream track sample. The structure consists of two remaining kinds of resonant backgrounds. The first is due to misidentification of a pion from a K_S^0 decay as a proton. This leads to identifying $B^0 \to J/\psi K_S^0$ ($B^0 \to \psi(2S) K_S^0$) as the signal $\Lambda_b^0 \to J/\psi \Lambda$ ($\Lambda_b^0 \to \psi(2S)\Lambda$). Since the B^0 meson is lighter than the Λ_b^0 baryon and the mass of the K_S^0 is reconstructed too high by applying the proton hypothesis to the pion, this contribution is shifted into the signal range. The reason that this background is dominant in the downstream track sample is due to the applied PID requirements for the proton in the long track sample (compare Sec. 5.2) that reduces this background largely. For the downstream track sample this requirement is not applied, because the $B^0 \to \psi K_S^0$ can be easily described in the fit. An alternative approach vetoing this background is checked and results in comparable values for the branching fraction.

The second kind of remaining background is partially reconstructed background from $\Xi_b^- \to J/\psi \Xi^-$ ($\Xi_b^- \to \psi(2S)\Xi^-$) decays. The dominant Ξ baryon decay is into a Λ baryon and a pion. As the additional pion is not reconstructed, the detector measures the signal final state. Since the mass difference between the Ξ baryon and the Λ baryon is similar than the pion mass, the missed pion has not much momentum. This leads to a distribution that starts at the mass difference of the Ξ_b mass and the pion mass, which is roughly the Λ_b^0 mass and has a large tail to lower masses. Due to the small kinematic differences and the same measured final state, this background is irreducible. The reason there is no sign of its contribution in the long track sample is due to the Ξ lifetime, which is 290 ± 9 ps and in the same order of magnitude as the lifetime of the Λ. Since the pion does not carry much momentum, the Ξ baryon and the resulting Λ baryon fly in the same direction. Thus, a Λ baryon is reconstructed with roughly two times the lifetime of its normal lifetime. Therefore, it is unlikely that the Λ baryon decays inside the VELO. Simulation of $\Xi_b^- \to J/\psi \Xi^-$ decays

shows that less than 5% of these decays are long-track reconstructible and therefore it is negligible in the long track sample. The contribution of this background is modelled in the fit.

5.4.2 Fit Model

To extract the signal yield, each component contributing to the reconstructed Λ_b^0 mass needs to be described. The used reconstructed Λ_b^0 mass is computed using a decay tree fit [73], where the J/ψ (or $\psi(2S)$) and the Λ candidates are constrained to their nominal masses.

The fit model consists of four components, a double-sided Hypatia function [75] for the signal, two independent empiric Gaussian kernel functions (GK) [76] for the two resonant backgrounds and an exponential function for the combinatorial background. Thus, the full model is given as

$$P = N_{\Lambda_b^0\to\psi\Lambda}\mathcal{J}(\mu,\sigma,\lambda) + N_{B^0\to\psi K_S^0}GK_{B^0} + N_{\Xi_b^-\to\psi\Xi^-}GK_{\Xi_b} + N_{\text{comb}}\exp(\alpha). \tag{5.3}$$

The Hypathia function is a generalised Crystal Ball function [77], which can model the Gaussian mass resolution and radiative tails from bremsstrahlung effects. The Hypathia function is defined as

$$\mathcal{J}(m;\mu,\sigma,\lambda,\zeta,\beta,a_1,n_1,a_2,n_2) \propto$$
$$\begin{cases} ((m-\mu)^2+A_\lambda^2(\zeta)\sigma^2)^{\frac{1}{2}\lambda-\frac{1}{4}}e^{\beta(m-\mu)}K_{\lambda-\frac{1}{2}}\left(\zeta\sqrt{1+\left(\frac{m-\mu}{A_\lambda(\zeta)\sigma}\right)}\right) & , -a_1 < \frac{m-\mu}{\sigma} < a_2 \\ \frac{G(\mu-a_1\sigma,\mu,\sigma,\lambda,\zeta,\beta)}{\left(1-m/(n\frac{G(\mu-a_1\sigma,\mu,\sigma,\lambda,\zeta,\beta)}{G'(\mu-a_1\sigma,\mu,\sigma,\lambda,\zeta,\beta)}-a_1\sigma)\right)^{n_1}} & , -a_1 > \frac{m-\mu}{\sigma} \\ \frac{G(\mu-a_2\sigma,\mu,\sigma,\lambda,\zeta,\beta)}{\left(1-m/(n\frac{G(\mu-a_2\sigma,\mu,\sigma,\lambda,\zeta,\beta)}{G'(\mu-a_2\sigma,\mu,\sigma,\lambda,\zeta,\beta)}-a_2\sigma)\right)^{n_2}} & , \quad a_2 < \frac{m-\mu}{\sigma} \end{cases} \tag{5.4}$$

with

$$G(m;\mu,\sigma\lambda,\zeta,\beta) = ((m-\mu)^2 + A_\lambda^2(\zeta)\sigma^2)^{\frac{1}{2}\lambda-\frac{1}{4}} e^{\beta(m-\mu)} K_{\lambda-\frac{1}{2}}\left(\zeta\sqrt{1+\left(\frac{m-\mu}{A_\lambda(\zeta)\sigma}\right)}\right),$$

and $A_\lambda^2(\zeta) = \dfrac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}$

G: generalised hyperbolic distribution

G': derivative of G

K_λ: cylindrical harmonics.

The parameters ζ and β, that model an asymmetry in the peak are set to zero resulting in a more robust fit. Further, the parameter a_i, which defines at which point the Gaussian-like core transits into the tail, is highly correlated with the parameter n_i, which defines the slope of the tail. Therefore, n_i is set to a fixed value, which improves the robustness of the fit model further. The remaining tail parameters (a_1 and a_2) are fixed to values obtained from fits to simulation. These fits and the resulting fit parameters are shown in Fig. 5.8 and Table 5.1.

Table 5.1: The parameters a_1, n_1, a_2, n_2, β, λ, ζ of the Hypatia function from fits to simulation for downstream and long-track candidates.

	$\Lambda_b^0 \to J/\psi\Lambda$		$\Lambda_b^0 \to \psi(2S)\Lambda$	
Parameter	downstream	long	downstream	long
n_1(fixed)	2	2	2	2
n_2(fixed)	2	2	2	2
a_1	2.90 ± 0.10	3.32 ± 0.33	2.88 ± 0.06	3.43 ± 0.29
a_2	3.08 ± 0.13	4.60 ± 0.84	3.09 ± 0.09	3.78 ± 0.40
β(fixed)	0	0	0	0
λ	-3.05 ± 0.19	-2.02 ± 0.13	-3.12 ± 0.15	-1.97 ± 0.11
ζ(fixed)	0	0	0	0

The GK functions, which describe the partially reconstructed background from Ξ_b decays are extracted from $\Xi_b^- \to J/\psi\Xi^-$ ($\Xi_b^- \to \psi(2S)\Xi^-$) simulation that is reconstructed as $\Lambda_b^0 \to J/\psi\Lambda$ ($\Lambda_b^0 \to \psi(2S)\Lambda$), which is shown in Fig. 5.9. GK functions are also used to describe misidentified $B^0 \to \psi K_\mathrm{S}^0$ decays using the corresponding simulation. The extracted models are given in Fig. 5.10. These empirical models are used in the mass fit on data with their yields left floating. This is the only practical method as the branching fractions of these decays are not well known especially in the $\psi(2S)$ case. The possible effect on the measurement is

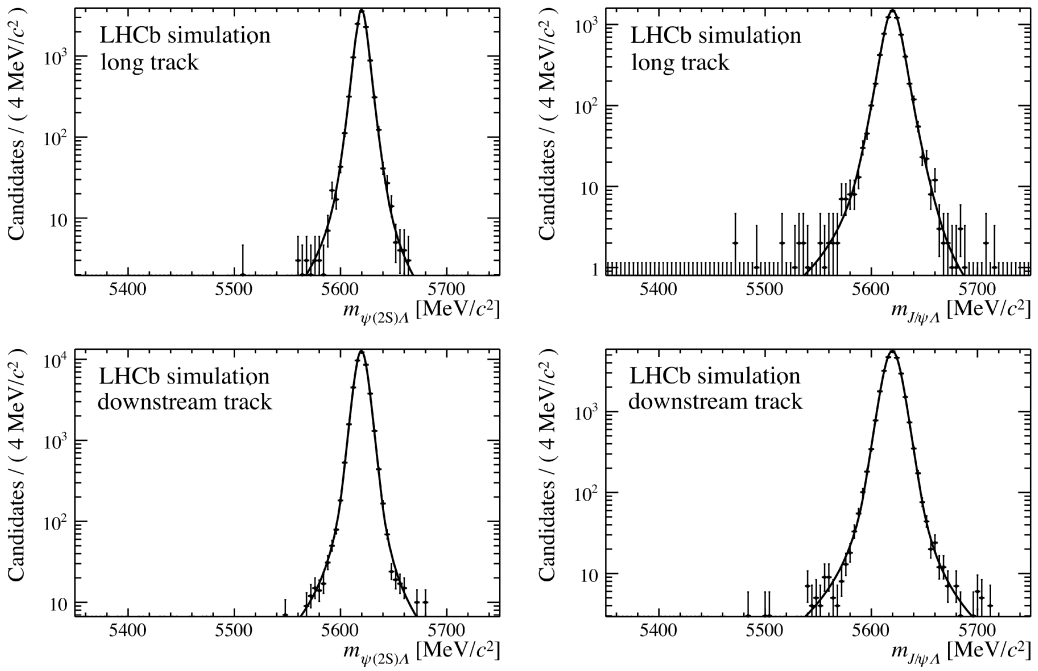

Figure 5.8: Fits to the simulated invariant-mass signal distributions of long-track candidates for $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$ and for downstream-track candidates for $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$ candidates using a double-sided Hypatia function.

minimal since the yield of these backgrounds is strongly correlated with the one of the combinatorial background and rather uncorrelated to the signal yield.

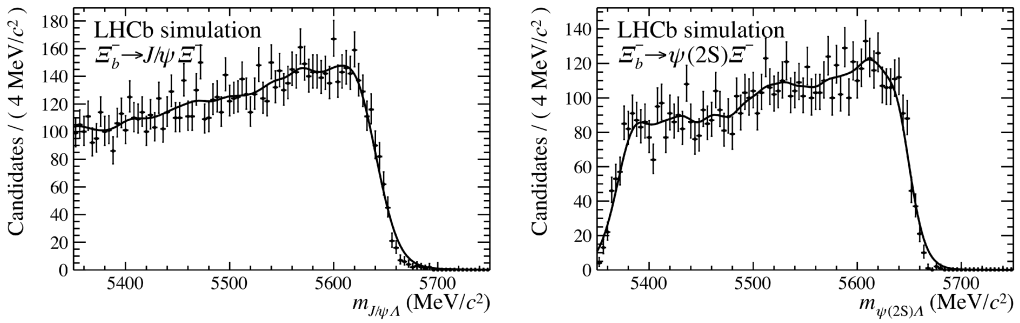

Figure 5.9: The reconstructed invariant mass of (left) $\Xi_b^- \to J/\psi\,\Xi^-$ and (right) $\Xi_b^- \to \psi(2S)\Xi^-$ simulation, which is reconstructed and fully selected as $\Lambda_b^0 \to J/\psi\,\Lambda$ and $\Lambda_b^0 \to \psi(2S)\Lambda$, respectively, with a fitted empiric Gaussian kernel function.

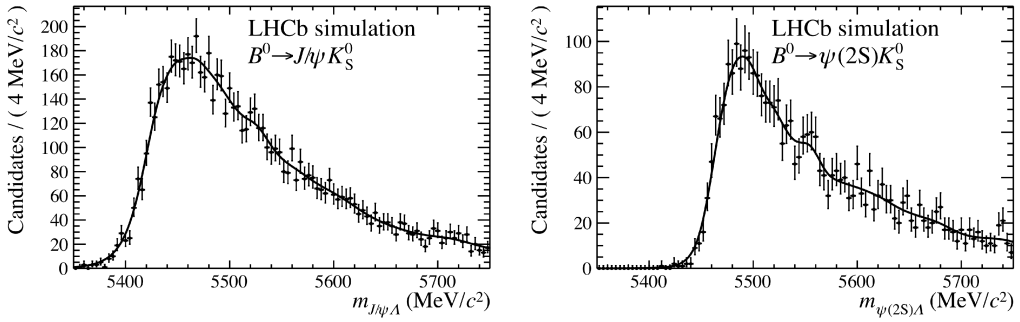

Figure 5.10: The reconstructed invariant mass of (left) $B^0 \to J/\psi\,K^0_S$ and (right) $B^0 \to \psi(2S)K^0_S$ simulation, which is reconstructed and fully selected as $\Lambda_b^0 \to J/\psi\,\Lambda$ and $\Lambda_b^0 \to \psi(2S)\Lambda$, respectively, with a fitted empiric Gaussian kernel function.

To extract the number of candidates, the complete model (Eq. (5.3)) is fitted to the selected data for each decay channel and track type. These fits describe the data very well and result in 14863 ± 138 $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates and 1131 ± 38 $\Lambda_b^0 \to \psi(2S)\Lambda$ candidates. The fits are shown in Fig. 5.11, while the yields of all components and all other parameters are given in Tables 5.2 and 5.3.

For the measurement of the ratio of branching fractions (Eq. (2.5)) the efficiency-corrected signal yields are needed. The method to extract these efficiency-corrected yields is to repeat the mass fits but weighting the candidates with a per-candidates weight, that is defined as the inverse of its efficiency. The per-candidate weight

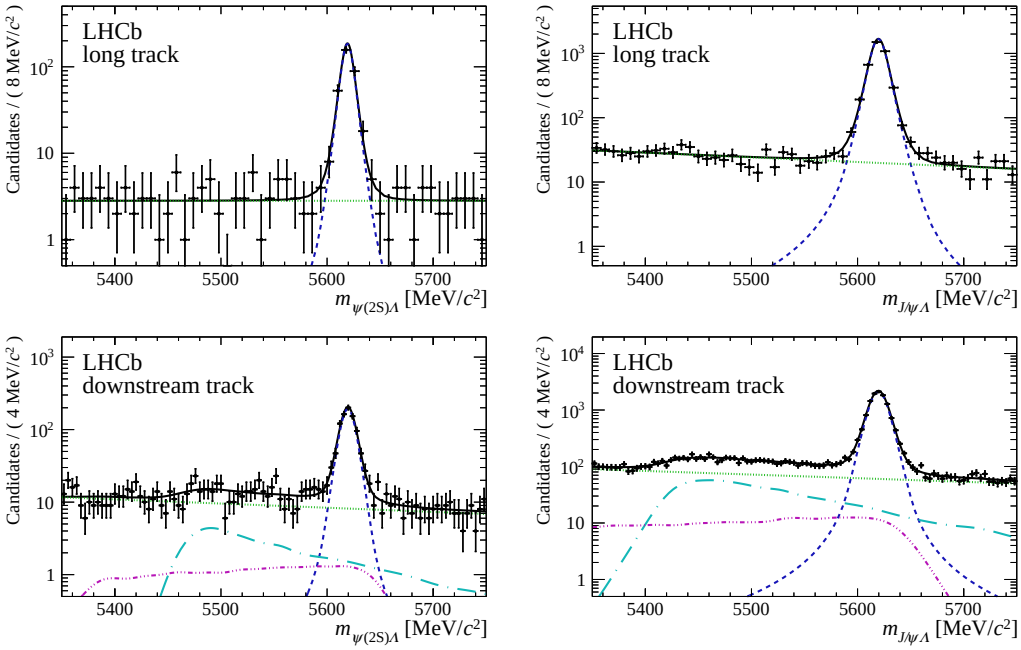

Figure 5.11: Fits to the (unweighted) invariant-mass distributions of long-track candidates for (upper-left) $\Lambda_b^0 \to \psi(2S)\Lambda$ and (upper-right) $\Lambda_b^0 \to J/\psi\Lambda$ and for downstream-track candidates for (lower-left) $\Lambda_b^0 \to \psi(2S)\Lambda$ and (lower-right) $\Lambda_b^0 \to J/\psi\Lambda$ candidates. The signal (blue, dashed), the combinatorial background (green, dotted), the $B^0 \to \psi K_S^0$ background (cyan, long-dash-dotted) and the $\Xi_b^- \to \psi\Xi^-$ background (violet, dash-triple-dotted) are indicated.

Table 5.2: Yields from the invariant-mass fits in the range 5350 to 5750 MeV/c^2 of (top) $\Lambda_b^0 \to J/\psi\,\Lambda$ decays and (bottom) $\Lambda_b^0 \to \psi(2S)\Lambda$ decays for each component.

track type	$\Lambda_b^0 \to J/\psi\,\Lambda$	$B^0 \to J/\psi\,K_S^0$	$\Xi_b^- \to J/\psi\,\Xi^-$	combinatorial
downstream	$11\,090 \pm 120$	$2\,330 \pm 210$	800 ± 400	$6\,790 \pm 240$
long	$3\,800 \pm 60$	–	–	$1\,130 \pm 40$
	$\Lambda_b^0 \to \psi(2S)\Lambda$	$B^0 \to \psi(2S)K_S^0$	$\Xi_b^- \to \psi(2S)\Xi^-$	combinatorial
downstream	819 ± 33	160 ± 60	60 ± 90	920 ± 60
long	317 ± 19	–	–	140 ± 13

Table 5.3: The results of the parameters in the fit with the fixed parameters from Table 5.1 separately for downstream and long track candidates for the signal decays.

	$\Lambda_b^0 \to J/\psi\,\Lambda$		$\Lambda_b^0 \to \psi(2S)\Lambda$	
Parameter	downstream	long	downstream	long
$\alpha\,[(\,\mathrm{MeV}/c^2)^{-1} \cdot 10^{-3}]$	-1.45 ± 0.17	-1.67 ± 0.27	-1.37 ± 0.40	$(0.00 \pm 7.67) \cdot 10^{-4}$
$\mu\,[\,\mathrm{MeV}/c^2]$	5619.60 ± 0.10	5619.63 ± 0.13	5619.62 ± 0.29	5619.22 ± 0.35
$\sigma\,[\,\mathrm{MeV}/c^2]$	9.83 ± 0.11	9.61 ± 0.17	7.84 ± 0.33	7.31 ± 0.44

is determined from the two dimensional efficiency functions described in Sec. 5.3. This method is only possible to use because the efficiency is uncorrelated to the reconstructed Λ_b^0 mass and thus the weighted distribution can be described with the same model as the unweighted distribution. Since weights are used in the fit, the uncertainty of the fit parameters is determined using a so-called sum weight squared correction. In this correction, the unweighted covariant matrix, V, replaced by the corrected covariance matrix, V_{corr}, which is calculated using a covariance matrix evaluated with squared weights, C, as

$$V_{\mathrm{corr}} = VC^{-1}V. \tag{5.5}$$

From this corrected covariance matrix, parameter uncertainties are estimated. Their coverage is tested using pseudo-experiments. The resulting fits are shown in Fig. 5.12, while the efficiency-corrected yields are listed in Table 5.4 and the other fit parameters are given in Table 5.5.

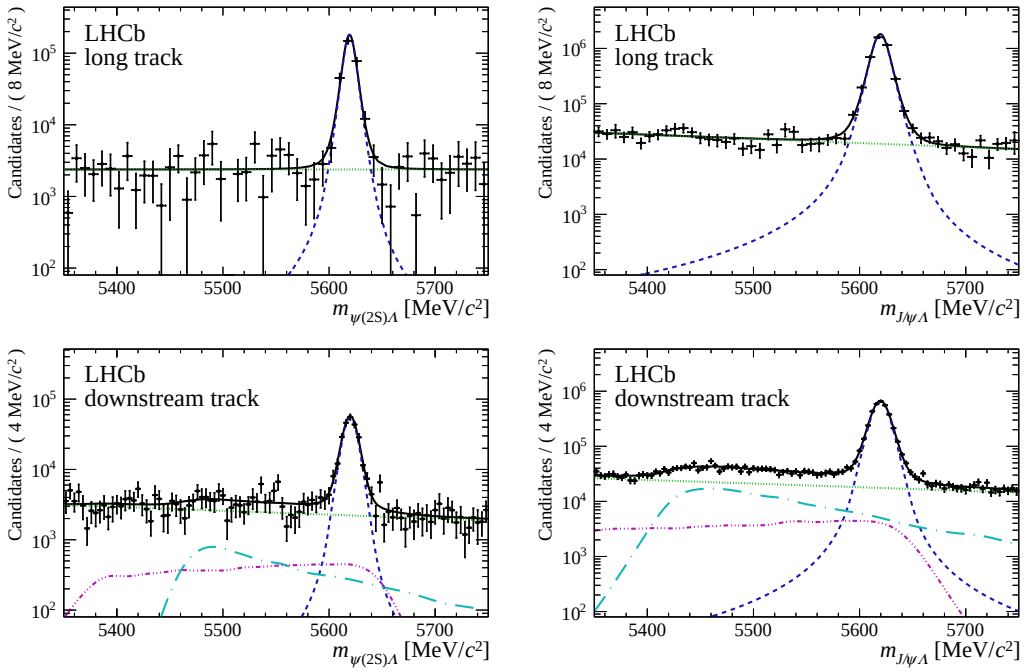

Figure 5.12: Fits to the weighted invariant-mass distributions of long-track candidates for (upper-left) $\Lambda_b^0 \to \psi(2S)\Lambda$ and (upper-right) $\Lambda_b^0 \to J/\psi\Lambda$ and for downstream-track candidates for (lower-left) $\Lambda_b^0 \to \psi(2S)\Lambda$ and (lower-right) $\Lambda_b^0 \to J/\psi\Lambda$ candidates. The signal (blue, dashed), the combinatorial background (green, dotted), the $B^0 \to \psi K_S^0$ background (cyan, long-dash-dotted) and the $\Xi_b^- \to \psi\Xi^-$ background (violet, dash-triple-dotted) are indicated.

Table 5.4: Efficiency-corrected yields of $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ signal decays from the fit to the weighted invariant mass for both track types.

track type	$N_{\Lambda_b^0 \to \psi(2S)\Lambda}$	$N_{\Lambda_b^0 \to J/\psi\,\Lambda}$
downstream	$223\,000 \pm 13\,000$	$3\,320\,000 \pm 50\,000$
long	$280\,000 \pm 18\,000$	$3\,980\,000 \pm 80\,000$

Table 5.5: The results of the parameters in the weighted fit with the fixed parameters from Table 5.1 separately for downstream and long track candidates for the signal decays.

	$\Lambda_b^0 \to J/\psi\,\Lambda$		$\Lambda_b^0 \to \psi(2S)\Lambda$	
Parameter	downstream	long	downstream	long
$\alpha\,[(\,\mathrm{MeV}/c^2)^{-1} \cdot 10^{-3}]$	-1.48 ± 0.33	-1.73 ± 0.22	-1.28 ± 0.76	$(0.00 \pm 1.56) \cdot 10^{-4}$
$\mu\,[\,\mathrm{MeV}/c^2]$	5619.57 ± 0.11	5619.52 ± 0.16	5619.62 ± 0.29	5619.22 ± 0.35
$\sigma\,[\,\mathrm{MeV}/c^2]$	9.58 ± 0.14	9.34 ± 0.19	7.51 ± 0.43	6.66 ± 0.43

5.5 Determination of the Ratio of Branching Fractions

Using the efficiency-corrected signal yields, N, and the known branching fractions of the ψ mesons to two muons [72], $\mathcal{B}(J/\psi \to \mu^+\mu^-)$ and $\mathcal{B}(\psi(2S) \to \mu^+\mu^-)$, the ratio of branching fractions is determined seperatly for long- and downstream-track reconstructed candidates as

$$\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)} = \frac{N_{\Lambda_b^0 \to \psi(2S)\Lambda}}{N_{\Lambda_b^0 \to J/\psi\,\Lambda}} \cdot \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi(2S) \to \mu^+\mu^-)}. \tag{5.6}$$

Assuming lepton universality, the value of the branching fraction of $\psi(2S)$ into two electrons, $\mathcal{B}(\psi(2S) \to e^+e^-) = (0.793 \pm 0.017)\%$ [72], is used for the decay into muons as the external measurement is much more precise in the electronic decay. With the external of the branching fraction J/ψ into two muons, $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033)\%$ [72], as well as the determined efficiency-corrected signal yields (given in Table 5.4), the ratio of branching fractions is determined to be

$$\left[\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)}\right]_{\text{long track}} = 0.528 \pm 0.036,$$

$$\left[\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)}\right]_{\text{downstream track}} = 0.504 \pm 0.029,$$

where the statistical uncertainty only includes the uncertainty on the measured signal yields. The two individual measurements are combined using a weighted average, where the weight is the variance of the individual measurements, into

$$\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)} = 0.513 \pm 0.023.$$

5.6 Systematic Uncertainties

Systematic uncertainties arise at multiple steps of the analysis. The effect of different assumptions or methods is used to get an estimate of the systematic uncertainties. Most sources of systematic uncertainty effect the efficiency determination and the fit to extract the signal yields. All assigned relative systematic uncertainties on the ratio of branching fractions are listed in Table 5.6. The largest single systematic uncertainty is the external value of the branching fraction of $\psi(2S)$ to two lepton. There is no dominant source of systematic uncertainty but multiple sources between $1-2\%$. All systematic uncertainties are summed up in quadrature to determine the total systematic uncertainty of 3.83%. Compared with the statistical uncertainty of the measurement, the systematic uncertainty is slightly smaller. The individual sources of systematic uncertainties are discussed in the following in detail.

Table 5.6: Systematic uncertainties on the ratio of branching fractions.

	value
Limited size of simulation	1.10%
Binning choice	1.61%
Interpolation	negligible
Trigger efficiency	1.21%
Fit model	1.57%
Resonant bkg in LL	negligible
PID requirements	negligible
Helicity structure	1.33%
$\mathcal{B}(c\bar{c} \to \ell\ell)$	2.22%
Total	3.83%
Total without $\mathcal{B}(c\bar{c} \to \ell\ell)$	3.13%

5.6.1 Limited Size of Simulation

As the per-candidate efficiency is determined with simulation and the size of the simulated data set is limited by the computing power, statistical fluctuations occur that lead to systematic uncertainty on the measured result. To estimate the effect of these fluctuations, pseudo-experiments are performed. For each pseudo-experiment the binned efficiencies $\epsilon(\eta, p_{\mathrm{T}})$ are varied for each bin assuming binomial uncertainty

$$\sigma_{\epsilon(\eta,p_{\mathrm{T}})} = \sqrt{\frac{\epsilon(\eta, p_{\mathrm{T}}) \cdot (1 - \epsilon(\eta, p_{\mathrm{T}}))}{N(\eta, p_{\mathrm{T}})}}. \tag{5.7}$$

If the variation leads to a negative efficiency in a bin, the variation is repeated. These varied binned efficiencies are then interpolated. This procedure is performed for each track type and decay channel separately. The varied efficiencies are then used as candidate-weights in the fit and the efficiency corrected yield are extracted as it is done in the nominal result. This procedure is repeated 4000 times for each decay channel and track type. The relative difference in the signal yields with a Gaussian fit to the distribution for each track type and decay channel is shown in Fig. 5.13. The distribution follows a Gaussian distribution as expected. A small bias to positive differences is visible, which occur due to the re-variation, when an efficiency is varied to a negative value. This bias is negligible compared to the width of the distribution and is therefore not taken into account. The width of the distribution $\sigma(N_{\text{MC size}})$, is assigned as systematic uncertainty due to the limited size of the simulation. The widths and a comparison to the naive estimation from unbinned efficiencies is given in Table 5.7. As expected the effect of the limited simulation sample size increases with the method of binning the efficiencies but the effect is small. The overall values are propagated to the ratio of branching fractions and result in an uncertainty of 1.1%, which is assigned as systematic uncertainty due to the limited size of the simulation.

Table 5.7: Uncertainty on the effective efficiencies $\sigma_{N_{\text{MC size}}}$ for both decay channels and track types. In parenthesis are the binomial uncertainties from simulation without using any binning.

	$\Lambda_b^0 \to \psi(2S)\Lambda$		$\Lambda_b^0 \to J/\psi\Lambda$	
Track type	long	downstream	long	downstream
$\sigma_{N_{\text{simulation size}}}$ [%]	1.34 (1.0)	0.66 (0.50)	1.94 (1.27)	0.80 (0.69)

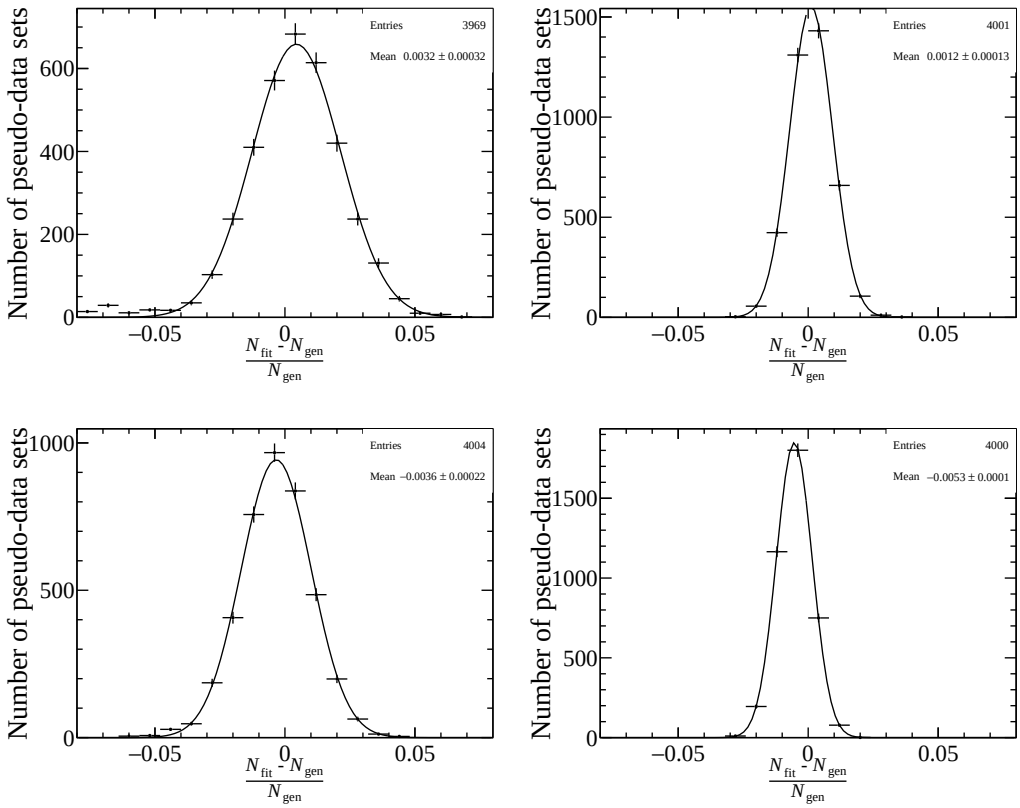

Figure 5.13: Relative difference between the fitted yield after varying the efficiencies and the nominal yield with a Gaussian fit to the distribution for (top) $\Lambda_b^0 \to J/\psi\Lambda$ and (bottom) $\Lambda_b^0 \to \psi(2S)\Lambda$ candidates for (left) long and (right) downstream track.

5.6.2 Binning Choice

For the baseline result a 7x7 binning scheme in $p_{\mathrm{T}}(\Lambda_b^0)$ and $\eta(\Lambda_b^0)$ is used. This binning scheme is chosen to ensure that every bin contains at least 100 reconstructed simulated events. As this number is rather arbitrary, the analysis is rerun with different binning schemes. For this, the interpolation of the binned efficiencies is performed in the same way as in the baseline result. With these new efficiency maps, the efficiency-corrected yields are re-determined. The relative difference of these results using these new determined yields to the nominal result are shown in Fig. 5.14. The maximum relative difference of the results for the different binnings from 4x4 to 10x10, which are reasonable binnings, to the nominal binning is 1.61%. This value is assigned as systematic uncertainty due to the choice of the binning scheme.

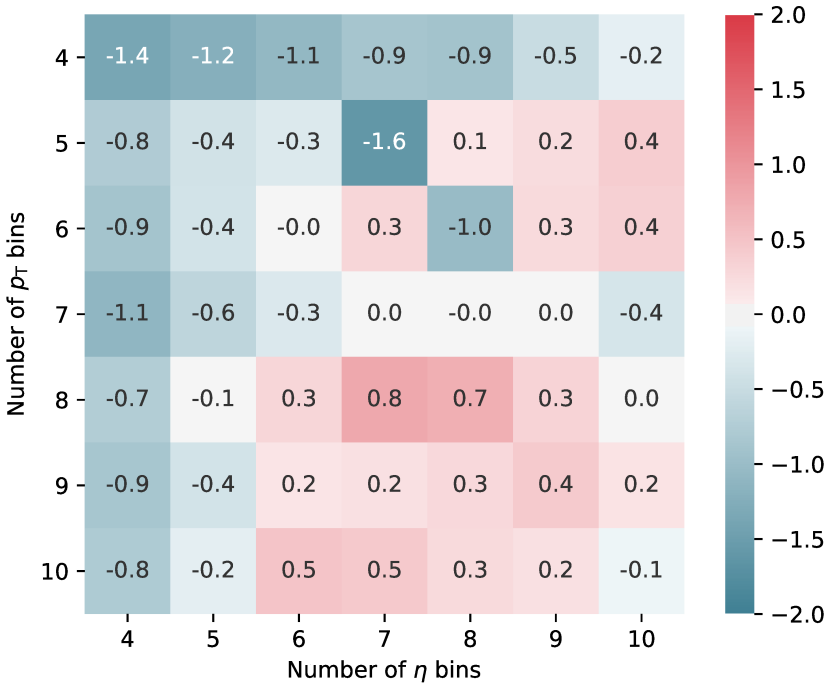

Figure 5.14: Difference of the ratio of branching fractions to the nominal result for different number of η and p_{T} bins used for the efficiency determination and the interpolation afterwards.

5.6.3 Bias from the linear Interpolation

The linear interpolation does not necessary conserve the integral over the two dimensional efficiency plane. To estimate a possible bias, new binned efficiencies are sampled from the nominal interpolated efficiency map and these binned efficiencies are rescaled to match the normalisation of the nominal binned efficiencies. These new binned efficiencies are linear interpolated. To estimate the difference between this new and the nominal efficiency map the $\Lambda_b^0(p_\mathrm{T}, \eta)$ distribution from simulation is used. The mean efficiency of the new efficiency map has a clear bias to lower values. The relative differences between the mean efficiencies $\Delta\epsilon$ for the different samples are

$$\begin{aligned}
\Delta\epsilon_{\mathrm{LL}}^{\psi(2S)} &= -4.22\%, \\
\Delta\epsilon_{\mathrm{DD}}^{\psi(2S)} &= -2.40\%, \\
\Delta\epsilon_{\mathrm{LL}}^{J/\psi} &= -3.82\%, \\
\Delta\epsilon_{\mathrm{DD}}^{J/\psi} &= -2.45\%.
\end{aligned}$$

This bias cancels largely in the ratio especially for the downstream candidates, which are more important for the weighted average. Propagating the remaining differences to the final result leads to a relative difference of 0.1% and is therefore negligible.

5.6.4 Trigger Efficiencies

It is known that the trigger responds and therefore the trigger efficiency are not optimal described in the simulation. To cross-check the used trigger efficiencies from simulation, the so called TISTOS method [78] is used to extract the trigger efficiency from recorded data. The uncertainty on the derived efficiency using this method depends on the available sample size. As the available number of candidates in the Λ_b^0 channels is not sufficient, two additional control channels $B^+ \to J/\psi K^+$ and $B^+ \to \psi(2S)K^+$ with high occupancy are used. Since only trigger requirements on the muons are placed, the difference between simulation and data is expected to be similar between $\Lambda_b^0 \to \psi\Lambda$ decays and the $B^+ \to \psi K^+$ decays. The muon momentum as a main characteristic in the trigger is very similar between the different mother particles, which is shown in Fig. 5.15.

To cancel further remaining differences, the ratio of the trigger efficiencies used. The trigger decision can be divided in trigger decisions that are independent of the signal particles (TIS) and decisions that are based on the signal particles (TOS)

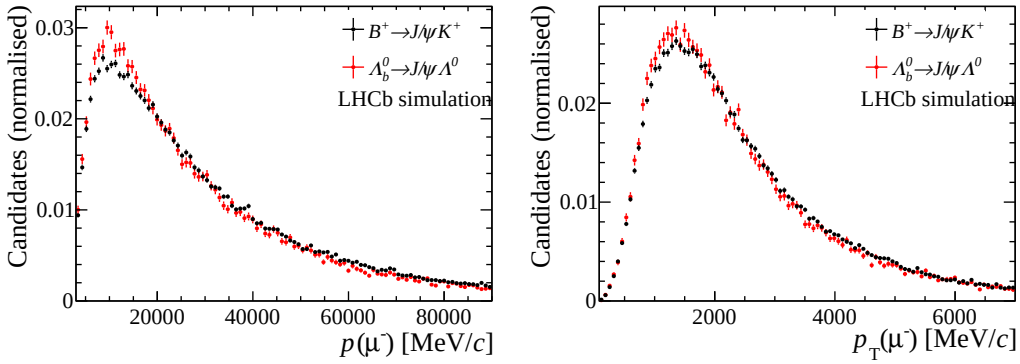

Figure 5.15: Distribution of the (left) muon momentum and (right) transverse momentum or $\Lambda_b^0 \to J/\psi\,\Lambda$ and $B^+ \to J/\psi\,K^+$ simulation.

The main assumption of the TISTOS method is that the trigger efficiency is the same on a sample that is TIS compared to all produced signal decays

$$\epsilon_{\text{trig}} = \frac{N_{\text{trig}}}{N_{\text{total}}} = \frac{N_{\text{trig}}}{\frac{N_{\text{TIS}}}{\epsilon_{\text{TIS}}}}, \tag{5.8}$$

where N_{trig} are the triggered events, N_{total} is the unknown number of total events, N_{TIS} is the number of TIS events and ϵ_{TIS} denotes the efficiency of the TIS trigger. Further, it is assumed that the TIS efficiency is independent of the TOS efficiency. Therefore, the TIS efficiency can be determined using the number of TOS events N_{TOS} and the number of TISTOS events N_{TISTOS}, which are TIS and TOS triggered, as

$$\epsilon_{\text{TIS}} = \frac{N_{\text{TISTOS}}}{N_{\text{TOS}}}. \tag{5.9}$$

Nevertheless this assumption is not true because b quarks are produced in pairs. Therefore, the underlying event in the signal event is more likely to cause a positive trigger decision than in an event without a b hadron. To reduce this correlation, the events are binned in phase space of transverse momentum p_{T} and longitudinal momentum p_{Z} of the b hadron. The correlation decreases with a smaller binning but the statistical uncertainty increases. To determine a optimal binning in p_{T} and p_{Z}, a grid search in the number of bins in the variables is used and the ratio of trigger efficiencies between J/ψ and $\psi(2S)$ channel, r_{trig}, is determined with the TISTOS method on simulation. This value is then compared to the true efficiency from simulation. The result of the grid search is given in Fig. 5.16. The relative difference between the true and the TISTOS efficiency ratio converges fast with the

number of p_T and p_Z bins to a small value, which is comparable with zero within one standard deviation. Thus, a binning in the converged area of four p_T bins and four p_Z bins is used. When using the TISTOS method with this binning on $B^+ \to J/\psi K^+$ and $B^+ \to \psi(2S)K^+$ recorded data an efficiency ratio of $r_\mathrm{trig}^\mathrm{TISTOS} = 1.056 \pm 0.013$ is determined, which differs from the efficiency ratio evaluated on simulated $B^+ \to J/\psi K^+$ and $B^+ \to \psi(2S)K^+$ data $r_\mathrm{trig}^\mathrm{MC} = 1.058 \pm 0.001$ by relative 0.19%. This difference is much smaller than the statistical uncertainty arising from the TISTOS method. Thus the relative statistical uncertainty on the efficiency ratio evaluated on recorded data with the TISTOS method $\sigma(r_\mathrm{trig}^\mathrm{TISTOS}) = 1.21\%$ is assigned as systematic uncertainty due to the trigger efficiency.

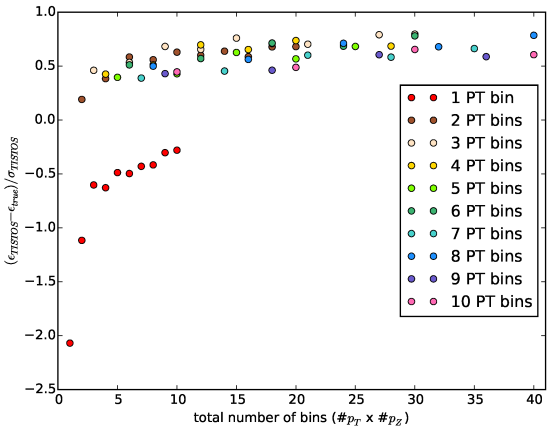

Figure 5.16: The relative bias of the TISTOS method for different binnings in p_T and p_Z.

5.6.5 Fit Model

There is no strong physical motivation for choosing a Hypatia PDF as signal model and an exponential function to describe the combinatorial background. There are other models, that describe the mass distribution reasonably. The considered alternative models are a double Gaussian as signal model and linear PDF as combinatorial background model. The effect of using these alternative models is checked using pseudo experiments. Between 9000 and 10000 data sets are generated according to the alternative descriptions, where the shape is determined with a fit to the $\Lambda_b^0 \to J/\psi\Lambda$ data. The generated yields are fluctuated according to the uncertainties from the fit. Three different combinations are tested:

- A Hypatia for the signal and linear PDF for the combinatorial background
- A double Gaussian for the signal and an exponential for the combinatorial background
- A double Gaussian for the signal and a linear PDF for the combinatorial background

The pseudo data sets are fitted with the baseline fit model. The difference between generated signal yield, N_{gen}, and the fitted yield, N_{fit}, relative to the generated yield $\frac{N_{\text{fit}}-N_{\text{gen}}}{N_{\text{gen}}}$ is filled in a histogram for each pseudo data set. This is shown in Fig. 5.17 separate for downstream- and long-track reconstructed candidates and for the different combinations. The mean of each of these histograms is a measure for the systematic uncertainty of the signal yield using the nominal fit model. There is only a very small difference between the different background models, while the main uncertainty arises from the signal fit model. The effect is expected to cancel largely between both decay channels since the same signal model is used. To be conservative, the largest uncertainty of 1.56 % is assigned as systematic uncertainty due to the fit model. A possible general fit bias is checked by generating toy data sets with the nominal fit model and using the same model for the fit, which results in a negligible difference.

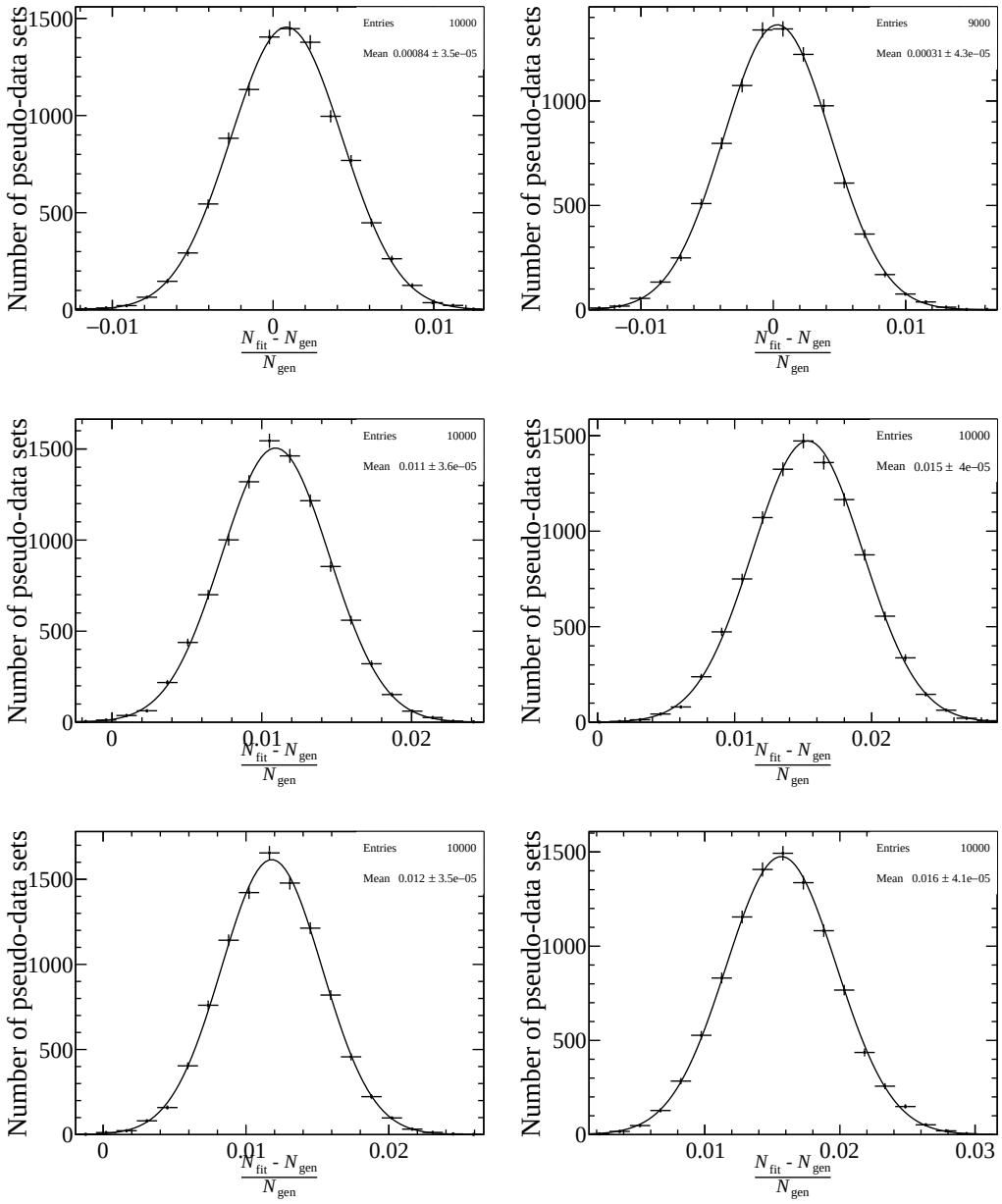

Figure 5.17: Distribution of the relative difference between generated signal yields and fitted signal yields (top) generated with a linear function for the combinatorial background, (middle) generated with a double Gaussian for the signal, and (bottom) generated with a linear PDF for the combinatorial background and with a double Gaussian for the signal for (left) downstream- and (right) long-track reconstructed candidates.

5.6.6 Neglecting Resonant Backgrounds for long-track reconstructed Candidates

In the long-track reconstructed $\Lambda_b^0 \to J/\psi\Lambda$ ($\Lambda_b^0 \to \psi(2S)\Lambda$) data sets, resonant $B^0 \to J/\psi K_S^0$ ($B^0 \to \psi(2S)K_S^0$) backgrounds decays are vetoed and therefore negected in the fit. Similary, $\Xi_b^- \to J/\psi\Xi^-$ ($\Xi_b^- \to \psi(2S)\Xi^-$) are negected in the long track data sets due to its kinematic structure. Both is discussed in detail in Sec. 5.4.1. To check that these backgrounds are negigble as determined, the fit is rerun but with the yields of these components left floating. The resulting fits are shown in Fig. 5.18 and the resulting yields are given in Table 5.8. Using these signal yields and calculating the ratio of branching fractions, leads to the very same result as the nominal result

$$\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\Lambda)} = 0.521.$$

Thus, the effect is negligible.

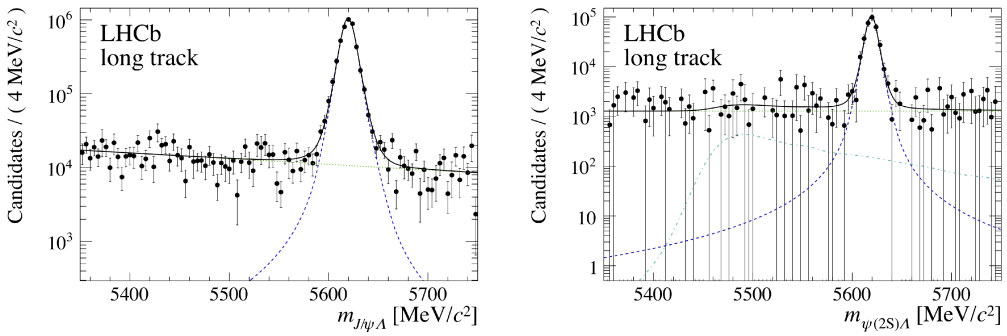

Figure 5.18: Mass fit on (left) $\Lambda_b^0 \to J/\psi\Lambda$ and (right) $\Lambda_b^0 \to \psi(2S)\Lambda$ data for long-track reconstructed candidates with the signal (blue, dashed), the combinatorial background (green, dotted), the peaking $B^0 \to J/\psi K_S^0 / B^0 \to \psi(2S)K_S^0$ background (turquoise, dash-dotted), and the peaking $\Xi_b^- \to J/\psi\Xi^- / \Xi_b^- \to \psi(2S)\Xi^-$ background (violet, dash-dotted) after the selection is applied.

Table 5.8: Yields of the three components of the fit for long-track reconstructed candidates in both decay channels with floating resonant background.

	N_{Signal}	$N_{B^0 \to \psi K_S^0}$	$N_{\Xi_b^- \to \psi\Xi^-}$	$N_{\text{combinatorial}}$
$\Lambda_b^0 \to J/\psi\Lambda$	455920 ± 92560	0 ± 48	0 ± 6	1237730 ± 91613
$\Lambda_b^0 \to \psi(2S)\Lambda$	325486 ± 21865	14720 ± 18396	0 ± 1	128213 ± 22176

5.6.7 Particle Identification Requirements

In the selection PID requirements are used. The simulation does not describe PID perfectly. Therefore, data-driven methods are used to evaluate the resulting effect on the measurement. The used PID requirements are

$$\mu : \mathrm{DLL}_{\mu\pi} > -3$$
$$p : \mathrm{DLL}_{p\pi} > -5 \text{ (only long track)}.$$

The distributions of these variables of sWeighted $\Lambda_b^0 \to J/\psi\Lambda$ data and simulation after the selection is shown in Fig. 5.19. The μ $\mathrm{DLL}_{\mu\pi}$ distribution is well described

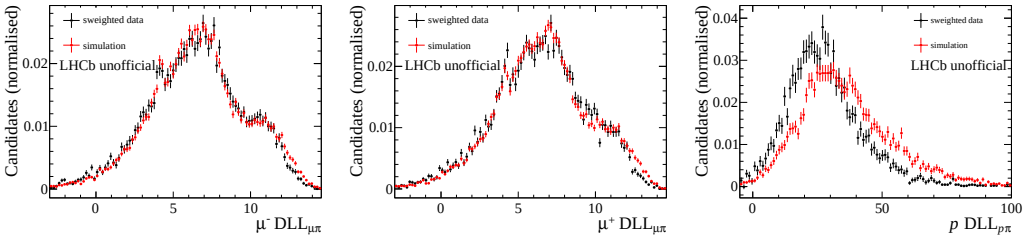

Figure 5.19: Comparison of the used PID variables: (left) μ^- $\mathrm{DLL}_{\mu\pi}$, (middle) μ^- $\mathrm{DLL}_{\mu\pi}$, (right) p $\mathrm{DLL}_{p\pi}$ for long-track reconstructed candidates between sWeighted data (black) and not corrected simulation (red) for $\Lambda_b^0 \to J/\psi\Lambda$ candidates after the preselection.

on simulation but the p $\mathrm{DLL}_{p\pi}$ is shifted. To correct the p $\mathrm{DLL}_{p\pi}$ distribution on simulation, an unbinned PID resampling using kernel density estimation method [79], provided by the LHCb PID group is used. With the assumption that the DLL variables depend mainly on the momentum, the pseudorapidity and the number of best tracks in the event, four dimensional PDFs can be extracted from very pure selected recorded data sets using unbinned kernel density estimation. With this PDF and the values for the three observables new per-candidate DLL values can be generated. The resampled p $\mathrm{DLL}_{p\pi}$ matches well the one on sWeighted data, which can be seen in Fig. 5.20. The difference between resampled and baseline efficiencies ($\eta(\Lambda_b^0)$ and $p_\mathrm{T}(\Lambda_b^0)$ dependent) is determined and added in the fit for the long track candidates. The resulting yields are lower by 1.18% and 1.14% for $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$, respectively. Thus, this effect cancels in the ratio and the difference in the branching ratio is below 0.05% and negligible.

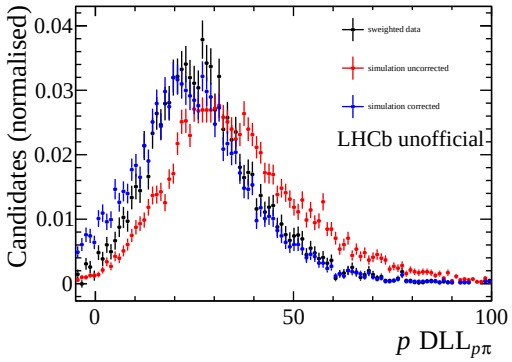

Figure 5.20: Comparison of p DLL$_{p\pi}$ between sWeighted data (black), not corrected simulation (red) and resampled simulation (blue) for long-track reconstructed $\Lambda_b^0 \to J/\psi\Lambda$ candidates.

5.6.8 Effect of Reweighting the Helicity Structure in Simulation to sWeighted Data

After reweighting the helicity structure of the decay to theory predictions (Sec. 5.3.1) for the simulation, the description of the relevant angular distributions is not perfect. Therefore, the effect of this mismatch is evaluated by reweighting the angular distributions $(\cos(\Theta_1), \cos(\Theta_2))$ to match those seen on sWeighted data. The reweighting is done independently for both angular distributions. No split in track types is done, since the fluctuations in the long track $\Lambda_b^0 \to \psi(2S)\Lambda$ data set would be too large. Therefore, the same reweighting is applied for both track types. The distributions after applying the reweighting in $\cos(\Theta_1)$ and $\cos(\Theta_2)$ to the simulation are shown in Figs. 5.21 and 5.22 for both decay channels and track types. Disagreement can be seen in the long track $\Lambda_b^0 \to \psi(2S)\Lambda$ distributions. As the simulation is reweighted to sWeighted data, which is a combined sample for both track types, the reweighting is more dominated by the downstream-track reconstructed data. This disagreement increases further, as the data fluctuates more in the $\Lambda_b^0 \to \psi(2S)\Lambda$ mode since the number of candidates is lower.

With these weights the efficiency maps are recalculated and the efficiency corrected yields are determined as for the baseline result. The ratio of branching fractions changes by 1.33%, which is assigned as systematic uncertainty.

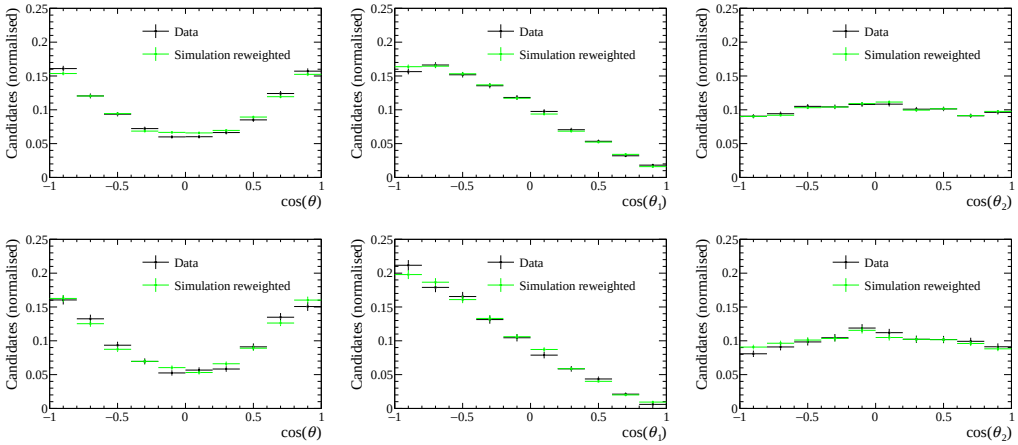

Figure 5.21: Distributions for (left) $\cos(\Theta)$, (middle) $\cos(\Theta_1)$, $\cos(\Theta_2)$ (right) for sWeighted Data (blue) and reweighted simulation (green) for (top) downstream- and (bottom) long track $\Lambda_b^0 \to J/\psi\,\Lambda$ decays.

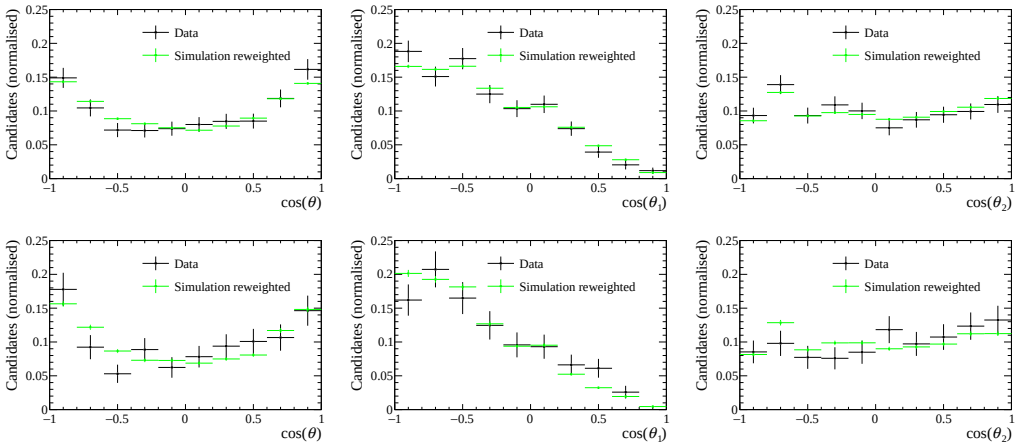

Figure 5.22: Distributions for (left) $\cos(\Theta)$, (middle) $\cos(\Theta_1)$, $\cos(\Theta_2)$ (right) for sWeighted Data (blue) and reweighted simulation (green) for (top) downstream- and (bottom) long track $\Lambda_b^0 \to \psi(2S)\Lambda$ decays.

5.7 Result

The ratio of branching fractions of $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ decays is determined to be

$$\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)} = 0.513 \pm 0.023\,(\text{stat}) \pm 0.016\,(\text{syst}) \pm 0.011(\mathcal{B}),$$

where the first uncertainty corresponds to the statistical uncertainty, the second corresponds to systematic uncertainty and the last one is due to used branching fractions of $\psi(2S)$ and J/ψ to two muons. The measurement it very compatible with the measurement by the ATLAS collaboration [4] with a similar systematic uncertainty but with a better statistical uncertainty. Further, the measurement confirms a tension seen to a theory prediction [5, 6].

6 Measurement of *CPV* in $B^0 \to \psi K^0_S$ Decays

The measurement of the CKM triangle through *CP* violation (*CPV*) measurements is a fundamental test of the SM as discussed in Sec. 2.3. For the CKM angle β the golden channel both from theoretical as well as experimental point of view is $B^0 \to J/\psi K^0_S$. The LHCb collaboration has published multiple measurements of $\sin(2\beta)$ [80, 81, 82] using the Run I data set of LHCb. The presented measurement in this thesis extends these measurement using the Run II data set. This data set is roughly 3.4 times the size of the Run I data set, due to the higher luminosity and the higher centre-of-mass energy, which enlarges the $b\bar{b}$ cross section. As in the published measurements not only the golden channel $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ is used but also J/ψ reconstructed through the decay into two electrons as well as the higher charmonium state $\psi(2S)$. To further increase the sensitivity compared to the Run I measurements, additional reconstruction methods are used for K^0_S mesons in the $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$. All used data sets are discussed in Sec. 6.1. With the increasing precision in the measurement of $\sin(2\beta)$, theoretical uncertainty arises due to so-called penguin pollution. Penguin pollution describe the effect of higher order (penguin) diagrams contributing to the decay, which do not involved the same phases as the tree diagram. Hence, these contributions dilute the measurement. One ingredient to estimate these penguin pollution in $B^0 \to J/\psi K^0_S$ is to measure *CPV* in $B^0_s \to J/\psi K^0_S$, where contribution from penguin diagrams are not heavily suppressed compared to the tree level diagrams. Since the measurement of *CPV* in $B^0_s \to J/\psi K^0_S$ is experimentally very similar to the other decay channels in this analysis, it is also included. The *CPV* in the decays is measured though a time-dependent *CP* asymmetry, which is theoretically discussed in Sec. 2.3.2. To measure this *CP* asymmetry, different ingredients are needed, which are discussed in the following. Starting with the signal selection described in Sec. 6.2, followed by the B meson decay time description (Sec. 6.4) and the flavour tagging discussed in Sec. 6.5. All these inputs are used in the final *CP* fit to measure *CPV* in all decay channels, which is described in Sec. 6.6. First systematic uncertainties are evaluated in Sec. 6.7. Finally, a discussion of the current status is given in Sec. 6.8. Since the analysis is in internal review and not all systematic uncertainties are fully evaluated, the analysis is blinded to ensure no unconscious analyst bias to be introduced. The blinding is done by shifting the central values of the measured CP parameters by a random number. This procedure allows to extract the sensitivity of the measurement.

6.1 Data Sets

The analysed data is collected by the LHCb experiment and corresponds to an integrated luminosity of 6 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$.

Up to four different track type combinations are used to reconstruct K^0_S mesons. Besides the standard combinations of long (LL) and downstream (DD) reconstructed K^0_S, also long-down (LD) and up-long (UL) combinations are used for $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ decays. The different track type combinations to reconstruct a K^0_S meson are discussed in Sec. 3.2. The two non-standard reconstruction methods are only used for $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ decays, since the corresponding data sets in $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi(e^+e^-)K^0_S$ would be too small. These additional types of K^0_S reconstruction add about 15% to the number of candidates, which is in the same order of magnitude as adding $B^0 \to \psi(2S)K^0_S$ or $B^0 \to J/\psi(e^+e^-)K^0_S$ decays. This will be the first measurement using UL reconstructed K^0_S at LHCb.

Since the differences between the conditions in the different years of data taking (2015-2018) is very small, all data is treated in the very same way, but the analysis is performed by splitting into the different track types and the different decay channels. To sum up, there are ten different data sets used, which divide into two data sets for $B^0 \to J/\psi(e^+e^-)K^0_S$ and $B^0 \to \psi(2S)K^0_S$ decays each, four data sets for $B^0 \to J/\psi K^0_S$ decays, and two data sets for $B^0_s \to J/\psi K^0_S$ decays, which is extracted from the $B^0 \to J/\psi K^0_S$ data sets after the selection.

6.2 Signal Selection

The selection of the signal candidates is based on [83] and extended to $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi(e^+e^-)K^0_S$ decays.

After a general cut based preselection and trigger requirements, which are summarised in Sec. 6.2.1, physical backgrounds are vetoed. These backgrounds and their removal is described in Sec. 6.2.2. Besides physical backgrounds, combinatorial background is present in the data sets, which is further suppressed using a multivariate classifier The used multivariate classifier and its input variables are further discussed in Sec. 6.2.3.

6.2.1 Preselection and Trigger Requirements

The general preselection is similar to the selection discussed in the $\Lambda^0_b \to \psi(2S)\Lambda$ analysis (Sec. 5.2) but with generally looser requirements. The detailed requirements

are given in Table A.2 for J/ψ or $\psi(2S)$ reconstructed from the decay to two muons. The preselection for the electron mode differs slightly and is summarised in Table A.3. On top of this general preselection, specific trigger requirements are set. In the muon modes it is required, that the signal muons triggered one of the corresponding J/ψ or $\psi(2S)$ lines. For the electron modes a different approach is used. Since electrons have a less clear signature in the detector, triggering electrons is less efficient. To increase the signal efficiency, candidates where the trigger decision is not based on the signal particles are also used. The number of possible trigger lines is enlarged, corresponding to nearly every trigger line can be used. The detailed trigger lines for all data sets are listed in Table A.4. This inclusive approach increases the amount of background candidates in the data set. The trigger efficiency ranges from 55% ($B^0 \to J/\psi(e^+e^-)K^0_S$ long track) to 82% ($B^0 \to \psi(2S)K^0_S$ downstream track).

6.2.2 Vetos for physical Backgrounds

The amount of physical backgrounds is small since the K^0_S meson in the final state has a long lifetime as discussed in Sec. 3.2. Thus, it is unique in the LHC environment, as other particles like B, D have significantly lower lifetimes or experimentally no lifetime *e.g.* K^*, ϕ or are stable within the detector like charged pions, kaons or neutrons. Requiring a flight distance longer than 0.5 ps of the K^0_S meson, which is very efficient for signal events, leads to suppress all physical background without a real K^0_S or Λ in the final state.

There are two different kinds of physical backgrounds: First, there is partial reconstructed backgrounds, where not all particles of the decay are reconstructed. The main irreducible partial background is $B^0 \to \psi K^*$, where the K^* meson decays to a K^0_S meson and a neutral pion, which is not reconstructed. The missing pion leads to a mass distribution of these decays, which starts at least a pion mass below the B^0 mass with a tail to lower masses. Thus, this contribution interferes only with the lower tail of the signal and is not an issue for the measurement. However, this is a irreducible background, that has to be modelled in the mass fit.

The second type of physical background occur due to misidentification. Here, the main component arises from $\Lambda^0_b \to \psi\Lambda$ decays, where the proton from the Λ baryon decay is misidentified as a pion. Due to the assigned pion mass, which is lower then the proton mass, the distribution starts below the Λ^0_b mass and has a large tail to lower masses, which includes the signal region. This background is removed by recomputing the Λ mass with the proton mass assigned to one of the pions and requiring a particle identification incompatible with a proton for this pion, if this combination is near the the known Λ mass ($1100 < m(p\pi) < 1130\,\mathrm{MeV}/c^2$) [16]. Due to the combination of the mass and particle identification requirement, this is

efficient for signal decays and suppresses most $\Lambda_b^0 \to \psi\Lambda$ decays present in the data set.

6.2.3 Boosted Decision Tree

The multivariate classifier, which is used to reduce the combinatorial background, is a gradient boosted decision tree (GBTD) [84]. A boosted decision tree is based on simple decision trees, where at each leaf a discussion is made depending on a variable to separate different classes. Here, two classes are used, signal and combinatorial background. The individual used decision trees have a low depth. These decision trees are iterative defined and added leading to a stronger classifier. Each added decision tree is designed to improve the classification. This procedure is called boosting. Gradient boosting is a special boosting algorithm, where a loss-function is optimized with each added decision tree. Here, the loss-function is the error rate of the classification. The GBTD is supervised trained using signal simulation as signal proxy and upper-mass sideband data as a proxy for combinatorial background. Upper-mass sideband data ($m_{\psi K_S^0} > 5450\,\mathrm{MeV}/c^2$) is used, since the other present physical backgrounds arises from missing particles or mis-identification and therefore mainly affect the lower-mass sideband. Between 18 and 26 variables are used in the GBDT depending on the used data set. These variables are very similar between the different decay channels and track types. They consist of vertex qualities, momenta and impact parameters to distinguish between detached signal particles and prompt background from the primary vertex. The detailed list is given in Tables A.5 and A.6.

To eliminate the effect of over-training, which describes the learning of features of statistical fluctuations in the training data set and to optimize the usage of the available training data, k-folding is used. The k-folding method describes the splitting in k data sets, where the classifier is trained using k-1 data sets and tested on the last one. This procedure results in k-1 different classifiers for every candidate in the training data set and k classifier for candidates that are not in the training data set. To further suppress over-training, the early stopping method [85] is used. Early stopping checks after each added decision tree, whether the performance on the test data set has improved. If the performance has not improved within the last n (here $n = 30$) added decision trees, the training is stopped and the best performing iteration is used. This results in a number of trees between 350-1000 for the different decay channels, depending on the available number of candidates in the training. Every classified candidate gets the mean of all classifiers, that have not seen the candidate during the training, assigned. The resulting classification is a float between 0 (background-like) and 1 (signal-like), which is shown for the

two largest data sets ($B^0 \to J/\psi K^0_S$ LL and DD) in Fig. 6.1, while the other data sets are given in Fig. A.2. No sign of over-training is seen by comparing the BDT classification on the training sample and the test sample.

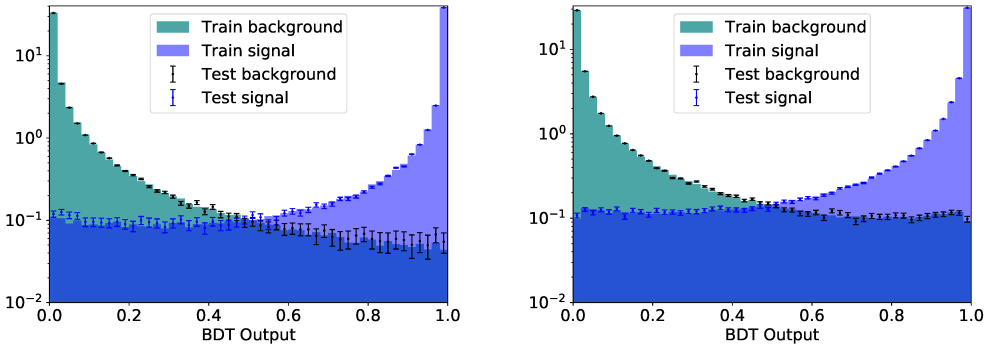

Figure 6.1: The BDT classification output for signal (blue) and background (green/black), separate the training (solid) and test data set (points) for LL (left) and DD (right) $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ data.

To optimise the cut on the resulting classification, a figure of merit (FOM) is maximised. The signal significance is used as FOM. It is defined as FOM $= S/\sigma(S)$, where S is the number of signal candidates and $\sigma(S)$ its uncertainty. This FOM is evaluated using fits to data using the mass model described in Sec. 6.3 for different requirements on the classification. The resulting curves of the FOM in dependence of the BDT cut with the maximum positions indicated are given in Fig. 6.2. The FOM trend is mostly steady but especially in the smaller data sets, larger peaks are possible, which are mainly do to imperfect fits. These maxima in the FOM are not used.

Due to the peaking nature of the classification, the FOM shows a larger plateau region, where the maximum is found. This ensures that the final result does not strongly depend on from the explicit cut value. Note that for $B^0_s \to J/\psi K^0_S$ decays the same classifier is used as for $B^0 \to J/\psi K^0_S$ decays but with a stronger cut due to the lower number of signal candidates. The selection efficiency split for each selection step is given in Tables 6.1 and A.7 to A.9. The difference in the preselection and reconstruction efficiency reflects the different occurrence of the track types. The differences in the other efficiencies are small between the track types. The trigger efficiency is mainly dependent on the ψ meson especially for the muon channels. The BDT efficiencies mainly differ due to the different BDT requirements, which are driven by the different signal yields between the track types.

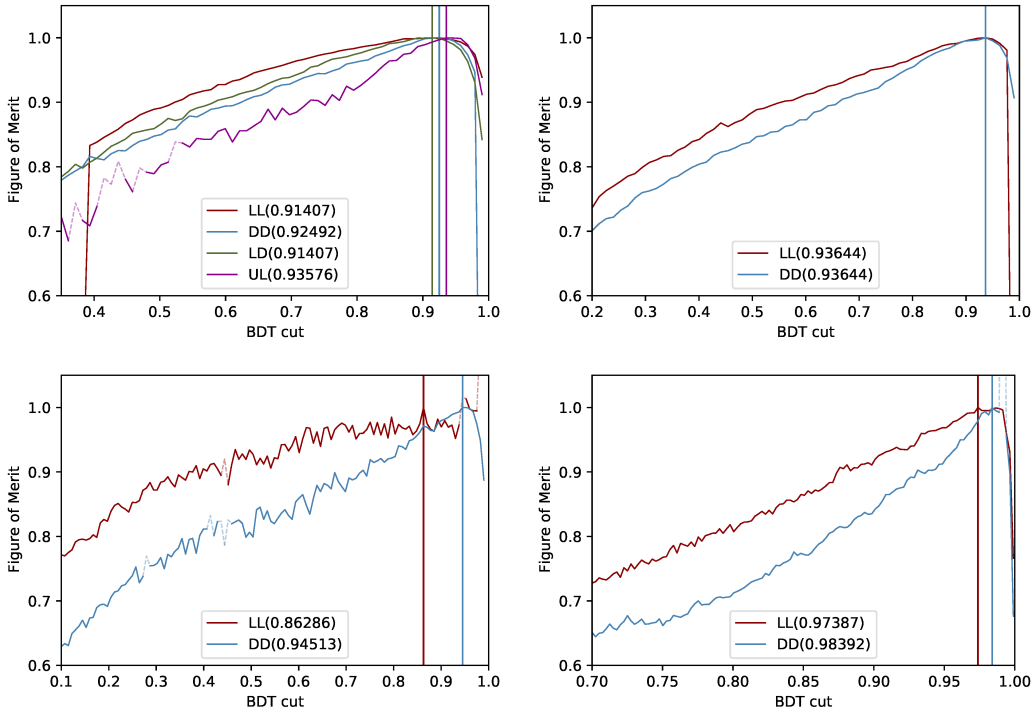

Figure 6.2: The Figures of merit normalised to the maximum in dependence of the BDT cut point for the modes $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ (upper left), $B^0 \to \psi(2S)K^0_S$ (upper right), $B^0 \to J/\psi(e^+e^-)K^0_S$ (lower left) and $B^0_s \to J/\psi\, K^0_S$ (lower right) for each track type, with the position of the maximum indicated.

Table 6.1: Selection and reconstruction efficiencies for $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates determined on simulation.

step	efficiency [%]			
	LL	DD	LD	UL
detector acceptance		19.897 ± 0.083		
presel. and reco.	2.1574 ± 0.0029	5.006 ± 0.004	0.3307 ± 0.0012	0.7292 ± 0.0017
trigger	80.79 ± 0.05	80.879 ± 0.031	82.69 ± 0.12	82.33 ± 0.08
BDT	86.45 ± 0.05	78.93 ± 0.04	78.90 ± 0.14	78.52 ± 0.10
vetos	96.576 ± 0.025	98.220 ± 0.012	98.92 ± 0.04	97.69 ± 0.04
total	1.4516 ± 0.0024	3.1314 ± 0.0032	0.2120 ± 0.0009	0.4555 ± 0.0013

6.3 Massfit and Signal Weight Extraction

The reconstructed B^0 mass is used to disentangle signal and background components. To describe the different components, different PDFs are used: The $B^0 \to \psi K^0_S$ and $B^0_s \to \psi K^0_S$ decays are described with a double-sided Hypatia PDF [75], where the tail parameters are fixed to values from simulation, while the width and the mean is left floating in the fit. The definition of the Hypatia function is given in Sec. 5.4.2. The mean of the $B^0_s \to \psi K^0_S$ Hypatia is shifted by the known B^0-B^0_s mass difference [16]. The combinatorial background is described using an exponential function. As discussed in Sec. 6.2.2, the residual backgrounds from partially reconstructed backgrounds needs to be modelled. This is done using a Gaussian function with an exponential tail, $\mathcal{G}(\sigma_{\text{part}}, \alpha_{\text{part}}, \mu_{\text{part}})$. The final mass fit model $\mathcal{P}$ is the sum of all these components

$$\mathcal{P} = N_{B^0 \to \psi K^0_S} \cdot \mathcal{J}_{B^0 \to \psi K^0_S} + N_{B^0_s \to \psi K^0_S} \cdot \mathcal{J}_{B^0_s \to \psi K^0_S} + N_{\text{partial}} \cdot \mathcal{G} + N_{\text{comb}} \cdot \text{exp}, \tag{6.1}$$

with individual yields that are floating in the fit.

The mass fits are performed individual for all the ten selected data sets in the range from 5100 to 5900 MeV/c^2. The resulting fits are shown in Figs. 6.3 and 6.4 for each of the data sets, and the resulting yields for each component are given in Table 6.2. Since in the *CP* fit of the $B^0_s \to J/\psi K^0_S$ candidates, the decay time acceptance is taken from $B^0 \to J/\psi K^0_S$ data (discussed in detail in Sec. 6.6), the mass range is decreased to $5320 - 5400\,\text{MeV}/c^2$ for the $B^0_s \to J/\psi K^0_S$ data set to decrease the number of events that are used twice in the analyses. Based on the fits on the whole range, sweights [68, 69] are calculated, that are used in the final CP fit.

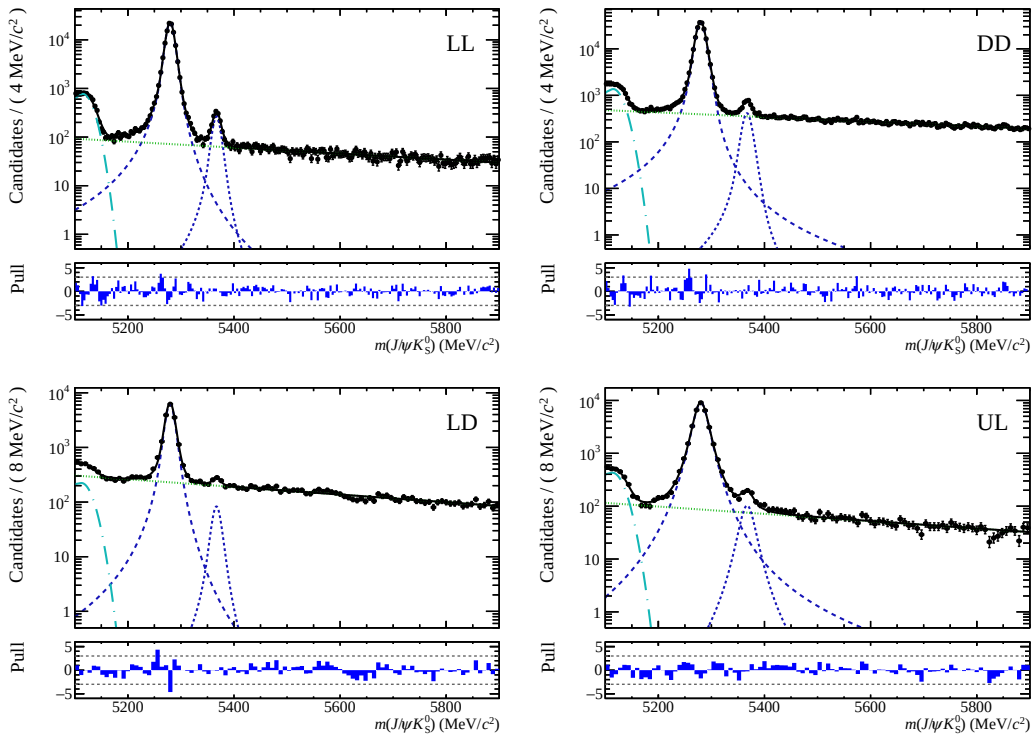

Figure 6.3: Mass fit on selected data for (upper-left) LL, (upper-right) DD, (lower-left) LD and (lower-right) UL $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ candidates with the signal (blue, long-dashed), the combinatorial background (green, dashed) and the partially reconstructed low-mass background (cyan, long-dashed-dotted) and the peaking $B^0_s \to J/\psi(\mu^+\mu^-)K^0_S$ background (blue, dashed) on the full range.

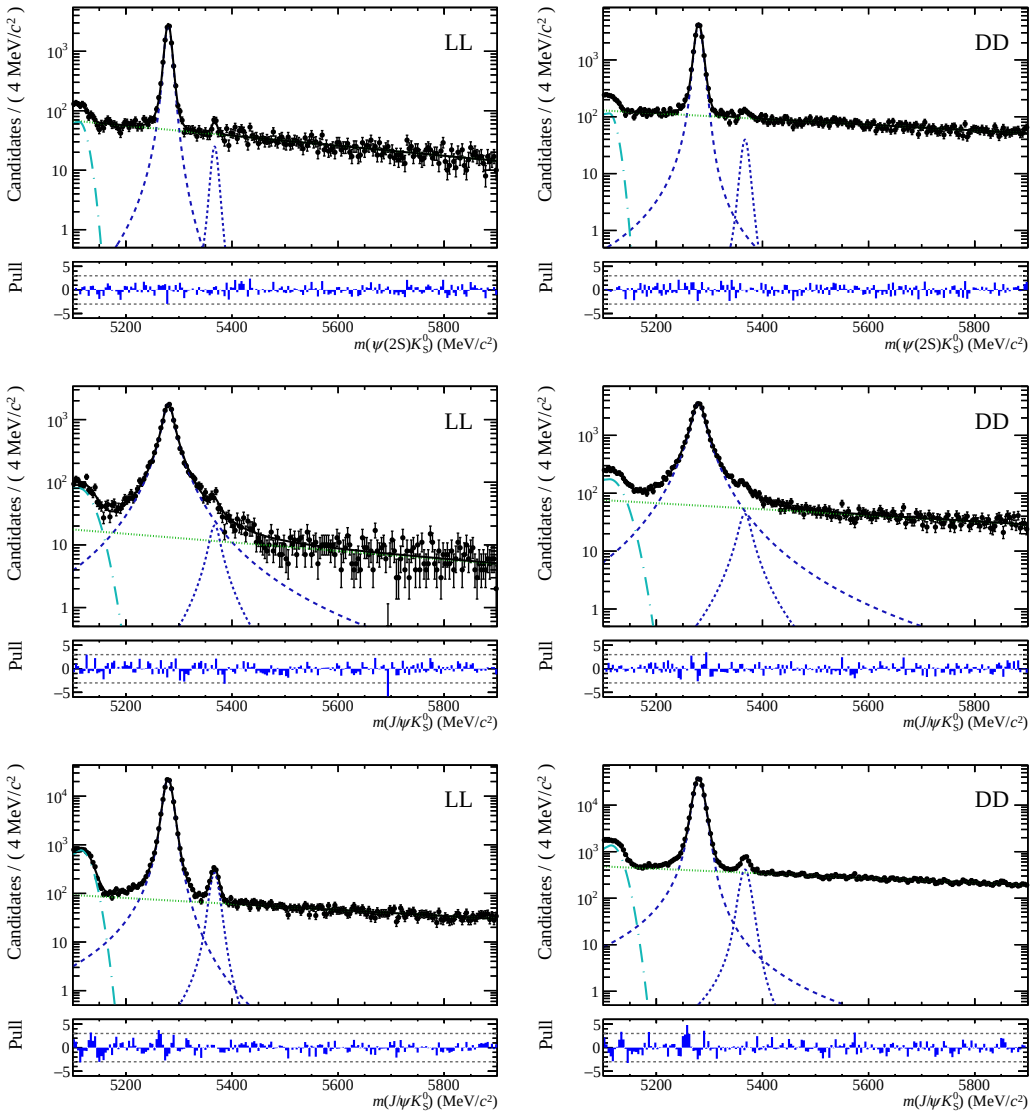

Figure 6.4: Mass fit on selected data for (left) LL and (right) DD to (top) $B^0 \to \psi(2S)K^0_S$, (middle) $B^0 \to J/\psi(e^+e^-)K^0_S$, and (bottom) $B^0 \to J/\psi K^0_S$ candidates with a stronger BDT requirement with the signal (blue, long-dashed), the combinatorial background (green, dashed) and the partially reconstructed background (cyan, long-dashed-dotted) and the peaking corresponding $B^0_s \to \psi K^0_S$ candidates (blue, dashed).

Table 6.2: Fitted yields for the different components for the different data sets split by track type.

Track type	$N_{B^0 \to J/\psi(\mu^+\mu^-)K^0_S}$	$N_{B^0_s \to J/\psi(\mu^+\mu^-)K^0_S}$	N_{comb}	N_{partial}
LL	$105\,324 \pm 340$	1219 ± 44	$11\,085 \pm 183$	6747 ± 98
DD	$199\,011 \pm 508$	2171 ± 79	$62\,997 \pm 425$	$12\,106 \pm 162$
LD	$16\,055 \pm 142$	210 ± 33	$17\,252 \pm 174$	994 ± 69
UL	$34\,051 \pm 205$	383 ± 34	6512 ± 152	2022 ± 66
	$N_{B^0 \to \psi(2S)K^0_S}$	$N_{B^0_s \to \psi(2S)K^0_S}$	N_{comb}	N_{partial}
LL	9925 ± 105	87 ± 18	6907 ± 101	461 ± 40
DD	$17\,384 \pm 143$	159 ± 29	$16\,849 \pm 156$	765 ± 52
	$N_{B^0 \to J/\psi(e^+e^-)K^0_S}$	$N_{B^0_s \to J/\psi(e^+e^-)K^0_S}$	N_{comb}	N_{partial}
LL	$14\,878 \pm 156$	172 ± 31	2000 ± 151	859 ± 61
DD	$32\,661 \pm 219$	391 ± 52	9730 ± 217	1833 ± 87
B^0_s selection	$N_{B^0 \to J/\psi(\mu^+\mu^-)K^0_S}$	$N_{B^0_s \to J/\psi(\mu^+\mu^-)K^0_S}$	N_{comb}	N_{partial}
LL	$97\,155 \pm 315$	1085 ± 37	3462 ± 94	6177 ± 85
DD	$149\,349 \pm 394$	1600 ± 50	$12\,580 \pm 171$	9073 ± 112

6.4 Decay-Time Description

Since time-dependent *CPV* is measured, the B^0 decay time needs to be fully understood and described in the fit. There are three different effects, that define the measured decay time on data. First, there is the physical effective lifetime of the B^0 and B^0_s meson. For B^0 mesons the effective lifetime is equal to the mean lifetime, since the decay width between the two mass eigenstates is very small in the B^0 system. For B^0_s decays, the effective lifetime is specific to the analysed decay, which is given by the parameter $\mathcal{A}_{\Delta\Gamma}$. The mean lifetimes and the decay width differences in the B^0_s system are more precise known, than what could be measured in the analysed decay channels, therefore the known values [16] are constrained within their uncertainties.

The second contribution is the decay-time acceptance, which is a decay time dependent inefficiency arising from the selection and reconstruction. As the selection suppresses mostly prompt backgrounds, which have a very low reconstructed decay time, signal decays with lower decay times are more often rejected by the selection. For the measurement of $\sin(2\beta)$ in B^0_d decays the effect of the decay-time acceptance is small. The parameter S, which translate to $\sin(2\beta)$in the analysed decays, is the coefficient in front of the $\sin(\Delta m \cdot t)$ term. The sine term is small for very small decay times. Hence, the signal candidates mainly affected by the decay-time acceptance have negligible sensitivity on $\sin(2\beta)$. Therefore, the decay-time acceptance can be modelled using cubic splines [86, 87], that are left floating in the fit, without decreasing the sensitivity on $\sin(2\beta)$. Smaller inefficiencies arise from the VELO reconstruction, leading to inefficiencies at very high decay times [88, 89].

In the measurement of *CPV* in $B^0_s \to J/\psi K^0_S$ decays, the decay-time acceptance must be handled differently. Here, the parameter $\mathcal{A}_{\Delta\Gamma}$, which is the coefficient in front of $\sinh(1/2\Delta\Gamma \cdot t)$ and describes the asymmetry of the light and heavy mass eigenstate in the decay is of interest. Letting $\mathcal{A}_{\Delta\Gamma}$ float in the fit corresponds to letting the effective lifetime float in the fit. As floating splines can describe the lifetime, the spline coefficient and $\mathcal{A}_{\Delta\Gamma}$ cannot be left floating at the same time. The solution is to extract the decay-time acceptance from $B^0 \to J/\psi K^0_S$ decays, which are kinematically very similar to $B^0_s \to J/\psi K^0_S$ decays and no measurable difference is expected. This means the extracted acceptance model is fixed in the *CPV* fit on $B^0_s \to J/\psi K^0_S$ data. The determination of the acceptance is further discussed in Sec. 6.4.1.

The third and last part is the decay-time resolution, which results from the momentum and vertex resolution of the detector. The effective decay time resolution varies from roughly 40 to 70 fs depending on the decay channel and selection. Comparing these values to the B^0-$\overline{B}^0$ oscillation frequency $\Delta m_d = 0.5065 \pm 0.0019\,\mathrm{ps}^{-1}$ [16],

the resolution is very small and has no major effect on time dependent measurements in the B^0 system. This is not true for *CPV* in $B_s^0 \to J/\psi K_S^0$ decays, where $\Delta m_s = 17.756 \pm 0.021\,\text{ps}^{-1}$ [16] is much larger compared to the B^0 system. The reconstruction calculates an estimate of the decay-time resolution for every candidate. To calibrate and determine the decay-time resolution, different methods using simulation and prompt data are used. The calibration of the decay-time resolution is fully described in Secs. 6.4.2 and 6.4.3. Further, to reduce the number of figures, all methods are shown using $B^0 \to J/\psi K_S^0$ LL simulation or data in the following, while the figures using the other data sets are given in Appendix A.2.

6.4.1 Decay-Time Acceptance

The decay-time acceptance is a side effect of the selection, which mainly tries to separate short-lived backgrounds from the long-living signal decays. Since for B_d^0 decays the B^0 lifetime is fixed in the fit, the acceptance can be empirically modelled using cubic splines, whose coefficients are left floating in the fit. The knot positions and the number of splines are optimised to describe the simulation as well as the data, while minimising the number of knots. Seven knots are sufficient to describe the acceptance in each decay mode. The B^0 decay time distribution is shown for simulation and data with the fitted model, knot positions and the full acceptance function indicated in Fig. 6.5. As expected, the main effect is at very low decay times below 1 ps and a flat acceptance for larger decay times.

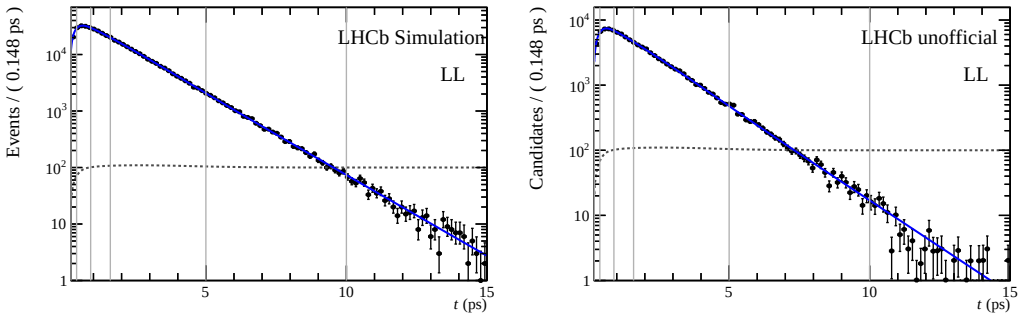

Figure 6.5: Fit to the reconstructed B^0 decay time on (left) signal simulation and (right) data for LL $B^0 \to J/\psi K_S^0$ candidates with the knot position (vertical lines) and the acceptance (dotted-grey). The fits to the other B_d^0 data sets are given in Figs. A.3 and A.4.

For $B_s^0 \to J/\psi K_S^0$ decays a simple decay-time fit is not possible, since in the B_s^0

system the decay width difference between the light and heavy mass eigenstate is not negligible. Therefore, the amount of $B^0_{s,\mathrm{H}}$ and $B^0_{s,\mathrm{L}}$ contributing to the decay, $\mathcal{A}_{\Delta\Gamma}$, needs to be known. As the kinematic differences between $B^0 \to J/\psi K^0_S$ and $B^0_s \to J/\psi K^0_S$ decays are negligible, the acceptance for $B^0_s \to J/\psi K^0_S$ decays is extracted using the $B^0 \to J/\psi K^0_S$ data set, which is selected in the very same way as the $B^0_s \to J/\psi K^0_S$ data set. This spline function is fixed in the final B^0_s CP fit, where $\mathcal{A}_{\Delta\Gamma}$ is a floating parameter.

6.4.2 Decay-Time Resolution on Simulation

The measurement of time-dependent *CPV* is in its core the determination of oscillation amplitudes. The effect of the decay-time resolution is that the oscillation is smeared out and thus the amplitude is dampened. The statistical effect can be estimated using the dilution, D, which is given by

$$D = \int_{-\infty}^{\infty} dt R(t) \cos(\Delta m t), \tag{6.2}$$

where $R(t)$ is the resolution function, and Δm is the oscillation frequency. The dilution describes, the loss in statistical sensitivity due to the time-resolution, where a value of 1 corresponds to a perfect resolution. For a Gaussian kernel with width, σ, which is assumed in the following, the dilution can be calculated as

$$D = \exp\left(-\frac{1}{2}\sigma^2 \Delta m^2\right). \tag{6.3}$$

The effect of multiple Gaussians with a common mean is that the dilutions arising from each Gaussian are added weighted by their fraction. The baseline method in this analysis is to calibrate the decay-time resolution estimate using signal simulation. On simulation, the true decay time is known and can be compared with the reconstructed decay time. Hence, the difference between both decay time results in a Gaussian-like distribution, that can be described by the sum of three Gaussian functions. A fit of this distribution is shown in Fig. 6.6. From these fits the effective mean resolution are determined, which are given in Table A.10 and the corresponding dilutions are given in Table A.11. The effective resolution in the B^0_d channel varies between 60 and 70 fs, which is much lower then the B^0_d mixing oscillation and therefore the effect arising from the decay-time resolution is negligible, resulting in a dilution of $> 99.94\%$. For B^0_s decays, where the mixing oscillation is much faster, the effect is relevant. This results of to a dilution, which is about 70% and corresponds to an effective resolution of 44 to 47 fs. Although the physical resolution is the same between B^0 and B^0_s decays, the effective resolution

evaluated from the dilution is different. This is due to the multi-Gaussian resolution, which is non-linear averaged for the total dilution.

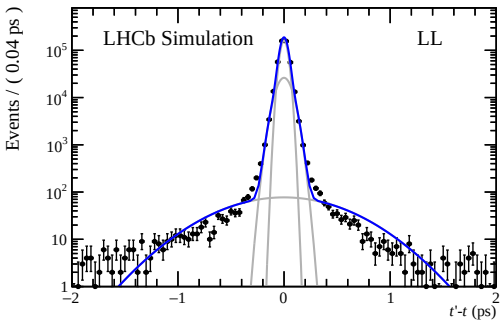

Figure 6.6: Fit to the decay time resolution on signal simulation for LL $B^0 \to J/\psi K^0_S$ candidates with the sum of the three Gaussians (blue) and the single Gaussians (grey). The same fits to the other B^0_d data sets are given in Figs. A.5 and A.6.

Since using a per-event based resolution results in a better sensitivity, the per-event decay-time resolution is used. This estimate from the reconstruction is calibrated with a linear function. To check whether a linear calibration is sufficient, the simulation data is binned in bins of the decay-time resolution estimate and the resolution is fitted in each of bin using the aforementioned model. The fits in the eight bins are shown in Fig. A.7, while the width of the narrow and medium Gaussian in each bin is shown in Fig. 6.7 for long track $B^0 \to J/\psi K^0_S$ simulation.

Two Gaussians are sufficient to describe the resolution in the bins, due to the size of the simulation. However, without the binning a clear need for a third Gaussian is seen. The linear calibration describes the resolution in each bin. To further improve this linear calibration, the calibration is extracted using an unbinned fit to the whole data set using a per-event calibration. The result of this unbinned fit is shown in Fig. 6.8 and the extracted calibration functions are given in Table 6.3.

As expected the calibration from the binned and unbinned fit are compatible but a small bias is seen in $B^0 \to J/\psi(e^+e^-)K^0_S$ simulation. The unbinned calibration is used in the baseline fits.

6.4.3 Decay-Time Resolution from prompt decays

Since there are possible effects that are not perfectly modelled in simulation, the decay-time resolution is also checked on recorded data. This is possible using prompt

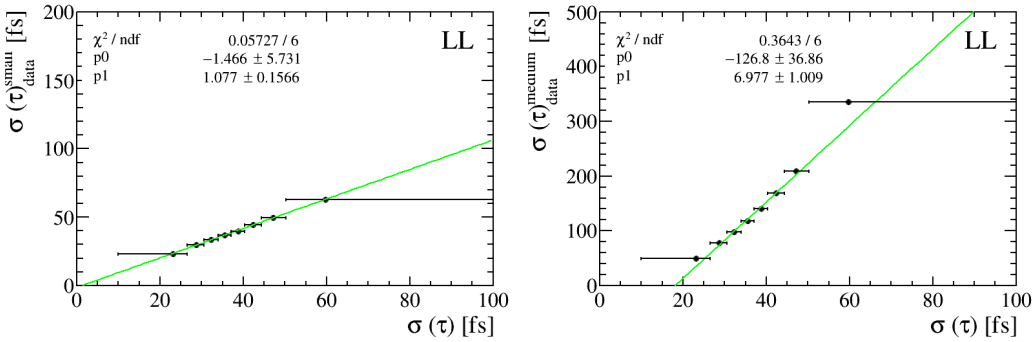

Figure 6.7: Width of the small (left) and the medium Gaussian (right) for long track $B^0 \to J/\psi K^0_S$ simulation in bins of the decay-time resolution estimate. The width are fitted with a linear function. The corresponding figures for the other $B^0 \to J/\psi K^0_S$ data sets are given in Figs. A.8 to A.13.

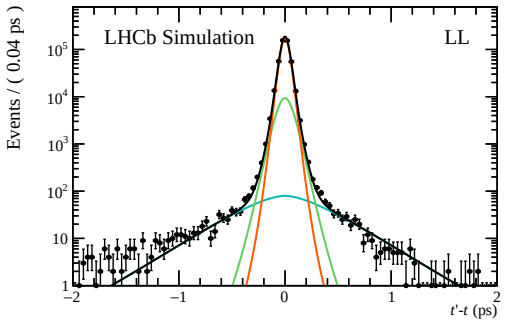

Figure 6.8: Per-candidate fit to the decay time resolution on signal simulation for LL $B^0 \to J/\psi K^0_S$ candidates with the sum of the three Gaussians (black) and the single Gaussians (coloured). The same fits to the other B^0_d data sets are given in Figs. A.14 and A.15

Table 6.3: Per-event (unbinned) calibration parameters for reconstructed candidates for each track type and B^0 decay channel determined on simulation. It is calibrated with a linear function with the slope c and a constant offset b. The fraction between the three Gaussian functions are denoted as f. The fraction f_3 is defined recursive as $1 - f_1 - f_2$.

	$B^0 \to J/\psi(\mu^+\mu^-)K^0_S$							
Track type	LL		DD		LD		UL	
b_1 / ps	−0.023	± 0.005	0.003	± 0.007	0.02	± 0.13	0.022	± 0.025
b_2 / ps	−0.0002	± 0.0006	0.0052	± 0.0005	0.0048	± 0.0025	0.0006	± 0.0013
b_3 / ps	−0.000 05	± 0.000 18	0.000 93	± 0.000 11	−0.0003	± 0.0005	0.0013	± 0.0004
c_1	5.78	± 0.22	4.75	± 0.27	6	± 8	3.5	± 1.0
c_2	1.58	± 0.04	1.479	± 0.024	1.60	± 0.10	1.46	± 0.06
c_3	0.959	± 0.008	0.9256	± 0.0035	0.959	± 0.013	0.918	± 0.015
f_1	0.0090	± 0.0006	0.0065	± 0.0004	0.0030	± 0.0020	0.0073	± 0.0020
f_2	0.177	± 0.015	0.137	± 0.007	0.121	± 0.020	0.208	± 0.032
	$B^0 \to \psi(2S)K^0_S$				$B^0 \to J/\psi(e^+e^-)K^0_S$			
Track type	LL		DD		LL		DD	
b_1 / ps	−0.027	± 0.008	0.011	± 0.007	0.008	± 0.009	0.037	± 0.010
b_2 / ps	0.0001	± 0.0009	0.0046	± 0.0007	0.0048	± 0.0021	0.0125	± 0.0015
b_3 / ps	0.000 49	± 0.000 25	0.001 17	± 0.000 16	0.0038	± 0.0007	0.0052	± 0.0004
c_1	6.25	± 0.35	4.06	± 0.29	3.4	± 0.4	2.87	± 0.35
c_2	1.52	± 0.05	1.42	± 0.04	1.48	± 0.17	1.34	± 0.06
c_3	0.925	± 0.011	0.905	± 0.006	0.87	± 0.04	0.820	± 0.012
f_1	0.0071	± 0.0006	0.0060	± 0.0006	0.028	± 0.009	0.0149	± 0.0031
f_2	0.170	± 0.022	0.142	± 0.012	0.24	± 0.08	0.204	± 0.023

decays. Prompt decays are candidates, where no real B^0 meson is present, but the reconstruction has a combination of a K^0_S and J/ψ coming from the PV. Thus, the true decay time of prompt candidates is zero and the decay-time resolution can be extracted from candidates with a negative reconstructed decay time. No differences between the resolution between $B^0_s \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S$ is expected. This study is performed with $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ pseudo candidates, because the decay-time resolution has a minor effect to the measurement in B^0 decays but a relevant effect in $B^0_s \to J/\psi K^0_S$ decays due to the higher oscillation frequency. Also the $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$ channel has a much larger sensitivity compared to the $B^0 \to \psi(2S)K^0_S$ or $B^0 \to J/\psi(e^+e^-)K^0_S$ channel, therefore the measurement is more sensitive to minor effects. As the baseline selection includes a requirement that the K^0_S and ψ are displaced from the PV, the prompt data sets (simulation as well as recorded data) has to fulfil a different preselection that does not include requirements on the decay time. Further, the BDT would cut out most of the prompt candidates and therefore no BDT is applied. As a first step to ensure that even with this different selection prompt candidates have a similar decay-time resolution, the mean resolution on prompt simulation is compared to the one on signal simulation. The fit to the resolution on prompt simulation is shown in Fig. 6.9. Since the number of candidates is very large in prompt data samples, especially on recorded data, a binned χ^2 fit is used instead of the unbinned maximum likelihood fit. This leads to shorter computing time and robust fits. The resulting effective resolution from prompt simulation is similar to the resolution extracted from the signal selection.

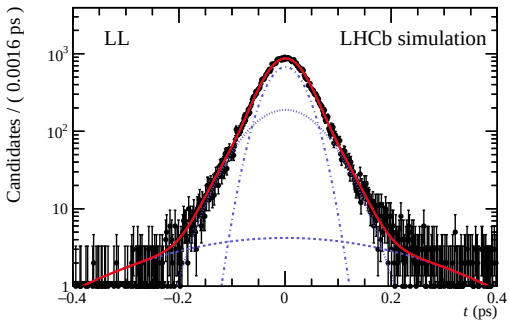

Figure 6.9: Mean decay time resolution on prompt simulation for LL $B^0 \to J/\psi K^0_S$ candidates with the sum of the three Gaussians (red) and the single Gaussians (blue).

In contrast to the prompt simulation, the prompt sample on data does not only contain prompt decays but a mixture ranging from true B decays, short lived

particles and purely combinatoric candidates. The distribution of these prompt data is shown in Fig. 6.10 with a fit. The fit model consist of a triple Gaussian

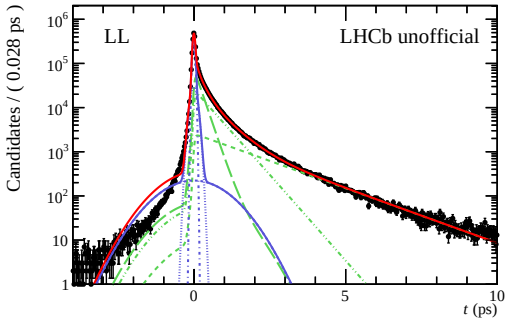

Figure 6.10: Fit to the reconstructed decay time from prompt LL data candidates with the total fit model (red), the sum of the three Gaussians (blue), and the individual Gaussians (purple) and exponential function (green).

for the prompt component and three exponential functions convolved with the same triple Gaussian as resolution function to model the other components in the sample empirically. A non-Gaussian component can be seen for negative decay times, which can not be described by the fit model. To be comparable with the baseline resolution, the resolution model is not changed for the prompt data set. The results of this first fit is given in Table A.12. The fits shows a significant decay-time bias around 10 fs, which means that the mean of the resolution is not zero. This is a known effect in the Run II LHCb data set due to a VELO misalignment in the reconstruction, which is not included in the simulation. One option would be to rerun the whole reconstruction, which would eliminate this bias but is very time and resource consuming. In this analysis the approach is to describe the seen bias and evaluate the effect on the measurement, which is fully described in Sec. 6.7. Since the decay-time bias arising from the misalignment is correlated with the decay kinematics as well as the decay-time resolution estimate, a correlation between the bias and the estimate is expected as well. Therefore to extract the calibration function from prompt data, the resolution calibration parameter are extracted in bins of the decay-time resolution estimate as on simulation but with the above described fit model to take the different components into account. The complete data set is split into 20 bins for each track type. The resulting fits for the different bins are shown in Figs. 6.11 to 6.13, while the parameters of interest, which are the scaling parameters of the decay-time resolution and the decay-time bias, are given in Fig. 6.14 with a linear and quadratic fits, respectively.

The scaling parameter can be described by a linear function, which is also seen on

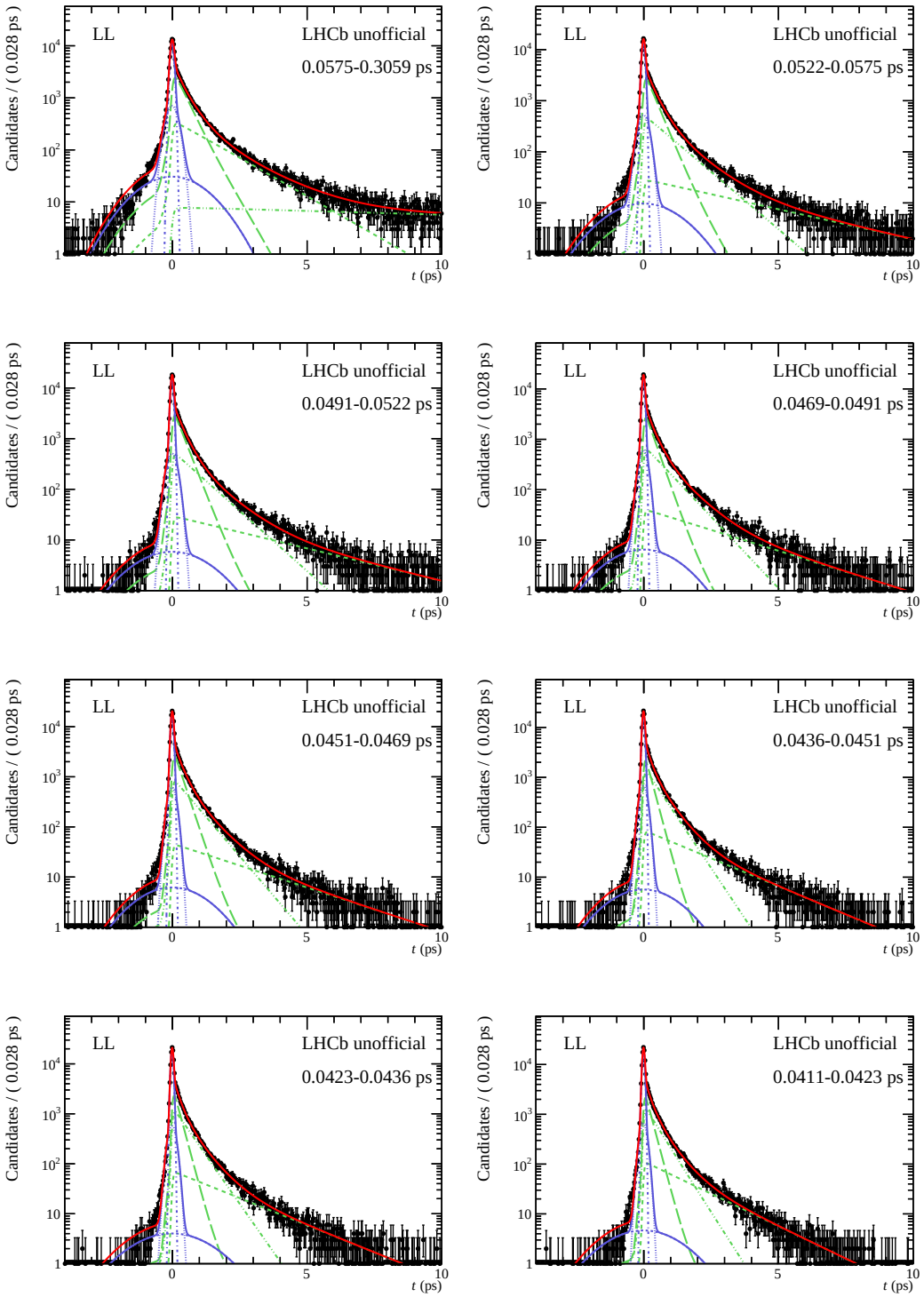

Figure 6.11: Fit to the mean decay-time resolution on prompt LL J/ψ data in different bins of the decay time error estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed like) with three exponential functions (green, dashed like) and the overall fit in red.

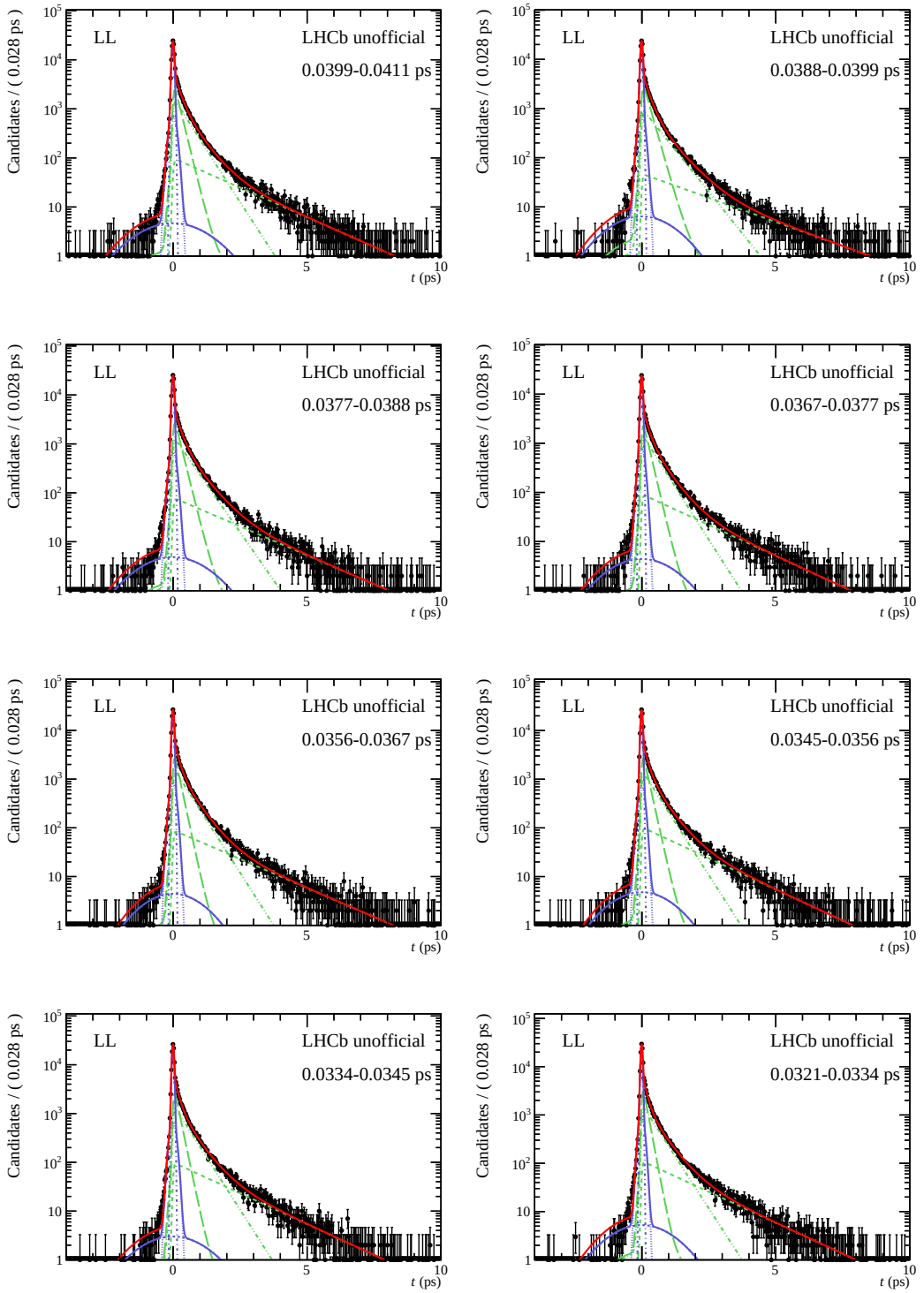

Figure 6.12: Fit to the mean decay-time resolution on prompt LL J/ψ data in different bins of the decay time error estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed like) with three exponential functions (green, dashed like) and the overall fit in red.

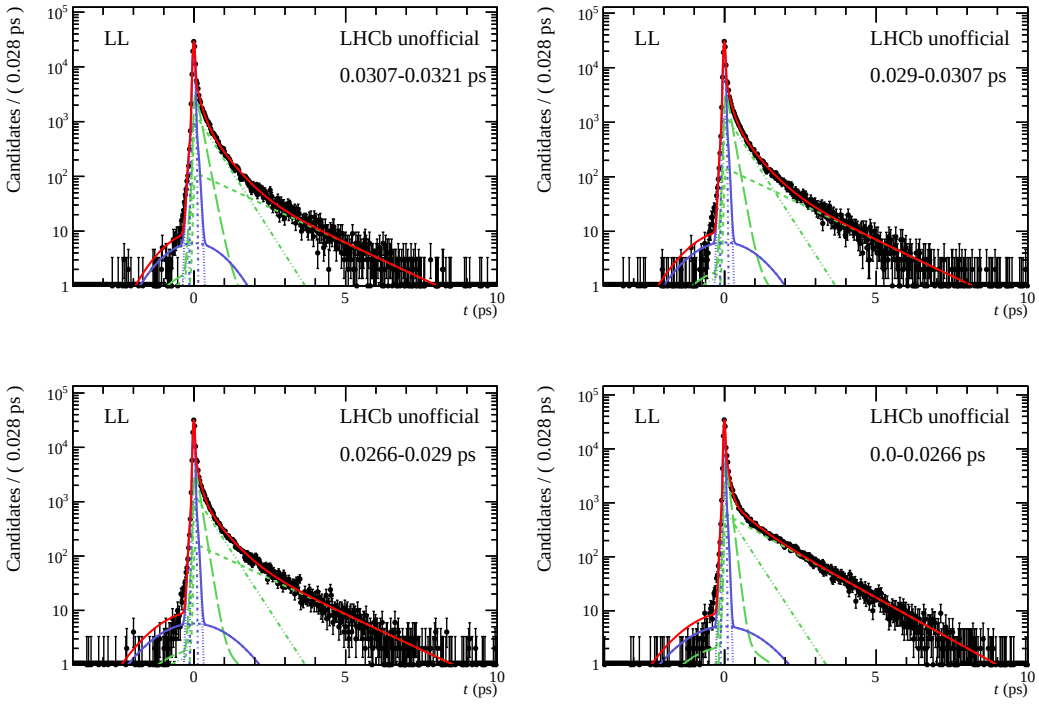

Figure 6.13: Fit to the mean decay-time resolution on prompt LL J/ψ data in different bins of the decay time error estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed like) with three exponential functions (green, dashed like) and the overall fit in red.

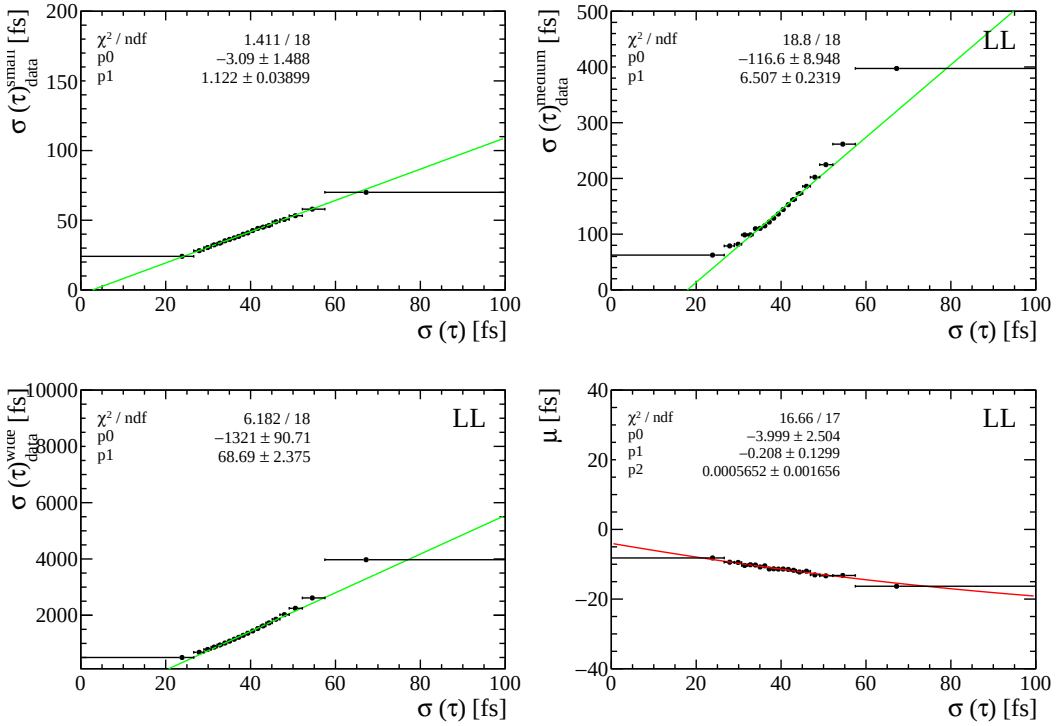

Figure 6.14: The width of the smallest (upper-left), the medium (upper-right), the widest Gaussian (lower-left), and the decay-time bias for different bins of the decay-time error estimate for LL candidates. With a linear or quadratic fit to the data points.

simulation, while the decay-time bias is described with a quadratic function. The extracted quadratic function of the decay-time bias is used to correct the decay time in the final *CP* fit on data, which is described in Sec. 6.6.

6.5 Flavour Tagging

A crucial ingredient in the measurement of time-dependent *CPV* in neutral B^0 mesons, is the knowledge of the production flavour of the B^0 meson. This is achieved using the flavour tagging [90]. In the p-p collisions at the LHC, b quarks are produced in b-$\bar{b}$ pairs. The different flavour tagging algorithms can be divided into two categories. One type of algorithm tries to reconstruct the decay of the non-signal b quark. If the non-signal b hadron decays in a flavour-specific channel, the flavour of the signal b and thus the flavour of the signal B^0 meson can be deduced. These types of algorithms are called opposite side algorithms (OS). The OS algorithms used in LHCb, are the OS Muon, OS Electron, OS Charm, OS Kaon and OS Vertex Charge. Most of these algorithms try to reconstruct the most common b decays, which are $b \to c\ell\overline{\nu}$ transitions. The OS Muon and Electron taggers try to select the resulting lepton, a muon or electron, respectively. From the charge of the lepton, the flavour of the b can be directly determined. The OS Charm tagger tries to reconstruct the c hadron coming from a $b \to cX$ transition, while the OS Kaon tagger goes one step in the decay chain further and tries to reconstruct a charged kaon, which possibly comes from the c hadron. In contrast, OS Vertex Charge does not reconstruct a particular physics process but sums up all charges in the possible b hadron decay vertex to give a tag decision.

The other type of tagging algorithm uses the hadronisation process of the signal B^0 meson. For a $\bar{b}$ to hadronise to a B^0 meson, a d quark is needed. This d quark is also produced in a pair with a $\overline{d}$ quark, which can hadronise with one or two $\overline{u}$ quarks to a negative charged pion or anti-proton, respectively. Therefore, the so-called same side (SS) tagging algorithms focus on reconstructing these pions or protons (SS Pion and SS Proton) and deduce the b quark flavour from their charge. For B^0_s mesons, the hadronisation process involves a s and therefore a charged kaon can be produced instead of a pion. Therefore, for B^0_s mesons the SS Kaon tagger is used. There is no analogue to the SS Proton tagger for B^0_s mesons since all dominant decays of strange baryons involve a Λ baryon, which has a very long lifetime. This long lifetime introduces large uncertainties by matching these baryons to the hadronsation process.

Since there is a large number of pions, protons and leptons in a typical LHCb collision, the flavour tagging is not perfect. Each tagging algorithm returns a tag

for the candidate B^0 meson, if possible, as well as a mistag estimate. This mistag estimate, η, is an output from a multivariate classifier and ranges from 0 (perfectly tagged) to 0.5 (untagged). To get a comparable quantity on the effect on a flavour-tagged measurement due to imperfect tagging, the so-called tagging power, $\epsilon_{\text{tag,eff}}$, is defined as

$$\epsilon_{\text{tag,eff}} = \epsilon \cdot (1 - \omega(\eta))^2. \tag{6.4}$$

Here, ϵ is the tagging efficiency of the tagger and $\omega(\eta)$ is the calibrated mistag probability. The tagging power can be interpreted as the efficiency, which reduces your data set due to the imperfect flavour tagging compared to a perfectly tagged data set. The tagging power depends mainly on the B^0 kinematics (p_{T}, pseudo-rapidity) and how many particles are in the total event. Therefore, the tagging power ranges from $3-8\%$ at LHCb [81, 91, 92] depending on the different decay channels and selections.

To verify and calibrate the mistag estimate, flavour specific B decays are used. These control channels need to be kinematicaly similar to the signal decay. For the OS taggers the decay $B^+ \to J/\psi K^+$ is used. $B^+ \to J/\psi K^+$ decays are experimentally well reconstructible and have a large branching fraction, resulting in a large number of signal candidates. Since the SS taggers analyse the hadronisation process of the B^0 meson, the calibration channel for the SS tagger needs to be a B^0 decay of the same flavour. For B^0_d mesons the used decay channel is $B^0 \to J/\psi K^*$. The decay can be called flavour specific, since the weak decay $K^* \to K^-\pi^+$ is negligible compared to the strong decay $K^* \to K^+\pi^-$. Since neutral B^0 mesons oscillate, the known oscillation frequency and the B^0 decay time has to be taken into account in the calibration. To decrease differences in kinematic distributions, for $J/\psi \to e^+e^-$ reconstructed $B^0 \to J/\psi K^0_{\text{S}}$ decays, the corresponding calibration channels are also reconstructed from $J/\psi \to e^+e^-$ decays. When calibrating the SS Kaon tagger for $B^0_s \to J/\psi K^0_{\text{S}}$ decays, the decay channel $B^0_s \to D^+_s\pi^-$ is used. Although the kinematic distributions and the selection differs, this calibration channel is the only one with sufficient number of signal candidates. After reweighting the kinematic distributions from $B^0_s \to D^+_s\pi^-$ decays to the one of $B^0 \to J/\psi K^0_{\text{S}}$ decays, the transferability is sufficient. This reweighting is done for all calibration channels.

First the calibration channel is selected using a BDT, that is trained for each channel analogous to the BDT of the signal channel. The cut on the BDT is optimized to ensure a very pure sample, which is assumed to be pure for the next steps. All calibration channels are reweighted to the signal distributions in the $p_{\text{T}}(B)$, $\eta(B)$, numbers of primary vertices and the number of tracks in the event. To mediate the curse of dimension [93], a multivariate reweighting algorithm [94] is used. After the reweighting, the mistag estimate is calibrated using a linear function. For

this the measured mistag is fitted in dependence of the estimate for each tagger. These individual calibrated taggers are then combined to a SS combination and OS combination, which are again calibrated. The combination of both tagging sides is done within the final *CP* fit. All of these steps are described in Sec. 6.5.1 for the SS taggers, while the OS tagger calibration is discussed in Sec. 6.5.2.

6.5.1 Same Side Tagger Calibration

The $B^0 \to J/\psi K^*$ candidates for the SS tagger calibration are selected using a cut-based selection followed by an BDT, that is similar to the signal BDT but retrained. The BDT is retrained using $B^0 \to J/\psi K^*$ simulation and upper-mass sideband data. The cut on the BDT is determined for the reconstructed data set to result in a signal purity of 99.9% (95%) for muon (electron) reconstructed $B^0 \to J/\psi K^*$ decays in the smallest window, that contains 95% of the signal, evaluated on the corresponding simulation. In Fig. 6.15 the mass distribution with a fit to a larger range with the narrow range indicated is shown for both decay modes.

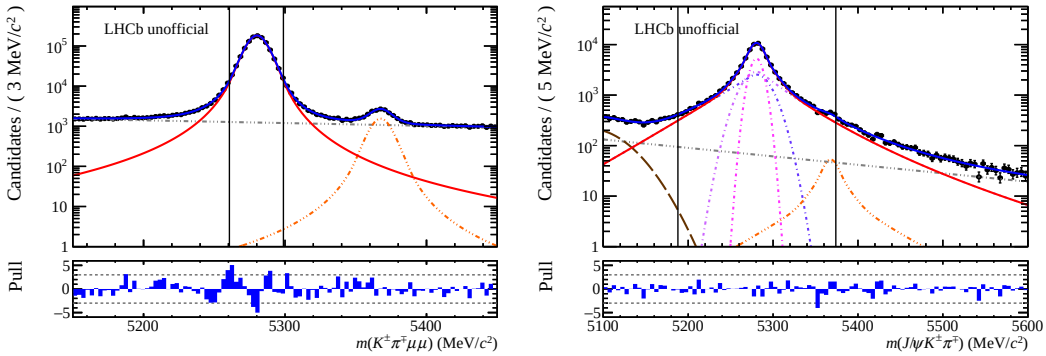

Figure 6.15: Mass fits of the SS calibration channels after full selection for (left) the muon and (right) the electron channel with the full model (blue), the $B^0 \to J/\psi K^*$ signal (red, solid), and $B^0_s \to J/\psi K^*$ background (red, dashed). The smaller range that is used in the following for the calibration is indicated with vertical lines.

From this the number of signal candidates is determined to be around 1 100 000 (85 000) in the narrow window for $J/\psi \to \mu^+\mu^-$ ($J/\psi \to e^+e^-$) reconstructed candidates. These candidates are reweighted to the signal kinematic and event multiplicity, separate for $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S)K^0_S$ decays. With these candidates, the SS Proton and SS Pion tagger are calibrated using the Espresso Performance Monitor (EPM) [95]. After calibration, both taggers are combined and the combination is

also calibrated. For the calibration, the control channel data is split and one half is used for the calibration of the single tagger and the other half is used for the combination. The resulting calibration plots for both steps are given in Fig. 6.16 for $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates, while the calibration for the other two decay channels are shown in Figs. A.29 and A.30.

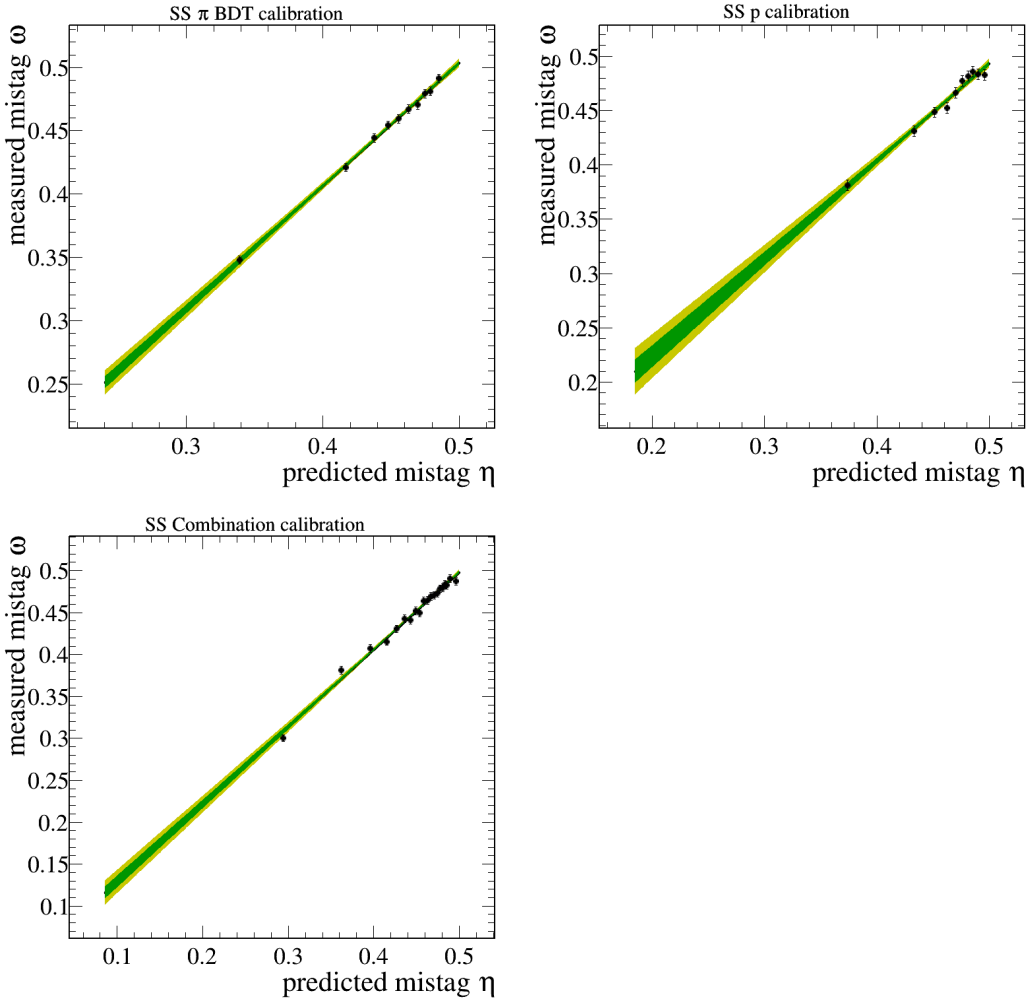

Figure 6.16: Flavour tagging calibration on $B^0 \to J/\psi K^*$ candidates of the SS single tagger (top-left) Pion and (top-right) Proton and (bottom) their combination for $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates.

The SS Kaon tagger for B^0_s mesons is calibrated using a selected $B^0_s \to D^+_s\pi^-$ data sample, which is selected and has sweights [68, 69] calculated by an other analysis

group [96]. In this decay sweights are needed, since the necessary purity can not be achieved due to remaining resonant backgrounds. The data set consists of about 170 000 $B^0_s \to D^+_s \pi^-$ signal candidates. These candidates are also reweighted to the $B^0_s \to J/\psi K^0_S$ signal. In Fig. 6.17 the calibration of the SS Kaon tagger is shown, while the resulting tagging power is given in Table 6.4.

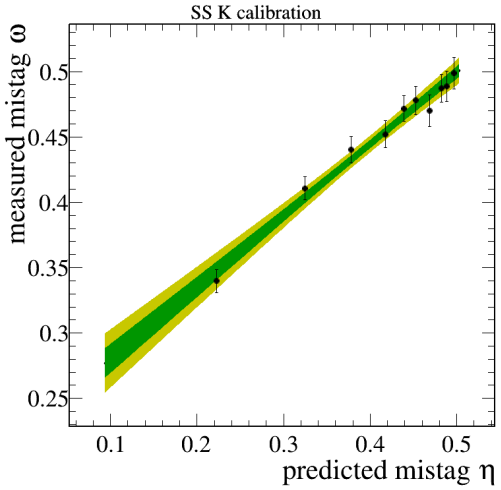

Figure 6.17: Flavour tagging calibration on $B^0_s \to D^+_s \pi^-$ candidates of the SS Kaon tagger for $B^0_s \to J/\psi(\to \mu\mu)K^0_S$ candidates.

Larger differences between the different decay channels and reconstructions can be seen. The $J/\psi \to e^+e^-$ reconstructed decay data set has a larger same side tagging power. This is due to the correlation of the B^0 kinematic with the tagging power, B^0 with a higher momentum can be better tagged and the selection requirements in the electron sample are tighter. This effect is larger than the quoted numbers. There is an irreversible bug in the central $B^0 \to J/\psi(e^+e^-)K^0_S$ preselection for 2018 data, which lead to no SS tagging information present for this sample. Thus, the SS tagging power in the 2015-2017 $B^0 \to J/\psi(e^+e^-)K^0_S$ data sets are around 2%.

6.5.2 Opposite Side Tagger Calibration

The $B^+ \to J/\psi K^+$ candidates are selected in a similar way than the $B^0 \to J/\psi K^*$ candidates. For $J/\psi \to e^+e^-$, a BDT cut is chosen with 95% purity in the smallest mass window that contain 95% of the signal evaluated on a corresponding simulation. For $J/\psi \to \mu^+\mu^-$, a purity of 99% is chosen. In Fig. 6.18 the mass distribution with a fit to a larger range is shown for both decay modes. After reweighting, the OS taggers are calibrated using the EPM. The calibrated single taggers are combined

Table 6.4: Tagging power in % for each single SS tagger and the SS tagger combinations after the FT calibration determined on sweighted signal data. The quoted uncertainty only contains the statistical uncertainty from the size of the data set and not uncertainties from the calibration.

Tagger	$B^0 \to J/\psi(\mu^+\mu^-)K^0_S$	$B^0 \to \psi(2S)K^0_S$	$B^0 \to J/\psi(e^+e^-)K^0_S$	$B^0_s \to J/\psi K^0_S$
SS Pion	1.012 ± 0.004	1.043 ± 0.016	0.974 ± 0.013	-
SS Proton	0.3899 ± 0.0026	0.402 ± 0.009	0.521 ± 0.010	-
SS Kaon	-	-	-	0.946 ± 0.006
SS Combination	1.267 ± 0.005	1.321 ± 0.018	1.254 ± 0.015	0.946 ± 0.006

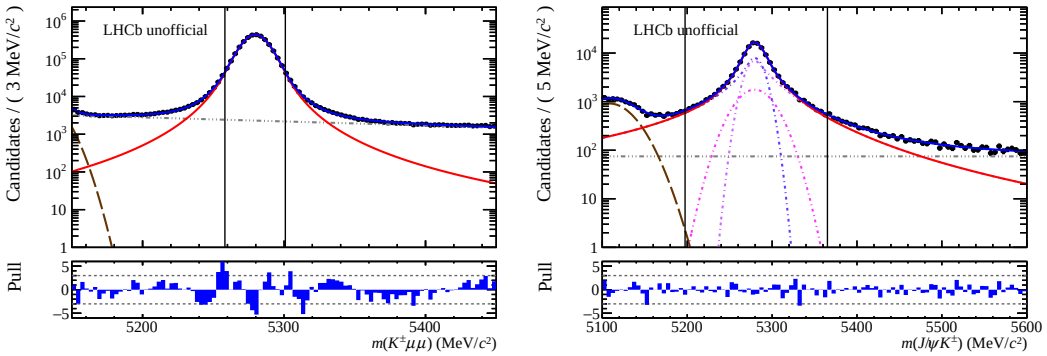

Figure 6.18: Mass fits of the OS calibration channels after full selection for (left) the muon and (right) the electron channel with the full model (blue), the $B^+ \to J/\psi K^+$ signal (red). The smaller range that is used in the following for the calibration is indicated with vertical lines.

and the combination is calibrated. The resulting OS calibration plots are shown in Fig. 6.19 for $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates, while the corresponding plots for the other decay channels are given in Figs. A.31 to A.33.

All tagging powers of the OS taggers for the different data sets are summarised in Table 6.5. As on the same side, larger differences between the electron and the muon modes are measured. Here, this is partially due to the different momentum spectra but mainly due to a different trigger requirements. Since the muons are easily reconstructible in the trigger, the candidates are triggered due to the particles of the signal decay. In contrast, electrons cannot efficiently triggered. Therefore, electron candidates that are trigged due to non-signal particles are also used. These non-signal particles are most likely from other b decays, which are well reconstructed and can be well used in the opposite side tagging.

Table 6.5: Tagging power in % for each single OS tagger and the OS tagger combinations after the FT calibration determined on sweighted signal data. The quoted uncertainty only contains the statistical uncertainty from the size of the data set and not uncertainties from the calibration.

Tagger	$B^0 \to J/\psi(\mu^+\mu^-)K^0_S$	$B^0 \to \psi(2S)K^0_S$	$B^0 \to J/\psi(e^+e^-)K^0_S$	$B^0_s \to J/\psi K^0_S$
OS Charm	0.3696 ± 0.0035	0.360 ± 0.012	0.413 ± 0.011	0.380 ± 0.004
OS Electron	0.333 ± 0.004	0.310 ± 0.015	0.385 ± 0.012	0.332 ± 0.005
OS Kaon	1.034 ± 0.006	1.017 ± 0.020	1.344 ± 0.019	1.049 ± 0.007
OS Muon	0.846 ± 0.008	0.805 ± 0.026	1.130 ± 0.025	0.837 ± 0.009
VtxCharge	1.046 ± 0.006	1.041 ± 0.020	1.386 ± 0.019	1.071 ± 0.007
OS Combination	2.685 ± 0.012	2.63 ± 0.04	3.37 ± 0.04	2.712 ± 0.014

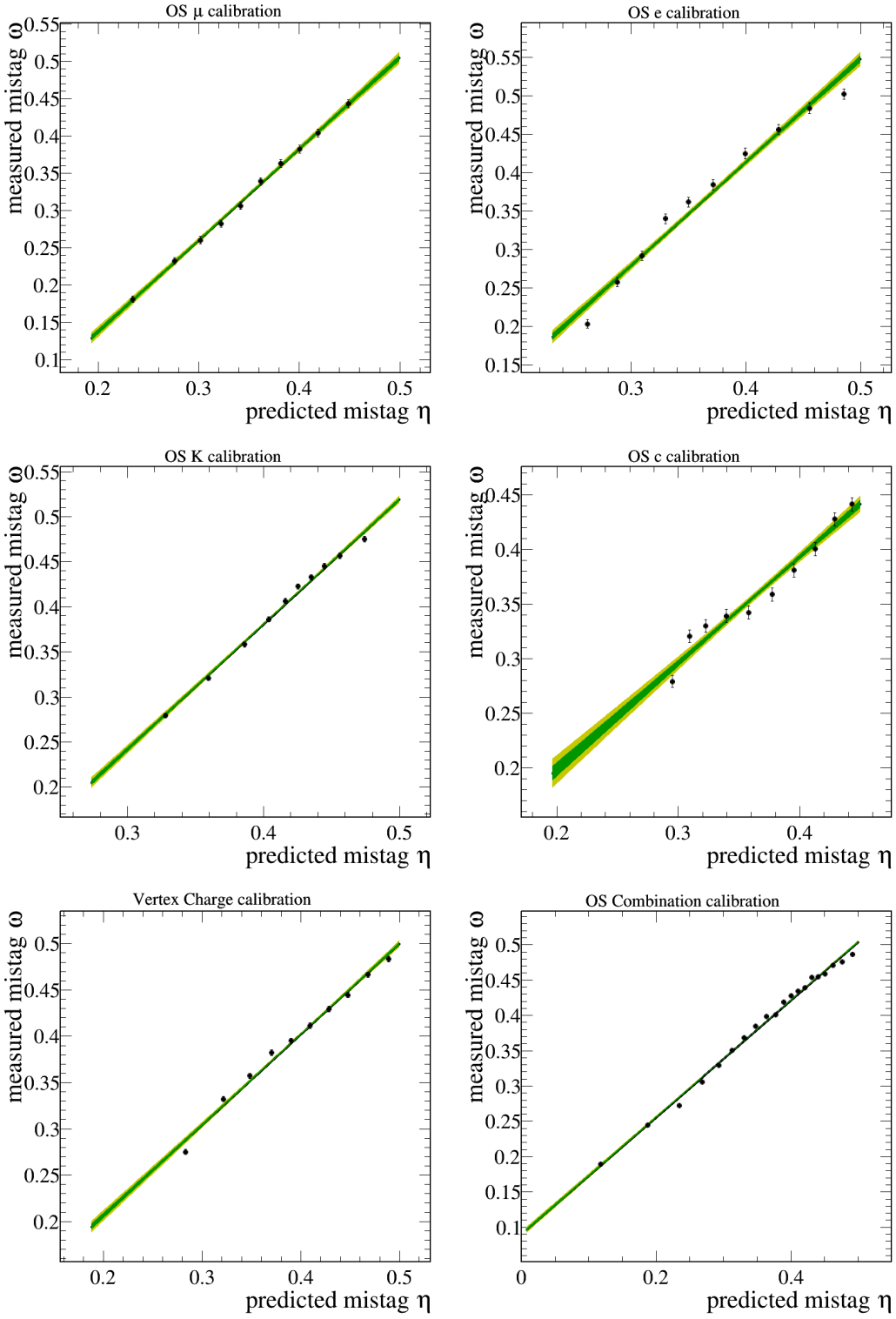

Figure 6.19: Flavour tagging calibration for $B^0 \to J/\psi(\to \mu\mu)K^0_S$ of the OS single tagger (upper-left) Muon, (upper-right) Electron, (middle-left) Kaon, (middle-right) Charm and (lower-left) vertex charge and (lower-right) the OS combination using $B^+ \to J/\psi\, K^+$ candidates.

6.6 Time-Dependent *CP* Asymmetry Fit

In the final fit to extract the *CP* parameter S, C and $\mathcal{A}_{\Delta\Gamma}$ (only for $B^0_s \to J/\psi K^0_S$) is described in the following. Starting with a discussion of the PDF that is fitted to the data in Sec. 6.6.1. To validate the fit procedure, fits to signal simulation are performed with different degrees of complexity, which is documented in Sec. 6.6.2. Different production rates of B^0 and $\overline{B}^0$ are possible, which leads to a production asymmetry, which affects the measurement of the *CP* parameters. This production asymmetry is determined using a control channel, which is discussed in Sec. 6.6.3. Finally, blinded fits to the recorded data sets are done that shows the sensitivity on the parameters of interest. These fits and the resulting sensitivities are shown in Sec. 6.6.4.

6.6.1 The *CP* Asymmetry Probability Density Function

The *CP* asymmetry is fitted time-dependent, the corresponding theoretical PDF, $\mathcal{P}(t, B^0, \overline{B}^0)$, can be derived from Eq. (2.26) as

$$\begin{aligned}\mathcal{P}(t, B^0, \overline{B}^0) &= \mathcal{P}(t, B^0) + \mathcal{P}(t, \overline{B}^0) \\ &= \exp^{-\Gamma t}\left[\mathcal{A}_{\Delta\Gamma}\sinh\left(\frac{1}{2}\Delta\Gamma t\right) + \cosh\left(\frac{1}{2}\Delta\Gamma t\right) + C\cos(\Delta m t) - S\sin(\Delta m t)\right] \\ &+ \exp^{-\Gamma t}\left[\mathcal{A}_{\Delta\Gamma}\sinh\left(\frac{1}{2}\Delta\Gamma t\right) + \cosh\left(\frac{1}{2}\Delta\Gamma t\right) - C\cos(\Delta m t) + S\sin(\Delta m t)\right],\end{aligned} \tag{6.5}$$

where Γ is the decay width, $\Delta\Gamma$ is the decay width difference between the heavy and light B^0 mass eigenstate and $\mathcal{A}_{\Delta\Gamma}$, S, and C are the *CP* observables, that are measured in the fit.

These theoretical PDFs are modified due to experimental effects. Time-independent asymmetries arise from detection, $\mathcal{A}_{\text{det}}$, and B^0 production $\mathcal{A}_{\text{prod}}$ [97, 98]. Since the analysed decay channels are charge symmetric in the final state, no detection asymmetry is present in this analysis. A non-zero production asymmetry is expected in LHC collisions since two protons are collided and not a proton and an anti-proton. The production asymmetry leads to offsets in both PDFs. Another experimental effect is the decay-time resolution and acceptance (discussed in Sec. 6.4), which leads the effect that the measured decay time is not the true decay time. This is incorporated in the PDF by a convolution with the decay resolution function, $\mathcal{R}(t'-t)$, while the acceptance, $\epsilon(t)$, adds a decay time dependent factor to the PDF. The last needed modification to the theoretical PDF is due to the imperfect flavour

tagging. As the definition of the mis-tag ω implies, the probability for one tagger to set the wrong tag decision, d', given the true tag, d, is ω. Hence, for $d' = d$ the probability is $1-\omega$. Here, $d = 1$ corresponds to a B^0, while $d = -1$ is a $\overline{B}^0$ meson. Further, a tag decision can be made with the tagging efficiency, ϵ_{tag}. Therefore, the likelihood for a tag decision d' given a true tag d given for one tagger is given as

$$\mathcal{P}(d'|d) = \epsilon_{\text{tag}}\delta_{d',1}\left(\frac{1 + d \cdot d'(1-2\omega)}{2}\right) + \delta_{d',0}(1-\epsilon_{\text{tag}}). \tag{6.6}$$

To simplify the following likelihood, the sum and the difference of both tag decisions is denoted as

$$\Sigma_{\text{tag}}(d') = \mathcal{P}(d'|B^0) + \mathcal{P}(d'|\overline{B}^0) \tag{6.7}$$

$$\Delta_{\text{tag}}(d') = \mathcal{P}(d'|B^0) - \mathcal{P}(d'|\overline{B}^0). \tag{6.8}$$

With this modification the complete likelihood function can be written as

$$\begin{aligned}\mathcal{P}(t', d') = \Bigg(\epsilon(t')\exp^{-\Gamma t'}\Bigg[\left(\Sigma_{\text{tag}} - \Delta_{\text{tag}}\mathcal{A}_{\text{prod}}\right)\left(\mathcal{A}_{\Delta\Gamma}\sinh\left(\frac{1}{2}\Delta\Gamma t'\right) + \cosh\left(\frac{1}{2}\Delta\Gamma t'\right)\right) \\ +(\Delta_{\text{tag}} - \Sigma_{\text{tag}}\mathcal{A}_{\text{prod}})(C\cos(\Delta m t') - S\sin(\Delta m t'))\Bigg]\Bigg) * \mathcal{R}(t'-t).\end{aligned} \tag{6.9}$$

The tag decision is a function of the mis-tag, which is a function of the mis-tag estimate η. Also the resolution depends on the decay-time uncertainty estimate σ. Therefore, the fit is done conditionally for each candidate taking the candidate's values of η and σ into account.

To include the uncertainties of input parameters in the fit, constraints are used. The input parameter is left floating in the fit but a Gaussian function of that parameter is multiplied to the likelihood. The mean and width of this Gaussian corresponds to the expected value and the uncertainty of the parameter. For correlated parameters, *e.g.* the slope and intercept of the linear flavour tagging calibration, multidimensional Gaussians with the corresponding correlations are used. A constrain lead to possibility that an input parameter differs from the expected value, if this helps to describe the data but a penalty is assigned to the likelihood through the Gaussian function. The constrained parameters in the baseline fit are the B^0 lifetime, Δm, $\Delta\Gamma_s$ (only for $B_s^0 \to J/\psi K_{\text{S}}^0$), and the tagging calibration parameters. The values and width are taken from Ref. [16] and the flavour tagging calibration, respectively.

The fit model is simplified for B^0 because $\Delta\Gamma_d$ is measured to be very small ($\Delta\Gamma_d/\Gamma_d = 0.001 \pm 0.01$ [16]) and therefore set to zero in the fit. This leads to

$\sinh\left(\frac{1}{2}\Delta\Gamma t'\right) = 0$ and $\cosh\left(\frac{1}{2}\Delta\Gamma t'\right) = 1$, therefore $\mathcal{A}_{\Delta\Gamma}$ cannot be measured in B^0 decays. In $B^0_s \to J/\psi K^0_S$ decays $\mathcal{A}_{\Delta\Gamma}$ is a parameter of interest and is fitted. Besides the usage of signal weights in the fit data, there is no difference between the fit to simulation (Sec. 6.6.2) and data (Sec. 6.6.4).

6.6.2 *CP* Fit on Simulation

To validate the fit procedure, the first step is to fit signal simulation. The generated values for the *CP* observables are known and listed in Table 6.6. Because no

Table 6.6: Generated values of the *CP* observables for $B^0 \to J/\psi K^0_S$, $B^0 \to \psi(2S)K^0_S$ and $B^0_s \to J/\psi K^0_S$ signal in the simulation.

Parameter	$B^0 \to J/\psi K^0_S$	$B^0 \to \psi(2S)K^0_S$	$B^0_s \to J/\psi K^0_S$
S	0.6997	0.6997	-0.3306
C	0.0	0.0	-0.0118
$\mathcal{A}_{\Delta\Gamma}$	-	-	0.9437

background is present in the simulated data sets, no sweights are needed. Another advantage is that the true production flavour is known, hence at a first check a fit is done using this truth information. This fit setup allows to check all other parameters entering the fit but the flavour tagging with a large statistical precision. Technically, the truth information is used as tag decision, while the mis-tag probability is set to zero. The decay-time resolution and acceptance is used as in the nominal fit. The results of the fits separate for each year of data taking and decay modes are listed in Figs. 6.20 and 6.21. In the B^0 decays most results are compatible with their generated values within the statistical uncertainties, which are much lower than the expected uncertainties on data, because no dilution from the flavour tagging is present. Smaller deviations are present in smaller simulated data sets *e.g.* $B^0 \to J/\psi(e^+e^-)K^0_S$ longtrack. In the $B^0_s \to J/\psi K^0_S$ simulation, S and C are compatible with the generated values, while a large deviation for $\mathcal{A}_{\Delta\Gamma}$ in long track decays is seen, which needs further investigation.

The next step is to use flavour tagging in the fit. Since the simulation does not describe the data perfectly, different calibrations are needed, especially because the underlying event is heavily used in the flavour tagging and this is known to be imperfect simulated. The calibration function is extracted in the very same way on signal simulation as on data, including the kinematic reweighting of the control channels simulation ($B^0 \to J/\psi K^*$ and $B^+ \to J/\psi K^+$). All fit results of the different *CP* parameters for the different years and decay channel are shown in Fig. 6.22.

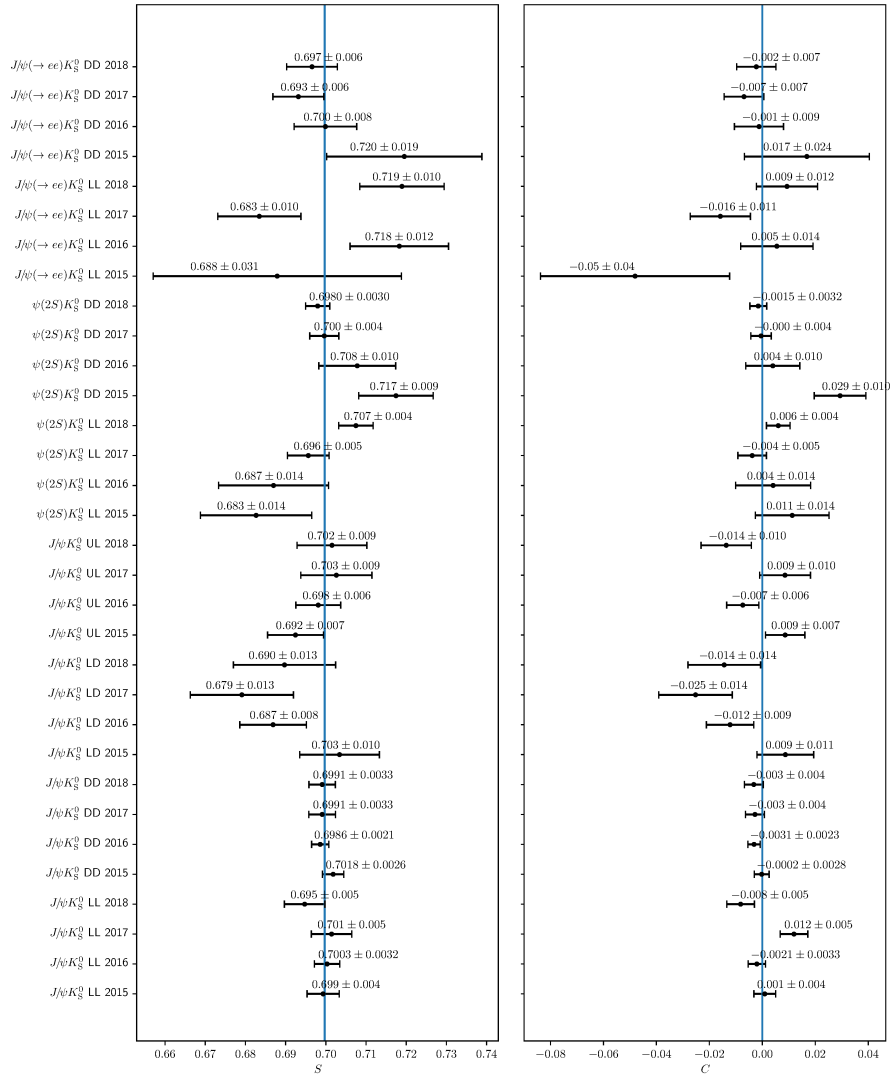

Figure 6.20: *CP* fit result of (left) S and (right) C of $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$, $B^0 \to J/\psi(e^+e^-)K^0_S$, $B^0 \to \psi(2S)K^0_S$ simulation using truth information instead of flavour tagging separate for each track type and year of data taking. The generated values are indicated by the vertical line.

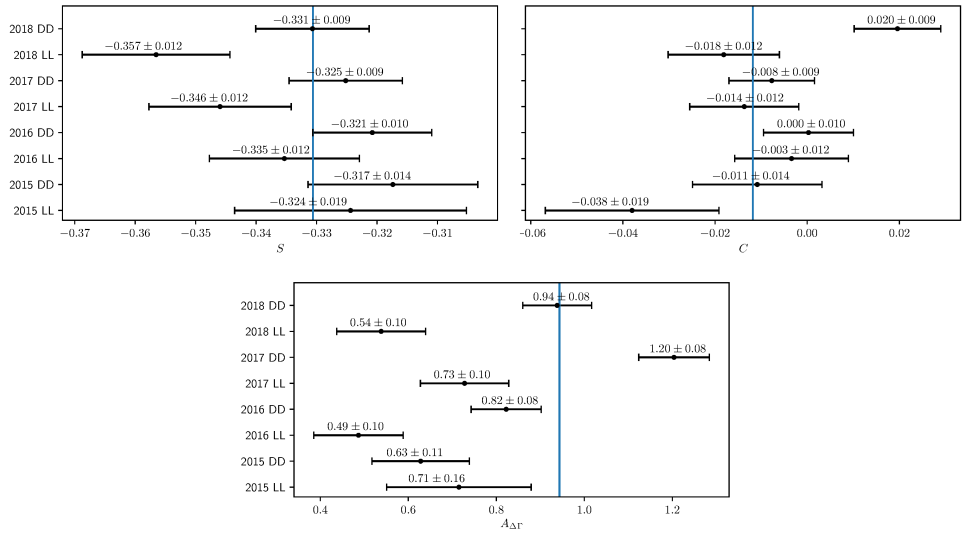

Figure 6.21: *CP* fit result of (top-left) S, (top-right) C, and (bottom) $\mathcal{A}_{\Delta\Gamma}$ of $B^0_s \to J/\psi K^0_S$ simulation using truth information instead of flavour tagging. The generated values are indicated by the vertical line.

The results are very similar to the ones using the true-ID calibration but with much larger statistical uncertainties due to the imperfect tagging. However, it confirms the given portability of the flavour tagging calibration after reweighting from the control channels to the signal channels.

6.6.3 Extracting the Production Asymmetry from $B^0 \to J/\psi K^*$ Decays

The production asymmetry, A_{prod}, can in principle be left floating in the final *CP* fit. However, since it is correlated to the *CP* parameters, this increases the statistical uncertainties on them. This increase is about 10% relative for the B^0 measurements. As the production asymmetry is independent of the final state, flavour specific decays, where the *CP* parameters are known can be used to extract it. For the B^0 mesons, the very same data set, that is used for the SS flavour tagging calibration, $B^0 \to J/\psi K^*$ decays, is used. Here, the final state is not reachable by both $B/\overline{B}$, therefore an additional set of *CP* parameters is introduced ($S_{\bar{f}}, C_{\bar{f}}$), which can be analogous defined as in Sec. 2.3.1. As the final state f ($\bar{f}$) can only be reached by B ($\overline{B}$), C ($C_{\bar{f}}$) is 1 (-1) while S ($S_{\bar{f}}$) is 0. With fixing these values in a *CP* fit, A_{prod} can be extracted. As the $B^0 \to J/\psi K^*$ final state is not charge symmetric, a detector asymmetry can be present, therefore A_{det} is left floating in the fit. With

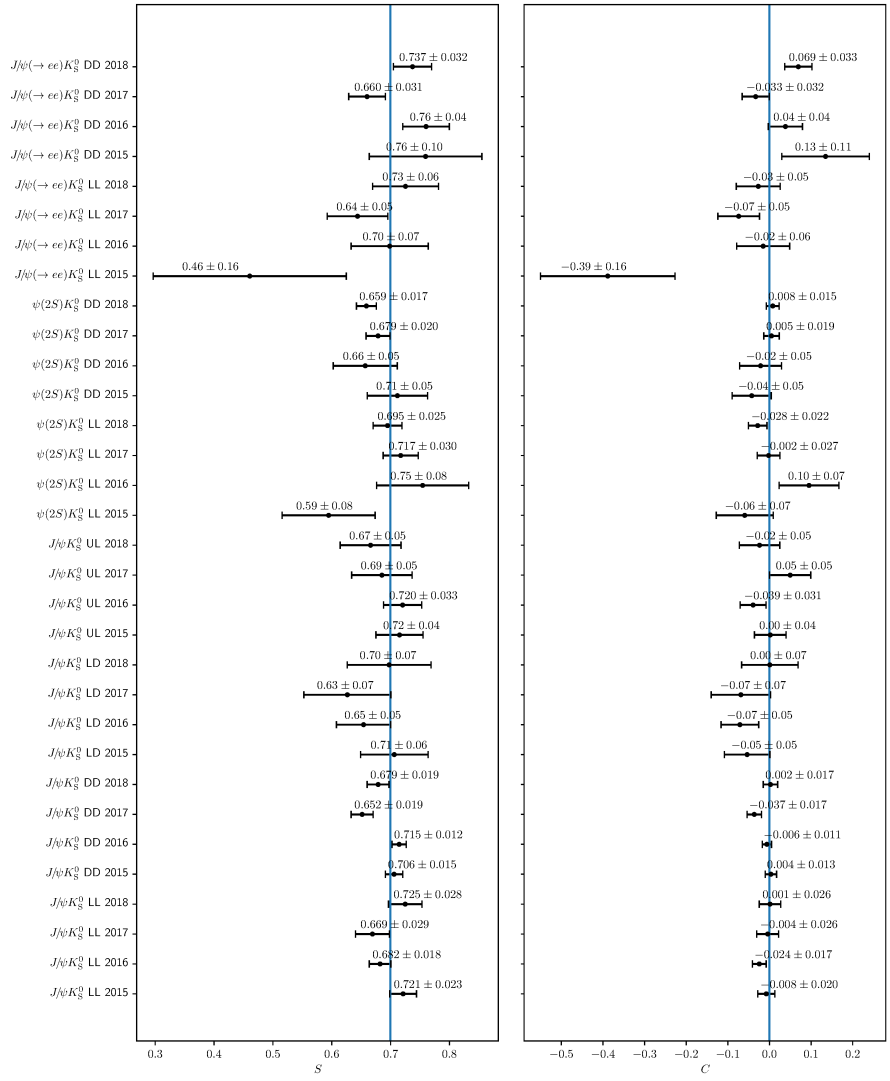

Figure 6.22: *CP* fit result of (left) S and (right) C of $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$, $B^0 \to J/\psi(e^+e^-)K^0_S$, $B^0 \to \psi(2S)K^0_S$ simulation separate for each track type and year of data taking. The generated values are indicated by the vertical line.

this fit A_{prod} is determined to be

$$A_{\text{prod}} = -0.0038 \pm 0.0018. \tag{6.10}$$

The corresponding decay time projection of this fit is given in Fig. 6.23. This value for A_{prod} is constrained in the *CP* fit in B^0_d decays.

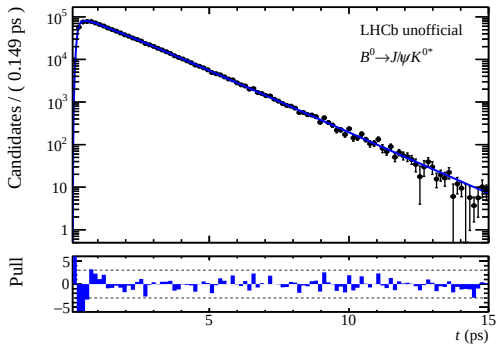

Figure 6.23: Decay time projection of the *CP* fit result $B^0 \to J/\psi\, K^*$ with fixed *CP* parameters to extract A_{prod}.

6.6.4 Blinded *CP* Fit on Data

In the fit to data, sweights are used. The coverage is tested using pseudo-experiments (Sec. 6.7.2), where only a slight over-coverage is found. At the moment, the analysis is blinded. The blinding is done by adding reproducible random numbers to the different *CP* parameters, while using the unblinded parameters internally. The blinded parameters are S and C and $\mathcal{A}_{\Delta\Gamma}$ in the B^0_s fit. Different shifts are used for S and C and for the B^0 and B^0_s decays, to avoid indirect unblinding. Therefore the following quoted central values of these fits are of no interest but the statistical uncertainty is the very same as without the blinding. Nevertheless, using the same random number between the different B^0 decays, allows to compare the compatibility of the blinded results between each decay channel. The blinded fit results for S and C are shown in Fig. 6.24 for the B^0 decay modes and in Fig. 6.25 for $B^0_s \to J/\psi K^0_S$ decays, while the decay time fit projections are shown in Figs. A.34 to A.36. The results are reasonably compatible between the difference decay modes and track types for the B^0 decays, while a large tension is present between long and downstream track reconstructed $B^0_s \to J/\psi K^0_S$ decays. All B^0 decays are fit simultaneously sharing the *CP* parameters, as well as the properties of the B^0 system

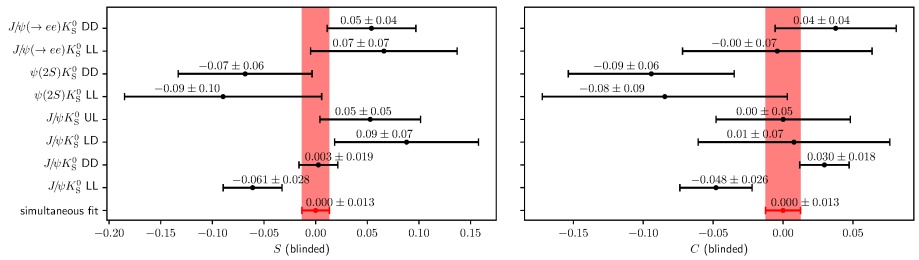

Figure 6.24: Blinded *CP* fit result of (left) S and (right) C split by each decay channels and track types and a simultaneous fit. The central values are shifted, that the simultaneous fit value is zero.

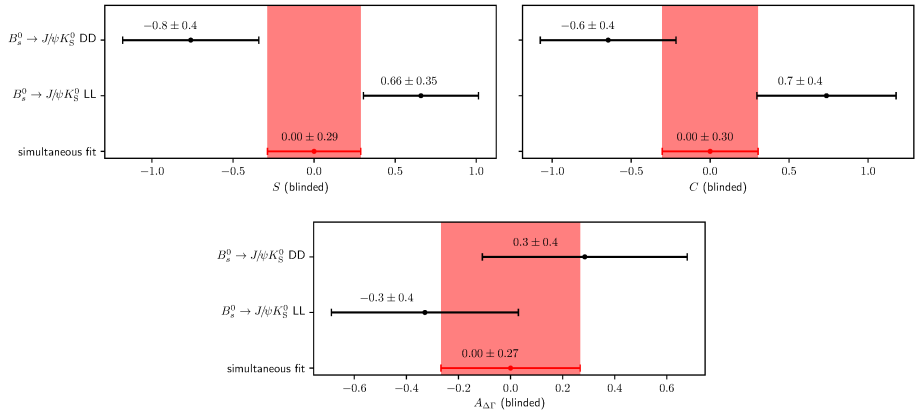

Figure 6.25: Blinded *CP* fit result of (top-left) S, (top-right) C, and (bottom) $\mathcal{A}_{\Delta\Gamma}$ for $B^0_s \to J/\psi K^0_S$ split by track types and a simultaneous fit. The central values are shifted, that the simultaneous fit value is zero.

(Δm, $\Delta\Gamma$, τ). The resulting combined statistical sensitivities are

$$\sigma(S)_{B^0\to\psi K^0_S} = 0.013,$$
$$\sigma(C)_{B^0\to\psi K^0_S} = 0.013.$$

Although the tension between long and downstream track reconstructed $B^0_s \to J/\psi K^0_S$ candidates, a simultaneous fit is done for completeness. This results in a sensitivity of

$$\sigma(S)_{B^0_s\to J/\psi K^0_S} = 0.29,$$
$$\sigma(C)_{B^0_s\to J/\psi K^0_S} = 0.30,$$
$$\sigma(\mathcal{A}_{\Delta\Gamma})_{B^0_s\to J/\psi K^0_S} = 0.27.$$

6.7 Systematic Uncertainties

The systematic uncertainties are mostly evaluated using pseudo-experiments. In pseudo-experiments pseudo-data sets are generated using PDFs extracted from data. These data sets are then fitted in the very same way as the nominal data sets. Then the resulting values for each parameter can be compared with the generated values. Since generating a pseudo-data set is much faster than full simulation, some thousand of these pseudo-experiments can be done to validate the method. Also the effects of certain parameters can be checked, *e.g.* using a different value for an input parameter in the generation as in the fit. A measure for potential bias in measurement arising from certain parameter choices is the mean and width of the resulting pull distribution of these pseudo-experiments. The pull, p, is defined as

$$p = \frac{(x_{\text{fit}} - x_{\text{gen}})}{\sigma(x_{\text{fit}})}, \tag{6.11}$$

where x_{fit} is the value of the fitted parameter on the pseudo-data set, $\sigma(x_{\text{fit}})$ the corresponding fit uncertainty, and x_{gen} is the generated value. Thus, a perfect pull distribution should have a mean compatible with zero otherwise the fit is biased. Further, the width of a good fit should be one, which is the definition, that the fit uncertainty is correctly estimated. An other measure is the residual, which is simply the difference between the generated and fitted values. The residual gives a measure of a possible absolute bias. Besides using pseudo-experiments to determine the effects from different input parameters, corresponding checks are done using the baseline data set. These checks do not show any deviation from the baseline results, that is not covered by the systematic uncertainty evaluated using pseudo-experiments. The

pseudo-experiment setup is discussed in Sec. 6.7.1, while the different scenarios are discussed in Secs. 6.7.2 to 6.7.6 and the assigned systematic uncertainties are listed in Table 6.7.

Table 6.7: Summary of the systematic uncertainties in % of the statical uncertainty. The values are evaluated for $B^0 \to J/\psi(\to \mu\mu)K^0_S$ decays. No larger differences are expected for the other B^0 decays, where the statistical uncertainties are much larger.

	rel. $\sigma(S)$ [%]	abs. $\sigma(S)$ [%]	rel. $\sigma(C)$ [%]	abs. $\sigma(C)$ [%]
Fit validation	1.9 ± 1.2	$0.000\,37 \pm 0.000\,23$	0.0 ± 1.3	0.0000 ± 0.0002
Common FT Calibration		negligible		
Production asymmetry	21.2 ± 2.1	$0.003\,42 \pm 0.000\,33$	28.6 ± 2.2	$0.003\,86 \pm 0.000\,29$
Neglecting $\Delta\Gamma_d$	29.0 ± 3.1	0.0053 ± 0.0006	11.1 ± 3.0	0.0017 ± 0.0005
Decay-time bias correction		negligible		
total	36 ± 4	0.0063 ± 0.0007	31 ± 4	0.0042 ± 0.0006

6.7.1 Pseudo-Experiment Setup

The pseudo-experiments are done by generating data sets according to PDFs, which are extracted using sweighted data. These generated data sets are fitted afterwards with the baseline fit model. The generated data sets also include the known background components to validate the sweight extraction and the use of sweights in the *CP* fit. Thus, the data set consists of four different classes: B^0 signal, B^0_s signal, combinatorial background and partially reconstructed background. Individual shapes are used for the different track types but no split in years of data taking is used, if not explicitly mentioned.

The first variable, which need to be generated is the reconstructed B^0 mass. The model used to generate the pseudo-experiments is the same as in the fits to data to extract the sWeights. The generated mass distribution of all signal and background pseudo candidates is fitted for each track type with the same PDF and sWeights are extracted using the same work flow as in the baseline analysis.

Since the final *CP* fit is conditional of the decay-time uncertainty, $\sigma(t)$, and the flavour tagging mistag estimates for SS and OS, $\eta_{\mathrm{OS,SS}}$, these values have to be generated before generating the tag decision and the decay time (compare Eq. (6.9)). Each of these variables are generated from a histogram PDF extracted from sweighted data. For the B^0 and B^0_s component as well as for the partial background component the baseline B^0 sWeights are used to construct these histograms, while for the combinatorial background sweights for the combinatorial background are used.

The last needed set of variables contains the decay time and the tag decision. These are generated using the *CP* PDF (Eq. (6.9)) using the generated per-candidate $\sigma(t)$ and $\eta_{\text{OS,SS}}$ values. For this PDF the flavour tagging calibration is used for each mass component as well as the decay time resolution calibration from the signal simulation. The lifetime is set to 1.52 ps for the B^0 and the partial component, while the B^0_s lifetime is set to 1.509 ps. For the Combinatorial component, the value is set to 0.8, which is roughly the value fitted on sweighted data for this component. The *CP* parameter S is set to 0.7 for the B^0 as well as for the partial component, while it is set to 0 for the B^0_s and combinatorial component. The *CP* parameter C is set to 0 for all components, as no direct *CPV* is expected. At this point all tests are done for the main decay channel ($B^0 \to J/\psi(\to \mu\mu)K^0_S$), with basic checks for the other analysed decay channels. No different behaviour is expected for the electron or $\psi(2S)$ channel, since the very same methods are used. Thus, similar absolute systematic uncertainties are expected in the other B^0 decays but relative to the statistic uncertainties, the effects are much smaller.

6.7.2 Validation of Pseudo-Experiment Setup

The first step using pseudo-experiments is to check that unbiased results are obtained when using the same PDF and parameter values in the generation as in the fit. This check also verifies the used sweight procedure to project out the background components in the fit. After 6000 pseudo-experiments a bias on S below 2 % of the statistical uncertainty is seen, with a significance of 1.5σ, which is considered to be negligible. Further, a small overestimation of the uncertainty is seen. Thus, the used weight correction does not perfectly correct the uncertainty obtained by the fit. Since the effect is small and an overestimation, the uncertainty is not further corrected to be conservative. The pull and residual distributions for S and C are given in Fig. 6.26 with a Gaussian fit.

6.7.3 Effect of Common FT Calibration

The baseline approach is to merge all years of data taking. This is mainly motivated due to the same centre-of-mass energy and no larger changes in the detector. Nevertheless slight differences in the reconstruction are expected, that could have an effect on the FT. The baseline flavour tagging approach is to use one calibration function per decay channel and tagging algorithm and not use a specific function for each track type and/or year of data taking. The reasoning behind this approach is that although the η distributions differ between the track type, no difference in the flavour tagging calibration function is expected. The smaller number of constraint

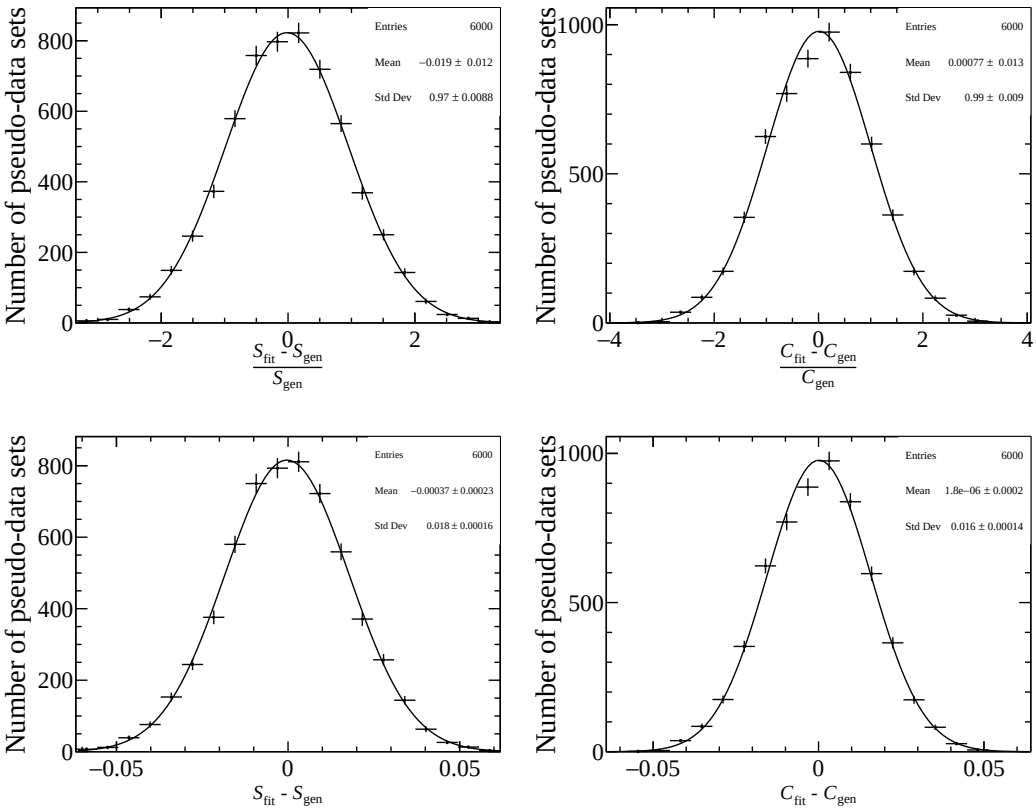

Figure 6.26: Pull (top) and residual (bottom) distribution for S (left) and C (right) for 6000 pseudo-experiments to validate the fitter with a Gaussian fit to each of the distributions.

parameters in the *CP* fit makes the fit more robust. To check for a possible bias, pseudo-experiments are generated using separate η, $m(B^0)$ and acceptance functions for each year and track type. The generating PDFs for these pseudo-experiments are extracted from sweighted data or a mass fit on data for each year separately. The generated data sets for the different track types and years of data taking are merged. The merged data sets are fit with baseline fit, which uses a common flavour tagging calibration. The resulting pull and residual distribution is shown in Fig. 6.27.

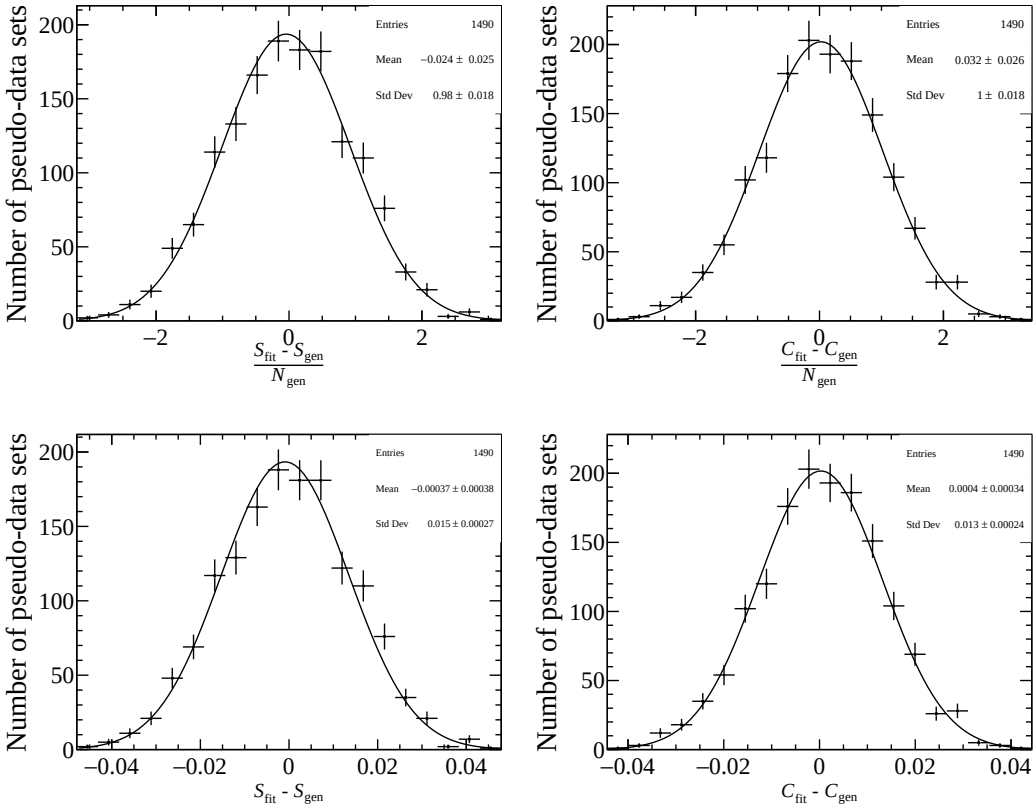

Figure 6.27: Pull (top) and residual (bottom) distribution for S (left) and C (right) with a Gaussian fit for 1490 pseudo-experiments where the data set is generated with independent PDFs for the flavour tagging in each year and track type.

No bias is seen with a sensitivity of approximately 3% of the statistical uncertainty and is covered by the uncertainty from the fit validation. Therefore, no additional systematic uncertainty is assigned due to the merging of the years of data taking.

6.7.4 Production Asymmetry

There is no official measurement of the B^0 production asymmetry, A_{prod}, for Run II so far. The Run I measured value is consistent with 0[99]. The baseline method is to extract the production asymmetry from $B^0 \to J/\psi K^*$ control data, which results in $A_{\text{prod}} = -0.0038 \pm 0.0018$ is described in Sec. 6.6.3. This extracted value does not include a proper evaluation of systematic uncertainties. However, in the Run 1 measurement [99] the systematic uncertainty for A_{prod} is small, which is expected to be similar in the presented measurement. To assign a systematic uncertainty due to the external value of the A_{prod}, an alternative value is extracted using the Run 1 measurement and scale it to the seen signal kinematics. This alternative value is $-1.2 \pm 0.6\%$ for the $B^0 \to J/\psi K^0_{\text{S}}$ kinematics. Pseudo-experiments are generated with this alternative value of A_{prod}, while the production asymmetry is constrained in the fit to the baseline value. The pull and residual distributions are shown in Fig. 6.28. The difference between generated and fitted values of the *CP* parameters for $B^0 \to J/\psi K^0_{\text{S}}$ decays is $-21.2 \pm 2.1\%$ of the statistical uncertainty for S and $28.6 \pm 2.2\%$ for C.

6.7.5 Effect of $\Delta\Gamma_d$

The decay width difference $\Delta\Gamma_d$ is known to be small. The current PDG value is $\Delta\Gamma_d/\Gamma_d = (0.1 \pm 1.0)\,\%$. In the baseline fit $\Delta\Gamma_d$ is set to zero. To evaluate the effect of a non zero value of $\Delta\Gamma_d$, pseudo-experiments are generated with a value of $\Delta\Gamma_d = 0.0073\,\text{ps}^{-1}$ or $\Delta\Gamma_d = -0.0073\,\text{ps}^{-1}$, corresponding to roughly one sigma up or down. Further, the value of $A_{\Delta\Gamma}$ is set to ± 0.7, which roughly satisfies the normalisation condition $S^2 + C^2 + A^2_{\Delta\Gamma} = 1$. These pseudo-data sets are fitted with the baseline fit model. The largest effect on the measured CP asymmetry arose from the term $A\Delta\Gamma \cdot \sinh(\frac{1}{2}\Delta\Gamma \cdot t)$, which is linear in $\Delta\Gamma$ for a small value of $\Delta\Gamma$. The $\cosh(\frac{1}{2}\Delta\Gamma \cdot t)$ term changes quadratically with $\Delta\Gamma$ and is therefore negligible. Two different setups are tested: $A_{\Delta\Gamma}$ and $\Delta\Gamma$ with the same sign and with different signs. The resulting pull distributions are given in Figs. 6.29 and 6.30 and show a bias of around 27% of the statistical uncertainty for both setups, where the bias has a different sign as expected.

6.7.6 Uncertainty of the Decay-time Bias Correction

In the CP fit input parameters with uncertainties are constrained. Therefore, statistical and systematic uncertainties from input parameters are covered by the statistical uncertainty of the fit. For the decay-time bias correction function,

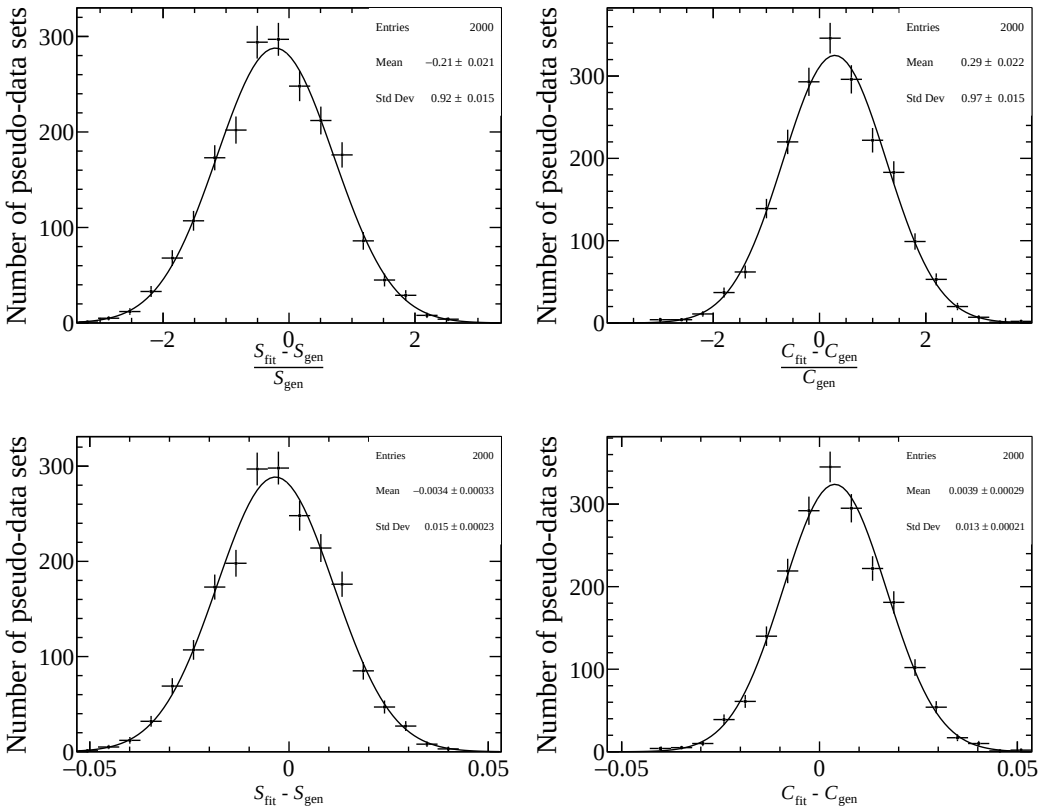

Figure 6.28: Pull (top) and residual (bottom) distribution for S (left) and C (right) with a Gaussian fit for 2000 pseudo-experiments where the production asymmetry is set to the alternative value of -1.2% in the generation and constrained to the baseline value from $B^0 \to J/\psi K^*$ -0.0038 ± 0.0018 in the fit.

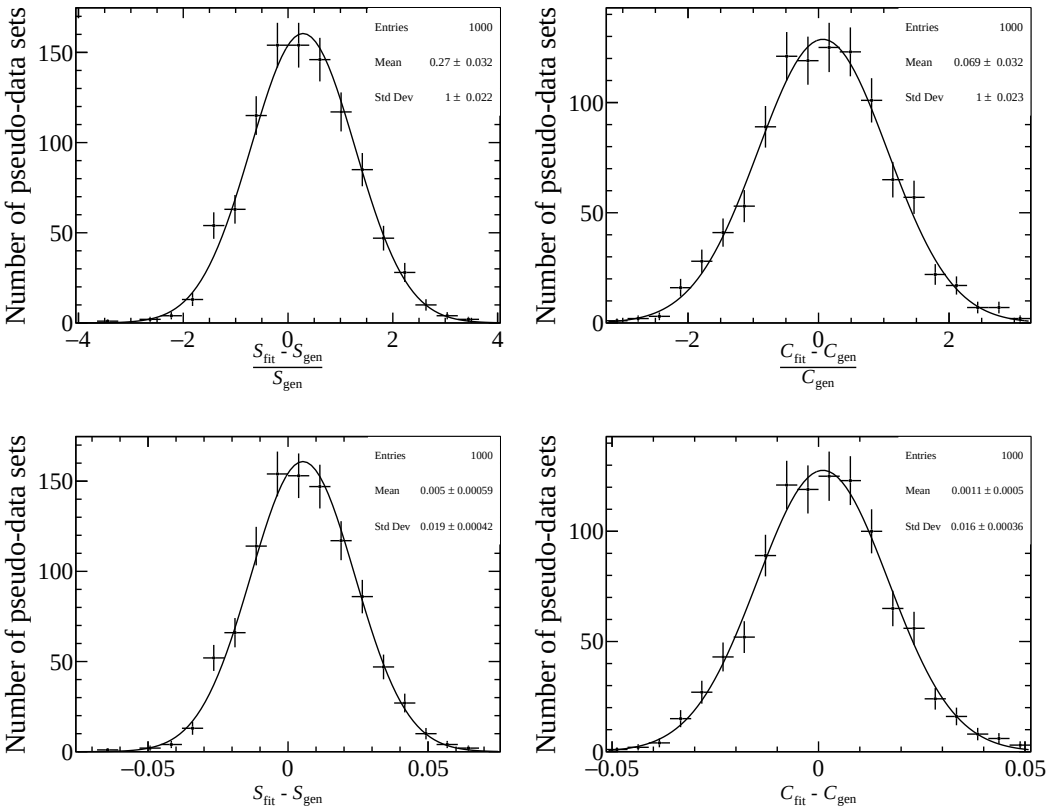

Figure 6.29: Pull (top) and residual (bottom) distribution for S (left) and C (right) with a Gaussian fit for 1000 pseudo-experiments where $\Delta\Gamma_d = 0.0073\,\mathrm{ps}^{-}1$ and $A_{\Delta\Gamma} = -0.7$ were used in the generation.

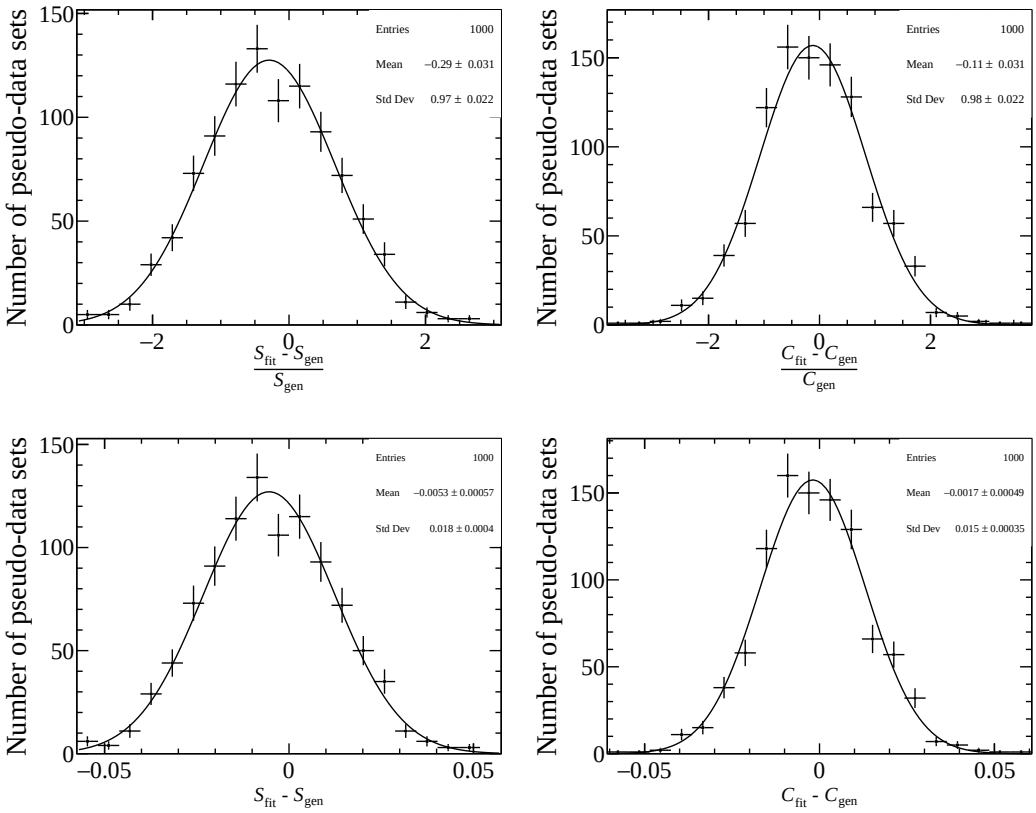

Figure 6.30: Pull (top) and residual (bottom) distribution for S (left) and C (right) with a Gaussian fit for 1000 pseudo-experiments where $\Delta\Gamma_d = -0.0073\,\mathrm{ps}^{-}1$ and $A_{\Delta\Gamma} = -0.7$ were used in the generation.

this is not easily possible for technical reasons. To transfer the uncertainties of the three parameters in the calibration for each track type to the parameters of interest, pseudo-data sets are created, where the parameters are varied within their uncertainty. In the variation, the parameters are assumed to be uncorrelated, which is a conservative assumption. These data sets are fitted using the baseline correction. The pulls and residuals distributions of these fits are shown in Fig. 6.31. The effect is compatible with zero, within the statistical uncertainty of the toy study of 2%. Thus, the possible systematic uncertainty, due to the uncertainty of the decay-time bias correction is negligible.

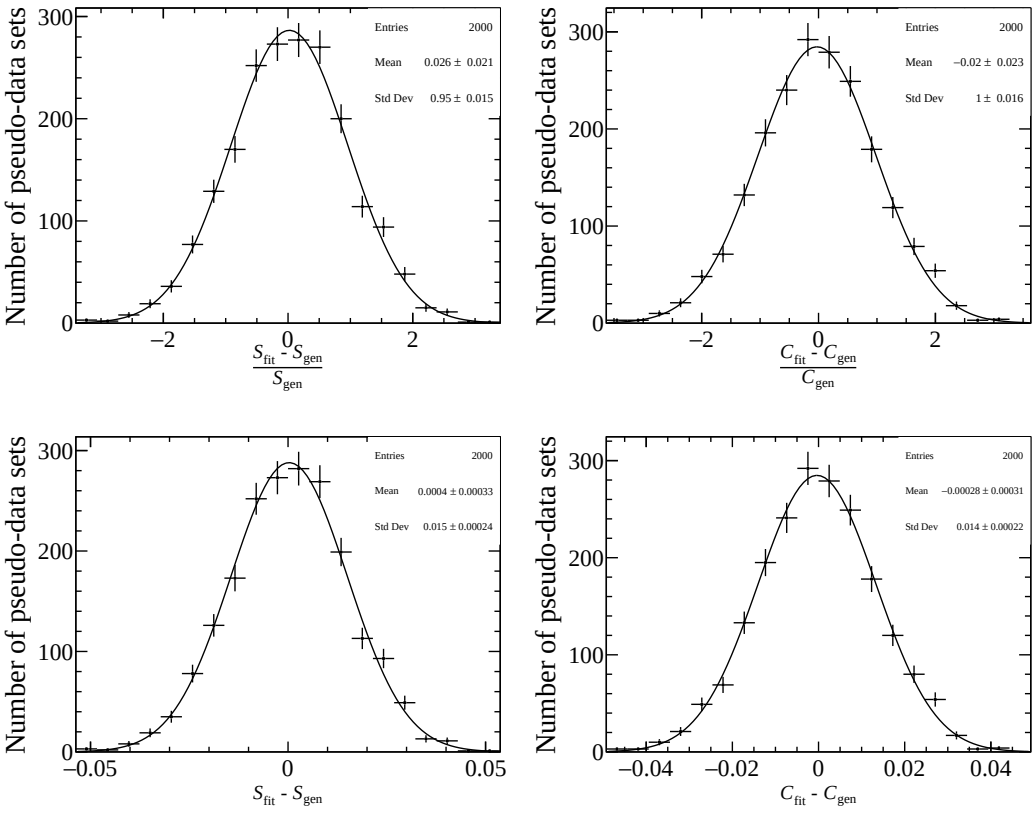

Figure 6.31: Pull (top) and residual (bottom) distribution for S (left) and C (right) with a Gaussian fit for 1000 pseudo-experiments where the decay-time bias calibration parameters are varied within their uncertainty in the generation.

6.8 Result

The simultaneous fit of the *CP* parameters in $B^0_d \to \psi K^0_S$ decays (fully described in Sec. 6.6) gives the statistical sensitivity on the two *CP* parameters in the decay. The combination of these statistical uncertainties with the systematic uncertainties, which are evaluated in Sec. 6.7, is given as

$$\sigma(S)_{B^0 \to \psi K^0_S} = 0.0133\,(\text{stat}) \pm 0.0063\,(\text{syst}),$$
$$\sigma(C)_{B^0 \to \psi K^0_S} = 0.0125\,(\text{stat}) \pm 0.0042\,(\text{syst}).$$

Here, the same systematic uncertainty is assumed for all three decay modes. This is a good assumption because the leading systematic uncertainties are the uncertainties from input values of A_{prod} and $\Delta\Gamma_d$, which are the very same between each decay channel. The combined uncertainty of $\sigma(S)_{B^0 \to \psi K^0_S} = 0.015$ corresponds to a sensitivity of about 0.6° on the CKM angle β. This is an improvement to the current world average $S = 0.699 \pm 0.017$ [12]. The statistical uncertainty on three *CP* parameter in $B^0_s \to J/\psi K^0_S$ decays is

$$\sigma(S)_{B^0_s \to J/\psi K^0_S} = 0.29,$$
$$\sigma(C)_{B^0_s \to J/\psi K^0_S} = 0.30,$$
$$\sigma(\mathcal{A}_{\Delta\Gamma})_{B^0_s \to J/\psi K^0_S} = 0.27.$$

In $B^0_s \to J/\psi K^0_S$ the systematic uncertainties need to be evaluated but this measurement will be statistically limited due to the small branching fraction.

7 Conclusion

Two measurements of neutral b hadrons to ψ and a neutral s hadron are presented. The ratio of branching fractions of $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$ decays is measured, and has been published [7]. The result is

$$\frac{\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\Lambda)} = 0.513 \pm 0.023\,(\text{stat}) \pm 0.016\,(\text{syst}) \pm 0.011(\mathcal{B}),$$

where the first uncertainty corresponds to the statistical uncertainty, the second corresponds to systematic uncertainty and the last one is due to the used branching fractions of $\psi(2S)$ and J/ψ decaying to two muons. The measurement confirms a tension seen by the ATLAS collaboration [4] to a theory prediction [5, 6], with an higher sensitivity in this new measurement.

The second ongoing presented measurement is the time-dependent *CP* violation measurement in $B^0 \to \psi K_S^0$ decays. Here, new reconstruction combinations of K_S^0 mesons are used to maximise the statistical sensitivity. Although the analysis is still blinded, the sensitivity is evaluated to

$$\sigma(S)_{B^0 \to \psi K_S^0} = 0.0133\,(\text{stat}) \pm 0.0064\,(\text{syst})^1,$$
$$\sigma(C)_{B^0 \to \psi K_S^0} = 0.0125\,(\text{stat}) \pm 0.0043\,(\text{syst})^1,$$

for the B^0 meson decays, corresponding to a combined sensitivity of about 0.6° on the CKM angle β.

The sensitivities for the *CP* parameters in $B_s^0 \to J/\psi K_S^0$ decays are evaluated to be

$$\sigma(S)_{B_s^0 \to J/\psi K_S^0} = 0.29,$$
$$\sigma(C)_{B_s^0 \to J/\psi K_S^0} = 0.30,$$
$$\sigma(\mathcal{A}_{\Delta\Gamma})_{B_s^0 \to J/\psi K_S^0} = 0.27.$$

The *CPV* measurements will be statistically limited. The $\sin(2\beta)$ measurement in the $B^0 \to \psi K_S^0$ decays and will supersede the current world average, while the $B_s^0 \to J/\psi K_S^0$ measurement will improve the former LHCb measurement [100].

In the future, the sensitivity of this measurement may be further improved by using a new tagging algorithm [90], which has access to the full event information using a

[1] This assumes the same systematic uncertainty for all three decay modes.

deep neural net instead of having individual taggers for individual physics processes. This tagging algorithm is currently under development. Since the tagging power is around 4-5% in an hadron collider environment, small absolute improvements result in a large relative improvement.

Additionally, as the sensitivity of *CPV* measurements in $B^0 \to \psi K^0_S$ decays increases in the future, the impact of penguin contributions will become more and more relevant to the measurement of the CKM angle β that can be performed at Belle II [101] and with Run 3 data at LHCb [102]. Continuing measurements of CPV in $B^0_s \to J/\psi K^0_S$ decays will help to constrain these penguin contributions, and will be a crucial part of these future *CPV* measurements.

A Appendix

A.1 Additonal Tables and Figures ($\Lambda_b^0 \to \psi(2S)\Lambda^0$)

Table A.1: The requirements of the used triggerlines decided in the three trigger stages [103, 104]. Additional requirements for 2012 data taking period are given in parenthesis.

L0	Muon	DiMuon		
p_T	$> 1.48(1.76)\,\text{GeV}/c$	-		
$p_{T_1} \cdot p_{T_2}$	-	$> 1.30^2(1.60^2)\,\text{GeV}^2/c^2$		
SPD hits	< 600	< 900		
Hlt1	TrackMuon	SingleMuonNoIP	TrackAllL0	DiMuonHighMass
Track p_T	$> 1\,\text{GeV}/c$	$> 4.8\,\text{GeV}/c$	$> 1.7\,\text{GeV}/c$	$> 0.5\,\text{GeV}/c$
Track p	$> 8\,\text{GeV}/c$	$> 8\,\text{GeV}/c$	$> 10\,\text{GeV}/c$	$> 6.0\,\text{GeV}/c$
Track χ^2/ndf	< 2	< 4	< 2.5	< 4
DOCA	-	-	-	$< 0.2\,\text{mm}$
Vertex χ^2	-	-	-	< 25
Track IP	$> 0.1\,\text{mm}$	-	$> 0.1\,\text{mm}$	-
Track IPχ^2	> 16	-	> 16	-
$m_{\mu\mu}$	-	-	-	$> 2.7\,\text{GeV}/c^2$
Hlt2	DiMuonDetached	DiMuonDetachedHeavy		
Track χ^2/ndf	< 5	< 5		
Vertex χ^2	< 25	< 25		
$m_{\mu\mu}$	$> 1\,\text{GeV}/c^2$	$> 2.95\,\text{GeV}/c^2$		
$p_{T_{\mu\mu}}$	$> 1.5\,\text{GeV}/c$	-		
FD χ^2	> 49	> 25		
χ^2(IP)	> 9	-		

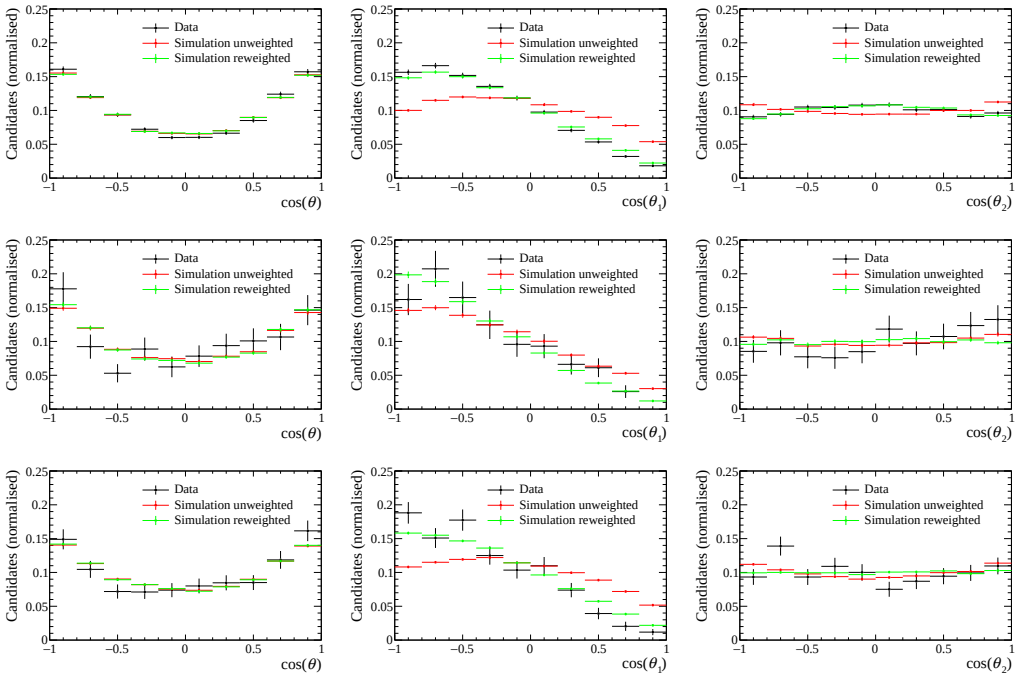

Figure A.1: Distributions for (left) $\cos(\Theta)$, (middle) $\cos(\Theta_1)$, (right) $\cos(\Theta_2)$ for sweighted Data (blue), unweighted MC (red) and reweighted MC (green) for and (top) downstream $\Lambda_b^0 \to J/\psi\Lambda$, (middle) long track $\Lambda_b^0 \to \psi(2S)\Lambda$ and (bottom) downstream $\Lambda_b^0 \to \psi(2S)\Lambda$ decays.

A.2 Additonal Tables and Figures ($B^0 \to \psi K^0_S$)

Table A.2: Preselection requirements for reconstructed $B^0 \to J/\psi\, K^0_S$ and $B^0 \to \psi(2S) K^0_S$ candidates. The pion subscripts denote the different reconstruction track types long (L), downstream (D) and upstream track (U).

Target	Observable	Requirement	Description
B^0	m	$\in [5,6]\,\mathrm{GeV}/c^2$	Reconstructed B^0 mass
	$\chi^2_{\text{Vtx, NDOF}}$	< 10	Vertex χ^2 per degree of freedom
J/ψ	$\lvert m - m_{\text{PDG}}\rvert$	$< 100\,\mathrm{MeV}/c^2$	mass difference to PDG value [16]
	$\chi^2_{\text{Vtx, NDOF}}$	< 20	
	$\lvert\min(\sigma_{\text{FD}})\rvert$	> 3	Decay length significance with respect to the PV
$\psi(2S)$	$\lvert m - m_{\text{PDG}}\rvert$	$< 100\,\mathrm{MeV}/c^2$	
	$\chi^2_{\text{Vtx, NDOF}}$	< 20	
	$\min p_{\text{T}}(\mu)$	$> 500\,\mathrm{MeV}/c$	
	$\lvert\min(\sigma_{\text{FD}})\rvert$	> 3	Min. decay length significance
$\mu^\pm$	p_{T}	$> 500\,\mathrm{MeV}/c$	
	DLL_μ	> 0	Difference between log-likelihood between muon and pion hypothesis
K^0_S	$\lvert\min(\sigma_{\text{FD}})\rvert$	> 5	
	χ^2_{Vtx}	< 25	Vertex χ^2
	$\lvert m - m_{\text{PDG}}\rvert$	$< 80\,\mathrm{MeV}/c^2$	
$\pi^\pm_L$	p	$> 2\,\mathrm{GeV}/c$	
	$\min(\chi^2_{\text{IP}})(\text{PV})$	> 9	Abs. difference of PV χ^2 with and without this track
$\pi^\pm_D$	p	$> 2\,\mathrm{GeV}/c$	
	$\min(\chi^2_{\text{IP}})(\text{PV})$	> 4	
$\pi^\pm_U$	p	$> 1\,\mathrm{GeV}/c$	
	$\min(\chi^2_{\text{IP}})(\text{PV})$	> 9	

Table A.3: Stripping requirements for $B^0 \to J/\psi(\to ee)K^0_S$ using the `Bd2JpsieeKSBd2JpsieeKSDetachedLine` from the leptonic stream. The kaon and pion selection differ for the LL and DD track type combination of the pions. The pion subscripts denote the track types L and D.

Target	Observable	Requirement	Description
B^0	m	$\in [4.4, 6]\,\mathrm{GeV}/c^2$	Reconstructed B^0 mass
	$\chi^2_{\mathrm{Vtx,\ NDOF}}$	$< 20/3$	Vertex χ^2 per degree of freedom
	τ	$> 0.2\,\mathrm{ps}$	Decay time
	DIRA(PV)	> 0.995	Cosine of opening angle of direction of flight w.r.t. the PV
J/ψ	m	$\in [2.3, 3.3]\,\mathrm{GeV}/c^2$	Reconstructed J/ψ mass
	$\chi^2_{\mathrm{vertex}}/\mathrm{NDOF}$	< 15	
	$\log(\chi^2_{\mathrm{IP}})$	> 0.5	χ^2 of the IP
$e^\pm$	p_{T}	$> 500\,\mathrm{MeV}/c$	
	DLL_e	> 0	Delta log-likelihood between electron and pion hypothesis
	$\chi^2_{\mathrm{Track}}/\mathrm{NDOF}$	< 5	Track fit χ^2 per degree of freedom
	$\max(\log(\chi^2_{\mathrm{IP}}))$	> 2.5	Max. of the log(IP χ^2)
	$\min(\log(\chi^2_{\mathrm{IP}}))$	> -0.5	Min. of the log(IP χ^2)
	$\mathrm{ProbNN_e}$	> 0.01	Neural net PID for electrons
K^0_S	$\lvert\min(\sigma_{\mathrm{FD}})\rvert$	> 5	Min. decay length significance
	χ^2_{Vtx}	< 20	Vertex χ^2
	$\lvert m - m_{\mathrm{PDG}}\rvert$	$< 64(35)\,\mathrm{MeV}/c^2$	For track types DD (LL)
	$\mathrm{FD_{ORIVX}}$	$> 25\,\mathrm{mm}$	Flight distance w.r.t. its origin vertex
	$\chi^2(\mathrm{FD_{ORIVX}})$	> 40	χ^2 of flight distance w.r.t. its origin vertex
	$\lvert m - m_{\mathrm{PDG}}\rvert$	$< 33(25)\,\mathrm{MeV}/c^2$	for track types DD (LL)
$\pi^\pm_L$	p	$> 2\,\mathrm{GeV}/c$	
	$\min(\chi^2_{\mathrm{IP}})(\mathrm{PV})$	> 9	Abs. difference of primary vertex χ^2 with and without this track
	p_{T}	$> 250\,\mathrm{MeV}/c$	
	$\chi^2_{\mathrm{Track,\ DOF}}$	< 2.4	Track fit χ^2 per degree of freedom
	$\max(\log(\chi^2_{\mathrm{IP}}))$	> 4	
$\pi^\pm_D$	p	$> 2\,\mathrm{GeV}/c$	
	$\min(\chi^2_{\mathrm{IP}})(\mathrm{PV})$	> 4	
	$\chi^2_{\mathrm{Track,\ DOF}}$	< 2.4	

Table A.4: Trigger lines for the analysed modes. For the muon mode one of the lines has to be triggered on signal (TOS) in each trigger stage. For $B^0 \to J/\psi(\to ee)K^0_S$ the lines trigger on the full event (DEC). The numbers in the brackets indicate the years, where the lines are used, if they are not used in all years.

stage	$B^0 \to J/\psi K^0_S$	$B^0 \to \psi(2S)K^0_S$	$B^0 \to J/\psi(\to ee)K^0_S$
L0	Muon DiMuon L0MuonHigh	Muon DiMuon L0MuonHigh	Electron Hadron Muon DiMuon Photon JetEl MuonNoSPD(15) MuonEW(16-18)
HLT1	DiMuonHighMass TrackMuon	DiMuonHighMass TrackMuon	TrackMVA TwoTrackMVA
HLT2	DiMuonJPsi DiMuonDetachedJPsi	DiMuonPsi2S DiMuonDetachedHeavy	Topo[2,3]Body RadiativeIncHHGamma(15) Topo[E,EE]2Body(16-18)

Table A.5: Used BDT features in the dimuon mode. The first of variables is used in all track types. The same variables are used for $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S)K^0_S$ candidates.

LL	DD	UL and LD
	IP(B^0)	
	IP(J/ψ)	
	IP(K^0_S)	
	$p_t(\pi^+)$	
	$p_t(\pi^-)$	
	$p_t(K^0_S)$	
	DTF fit χ^2	
	$\eta(B^0)$	
	B^0 vertex χ^2/ndof	
	FD(K^0_S) to PV	
	DOCA(π^+, π^-)	
	$\eta(K^0_S)$	
	$p_z(K^0_S)$	
	Min IP($\mu^\pm$)	
	Max $p_t(\pi^\pm)$	
DOCA(J/ψ, K^0_S)	IP(π^+)	IP(π^+)
Min χ^2(DOCA(J/ψ, K^0_S))	IP(π^-)	IP(π^-)
Min IP(K^0_S) wrt all PVs	Decay time(K^0_S)	DOCA(J/ψ, K^0_S)
———	Min(IP($\pi^\pm$))	Min χ^2(DOCA(J/ψ, K^0_S))
———	Min IP(π^-) wrt all PVs	Min IP(K^0_S) wrt all PVs
———	Min IP(π^+) wrt all PVs	Decay time(K^0_S)
———	———	Min IP($\pi^\pm$)
———	———	Min IP(π^-) wrt all PVs
———	———	Min IP(π^+) wrt all PVs

Table A.6: Used BDT features for $B^0 \to J/\psi(e^+e^-)K^0_S$ candidates. The first set of variables is used in both track types.

LL	DD
IP(B^0)	
IP(J/ψ)	
IP(K^0_S)	
IP(π^+)	
$p_t(\pi^+)$	
IP(π^-)	
$p_t(\pi^-)$	
$p_t(K^0_S)$	
DTF fit χ^2	
$\eta(B^0)$	
B^0 vertex χ^2/ndof	
Min(IP($e^\pm$))	
FD(K^0_S) to PV	
DOCA(π^+, π^-)	
$\eta(K^0_S)$	
$p_z(K^0_S)$	
Decay time(K^0_S)	
Min(IP($\pi^\pm$))	
Max($p_t(\pi^\pm)$)	
DOCA(J/ψ, K^0_S)	———
Min χ^2(DOCA(J/ψ, K^0_S))	———

Table A.7: Selection and reconstruction efficiencies for $B^0 \to \psi(2S)K^0_S$ candidates determined on signal simulation.

step	efficiency [%]			
	LL		DD	
geometric	19.57 ± 0.05			
presel. and reco.	2.542	± 0.005	5.090	± 0.007
trigger	81.47	± 0.07	81.99	± 0.05
BDT	84.76	± 0.07	82.15	± 0.06
vetos	96.93	± 0.04	98.152	± 0.019
total	1.701	± 0.004	3.357	± 0.006

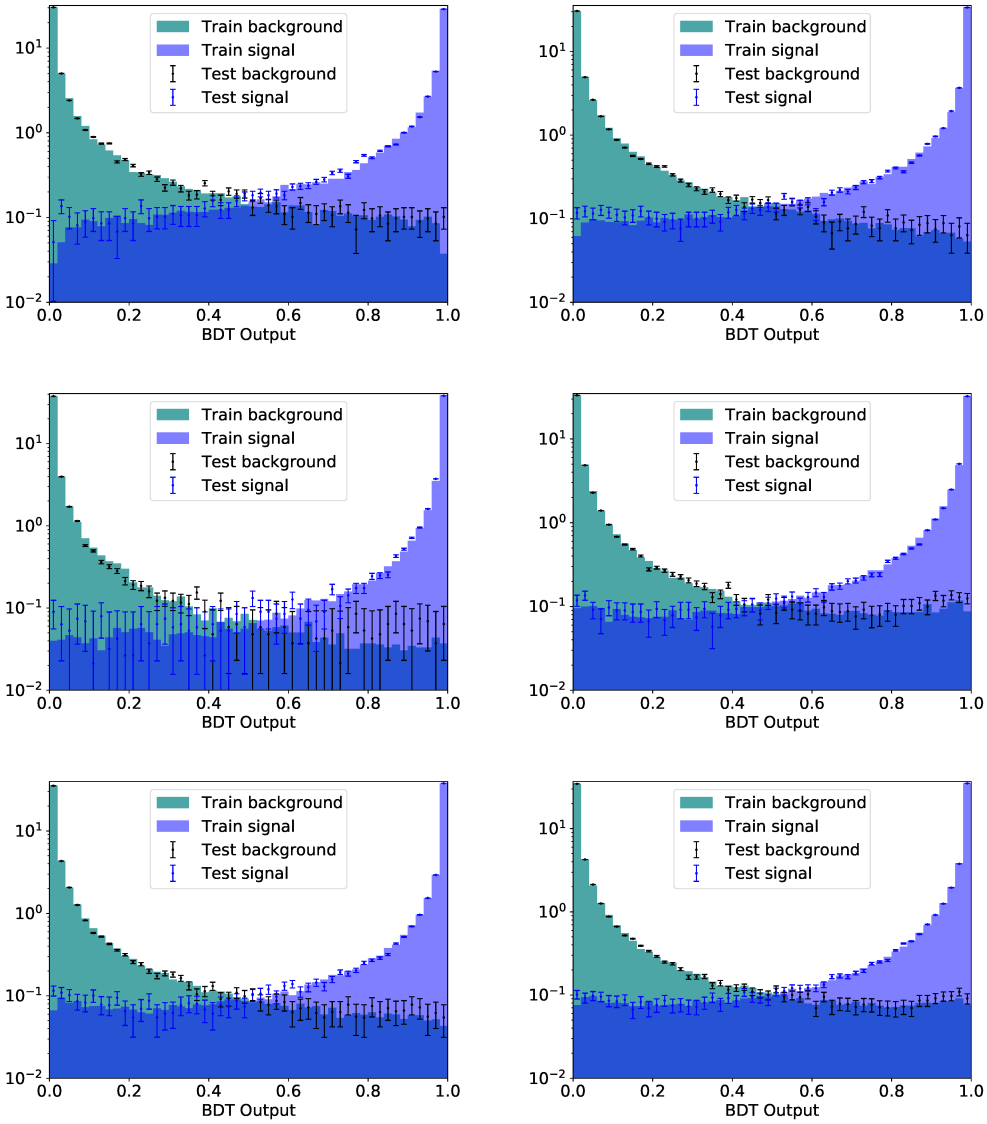

Figure A.2: The BDT classification output for signal (blue) and background (green/black), separate the training (solid) and test data set (points) for LD (upper-left) and DD (upper-right) $B^0 \to J/\psi(\mu^+\mu^-)K^0_S$, for LL (middle-left) and DD (middle-right) $B^0 \to J/\psi(e^+e^-)K^0_S$ and for LL (bottom-left) and DD (bottom-right) $B^0 \to \psi(2S)K^0_S$ candidates.

Table A.8: Selection and reconstruction efficiencies for $B^0 \to J/\psi(\to ee)K^0_S$ candidates determined on signal simulation.

step	efficiency [%]			
	LL		DD	
geometric	19.75 ± 0.05			
presel. and reco.	0.7978	± 0.0024	2.503	± 0.004
trigger	46.69	± 0.15	41.42	± 0.08
BDT	92.52	± 0.12	77.88	± 0.11
vetos	98.79	± 0.05	98.321	± 0.034
total	0.3402	± 0.0016	0.7920	± 0.0023

Table A.9: Selection and reconstruction efficiencies for $B^0_s \to J/\psi(\to \mu\mu)K^0_S$ candidates determined on signal simulation.

step	efficiency [%]			
	LL		DD	
geometric	19.86 ± 0.05			
presel. and reco.	2.023	± 0.007	4.785	± 0.011
trigger	78.33	± 0.15	77.15	± 0.10
BDT	80.46	± 0.16	61.95	± 0.13
vetos	96.55	± 0.08	98.230	± 0.035
total	1.231	± 0.005	2.242	± 0.008

Table A.10: Effective mean resolution in fs for each decay channel and track type, evaluated on signal simulation. The difference between B^0_s and B^0 decays is due to the non-linear average of the three Gaussians.

Track type	LL	DD	LD	UL
$B^0 \to J/\psi(\mu^+\mu^-)K^0_S$	59.9 ± 0.6	59.1 ± 0.4	64.5 ± 1.2	67.1 ± 1.7
$B^0 \to J/\psi(e^+e^-)K^0_S$	58.8 ± 1.7	59.7 ± 1.1	-	
$B^0 \to \psi(2S)K^0_S$	59.0 ± 1.0	54.3 ± 0.6	-	
$B^0_s \to J/\psi K^0_S$	43.99 ± 0.26	46.64 ± 0.16	-	

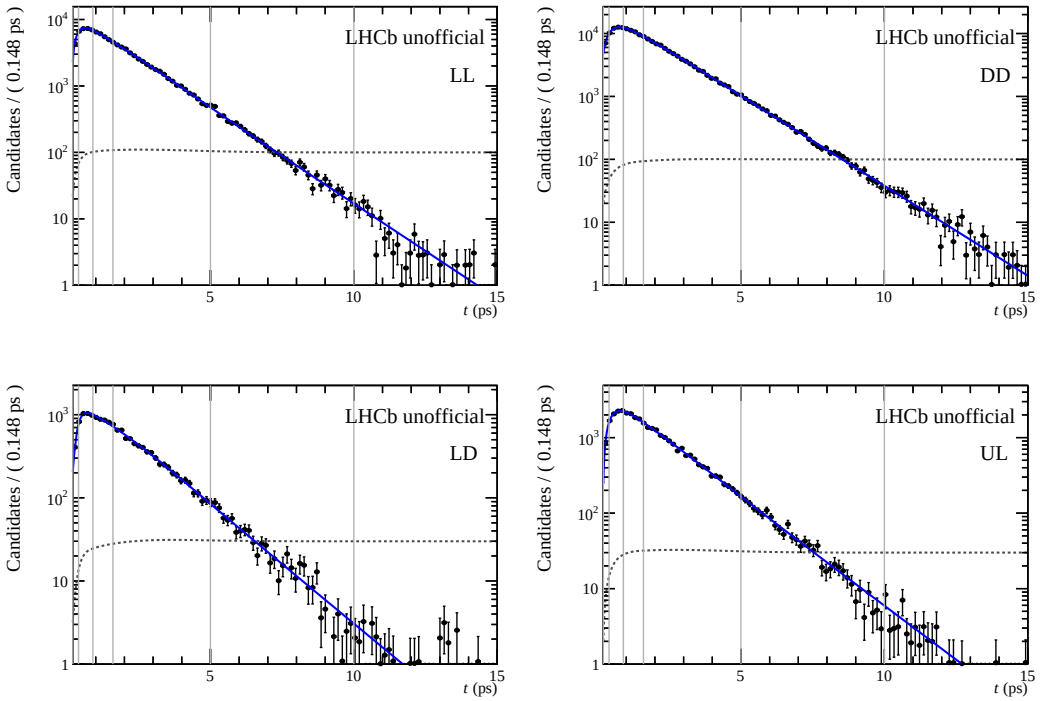

Figure A.3: Fit to the reconstructed B^0 decay time for (upper-left) LL, (upper-right) DD, (lower-left) LD and (lower-right) UL $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates with the knot positions (vertical lines) and the acceptance (dotted-grey).

Table A.11: Dilution in % for each decay channel and track type, evaluated on signal simulation. The difference between B^0_s and B^0 decays is due the large difference between Δm_d and Δm_s.

Track type	LL	DD	LD	UL
$B^0 \to J/\psi(\mu^+\mu^-)K^0_S$	99.9540 ± 0.0009	99.9552 ± 0.0006	99.9466 ± 0.0020	99.9423 ± 0.0030
$B^0 \to J/\psi(e^+e^-)K^0_S$	99.9556 ± 0.0025	99.9544 ± 0.0016	-	
$B^0 \to \psi(2S)K^0_S$	99.9553 ± 0.0016	99.9622 ± 0.0008	-	
$B^0_s \to J/\psi K^0_S$	73.73 ± 0.27	70.99 ± 0.17	-	

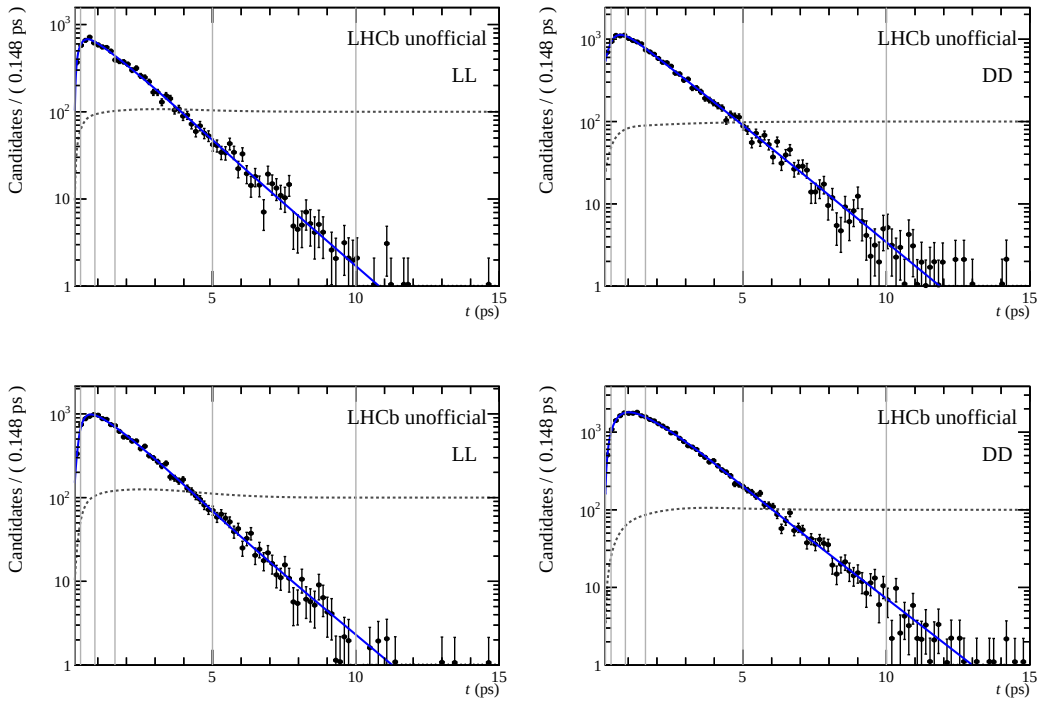

Figure A.4: Fit to the reconstructed B^0 decay time for (left) LL and (right) DD for (top) $B^0 \to \psi(2S)K^0_S$ and (bottom) $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates with the knot positions (vertical lines) and the acceptance (dotted-grey).

Table A.12: Effective mean resolution, σ_{eff}, and the mean bias, μ, for prompt $B^0 \to J/\psi\, K^0_S$ candidates separate for each track type.

	prompt data			
Track type	LL	DD	LD	UL
μ [fs]	-9.95 ± 0.06	-11.35 ± 0.05	-11.25 ± 0.09	9.98 ± 0.07
σ_{eff} [fs]	114.9 ± 0.7	159.6 ± 0.8	120.3 ± 0.9	144.5 ± 1.2

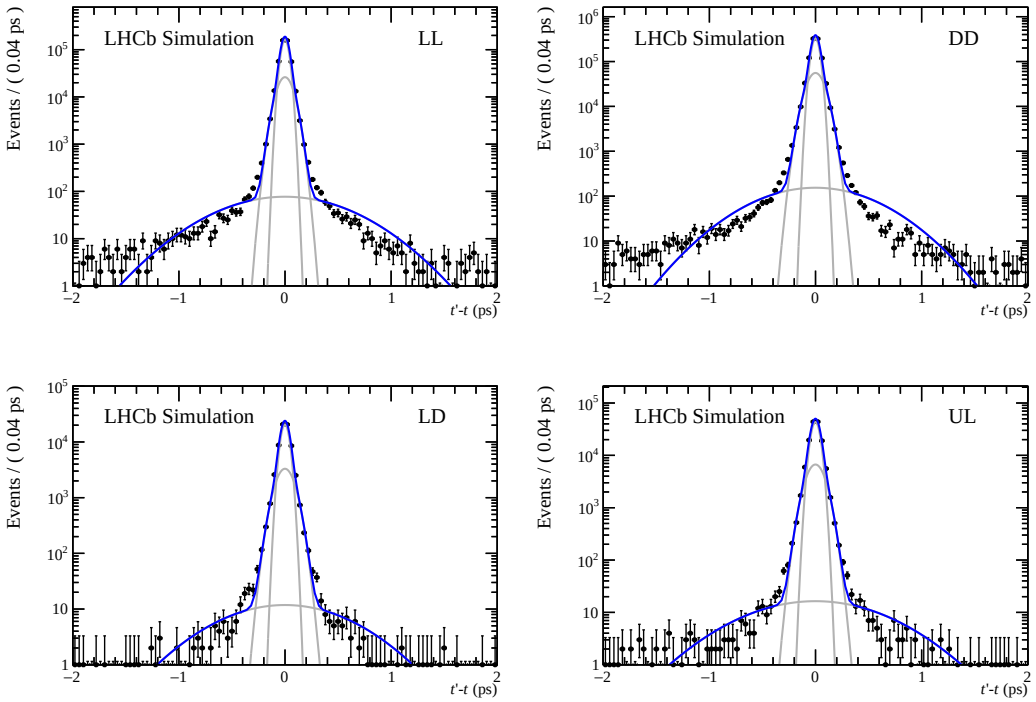

Figure A.5: Fit to the decay time resolution on signal simulation for (upper-left) LL, (upper-right) DD, (lower-left) LD and (lower-right) UL $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates with the sum of the three Gaussians (blue) and the single Gaussians (grey).

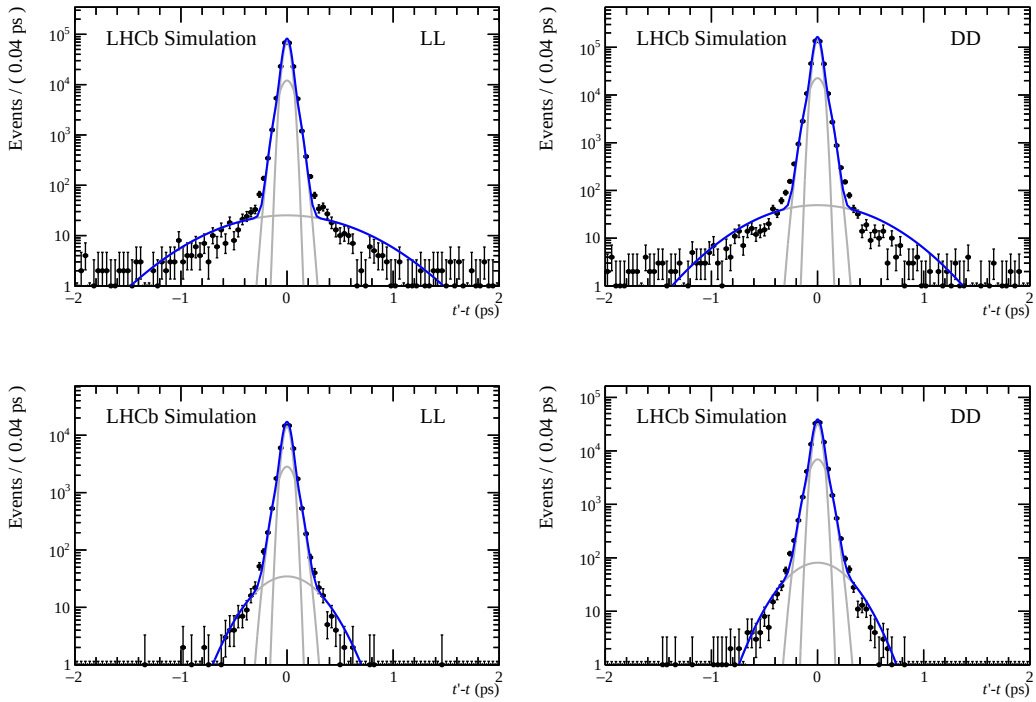

Figure A.6: Fit to the decay time resolution on signal simulation for (left) LL and (right) DD for (top) $B^0 \to \psi(2S)K^0_S$ and (bottom) $B^0 \to J/\psi(\to ee)K^0_S$ candidates with the sum of the three Gaussians (blue) and the single Gaussians (grey).

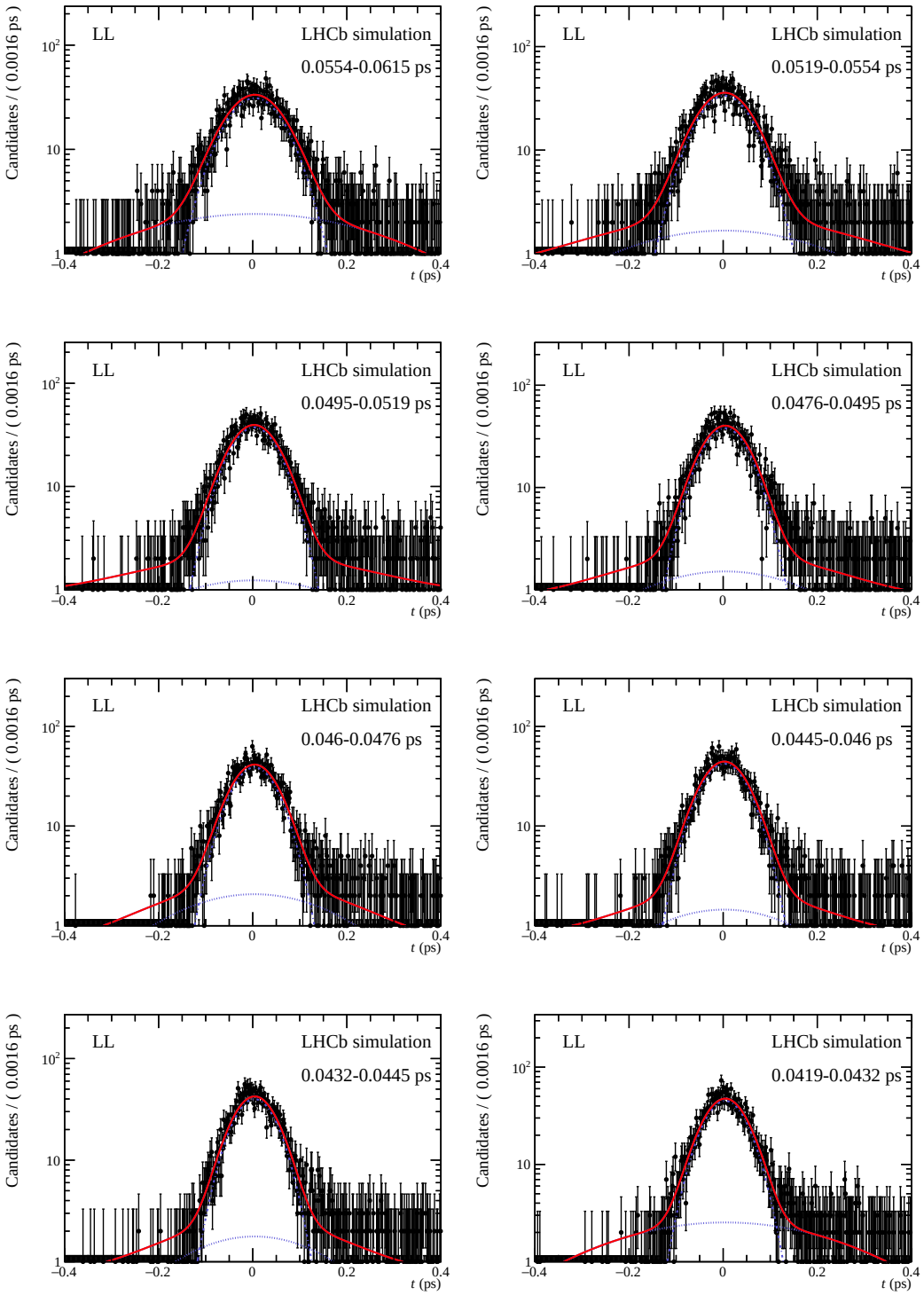

Figure A.7: Fit to the mean decay-time resolution on long track $B^0 \to J/\psi K^0_S$ simulation in different bins of the decay-time resolution estimate with the sum of the three Gaussians (red, solid) and the single Gaussians.

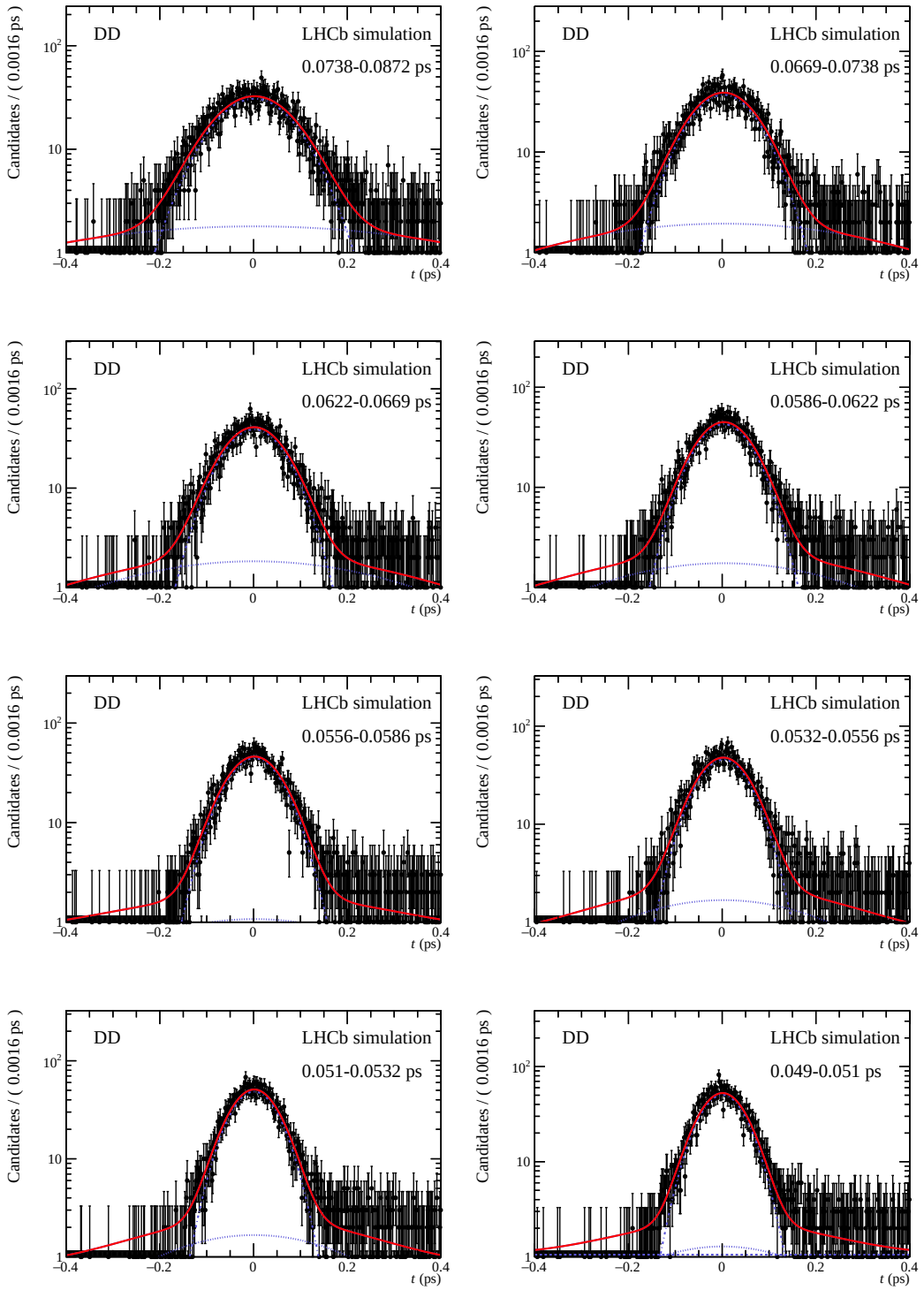

Figure A.8: Fit to the mean decay-time resolution on downstream track $B^0 \to J/\psi K^0_S$ simulation in different bins of the decay-time resolution estimate with the sum of the three Gaussians (red, solid) and the single Gaussians.

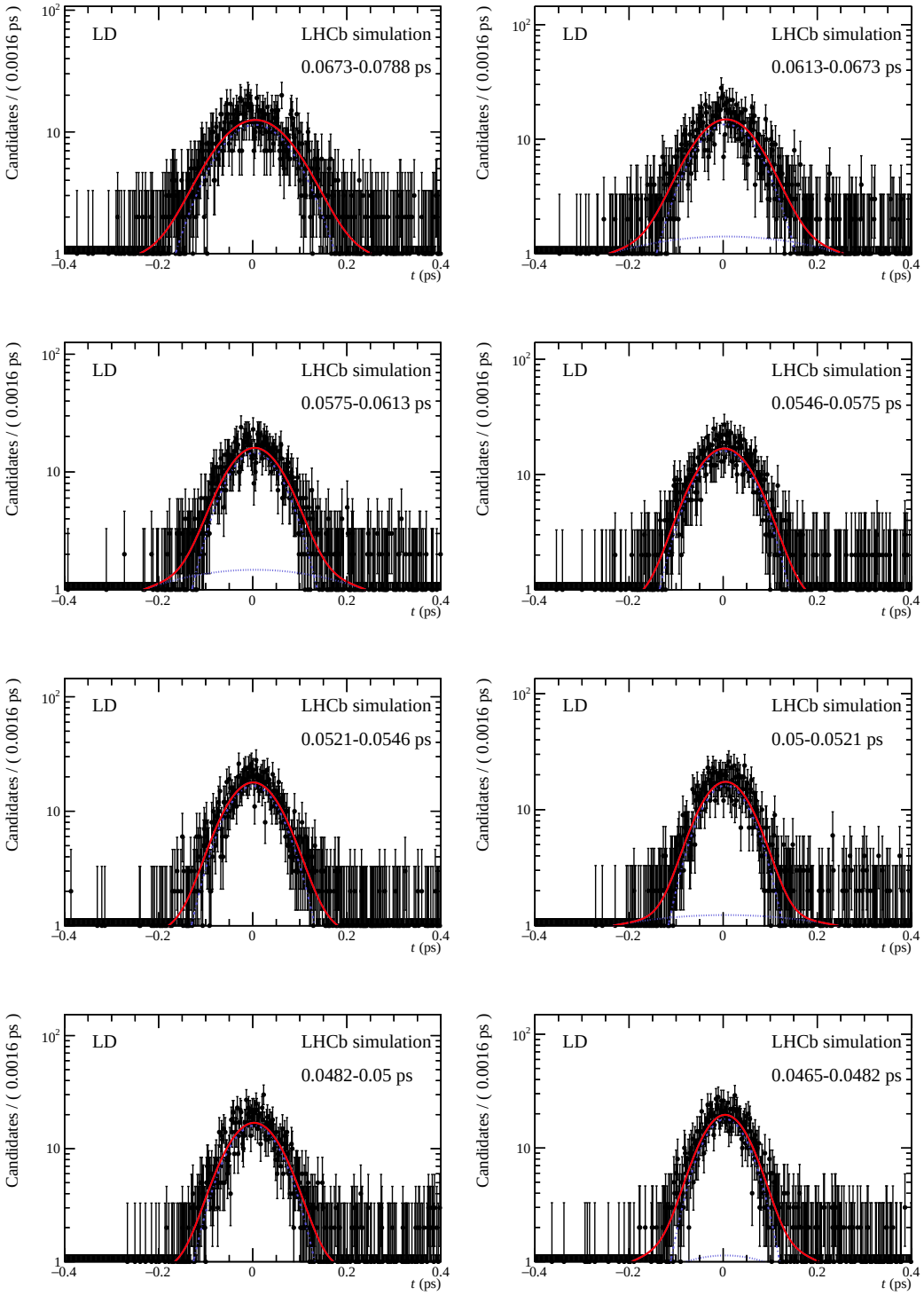

Figure A.9: Fit to the mean decay-time resolution on long-downstream track
124 $B^0 \to J/\psi K^0_S$ simulation in different bins of the decay-time resolution estimate with the sum of the three Gaussians (red, solid) and the single Gaussians.

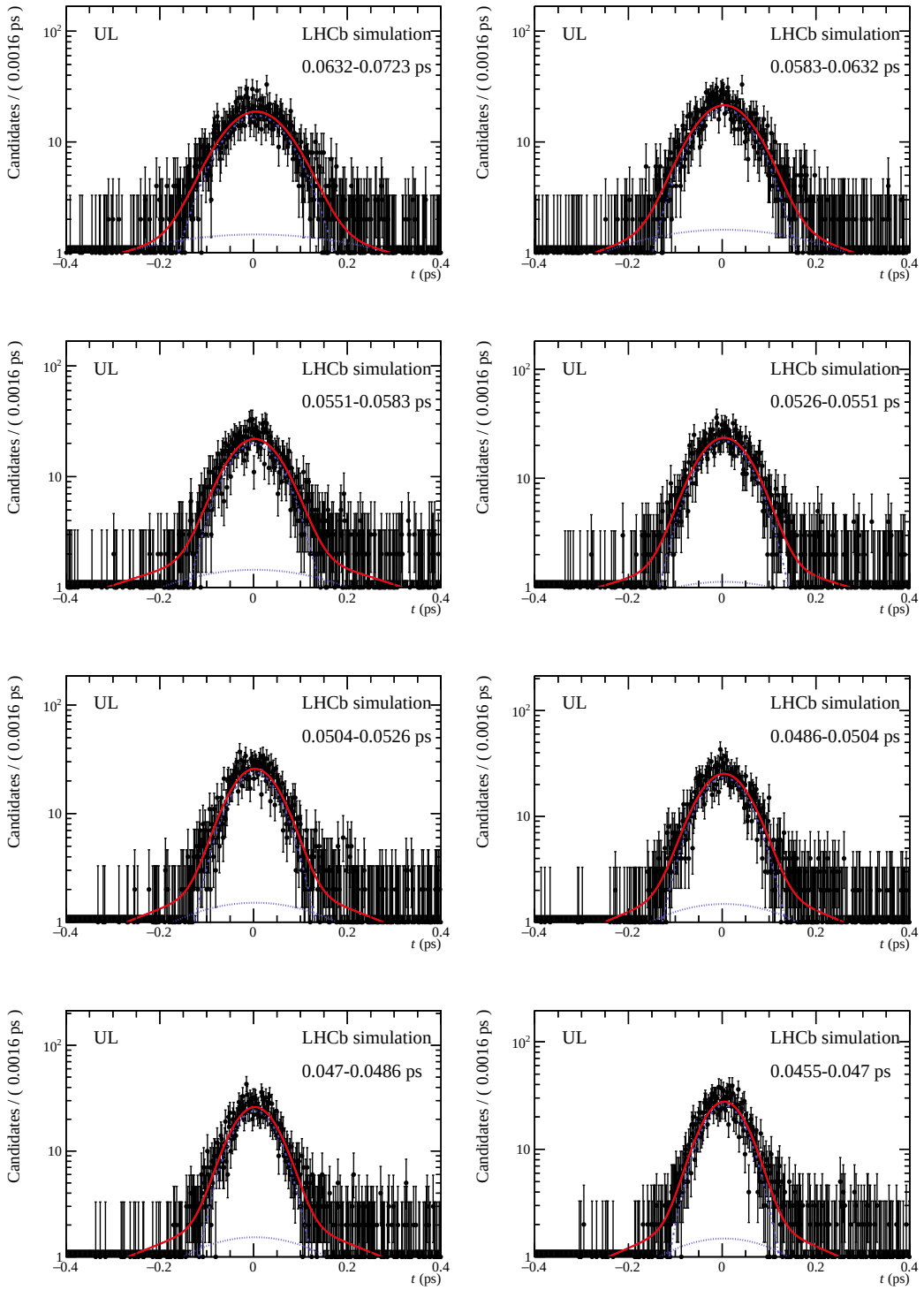

Figure A.10: Fit to the mean decay-time resolution on upstream-long track $B^0 \to J/\psi K^0_S$ simulation in different bins of the decay-time resolution estimate with the sum of the three Gaussians (red, solid) and the single Gaussians.

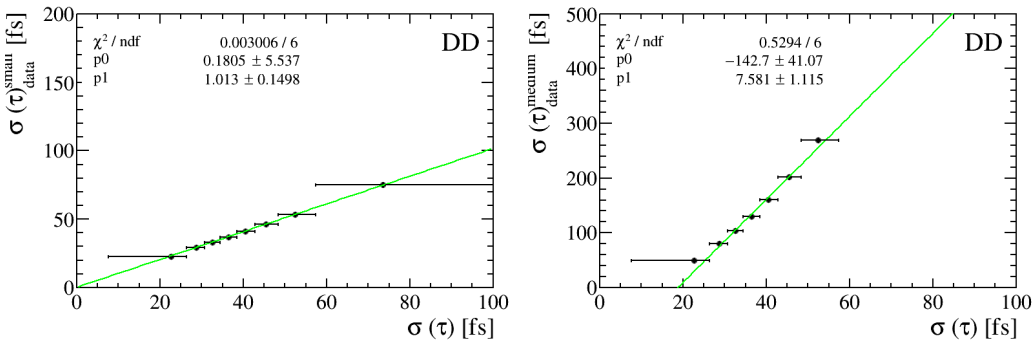

Figure A.11: Width of the small (left) and the medium Gaussian (right) for downstream track $B^0 \to J/\psi K^0_S$ simulation in bins of the decay-time resolution estimate. The width are fitted with a linear function.

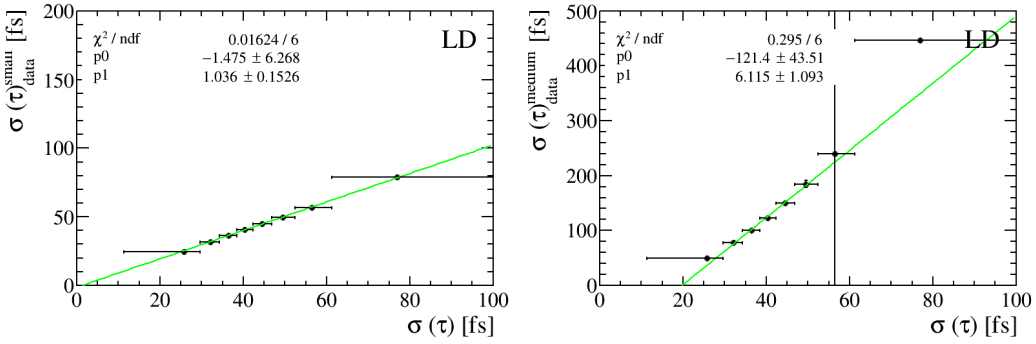

Figure A.12: Width of the small (left) and the medium Gaussian (right) for long-downstream track $B^0 \to J/\psi K^0_S$ simulation in bins of the decay-time resolution estimate. The width are fitted with a linear function.

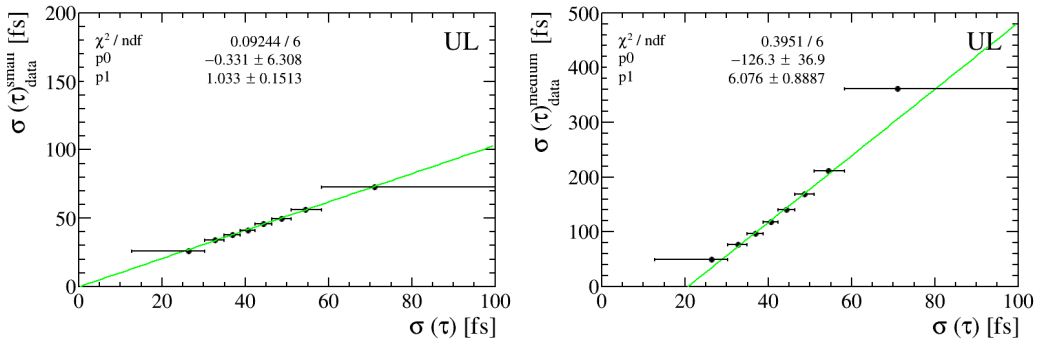

Figure A.13: Width of the small (left) and the medium Gaussian (right) for upstream-long track $B^0 \to J/\psi K^0_S$ simulation in bins of the decay-time resolution estimate. The width are fitted with a linear function.

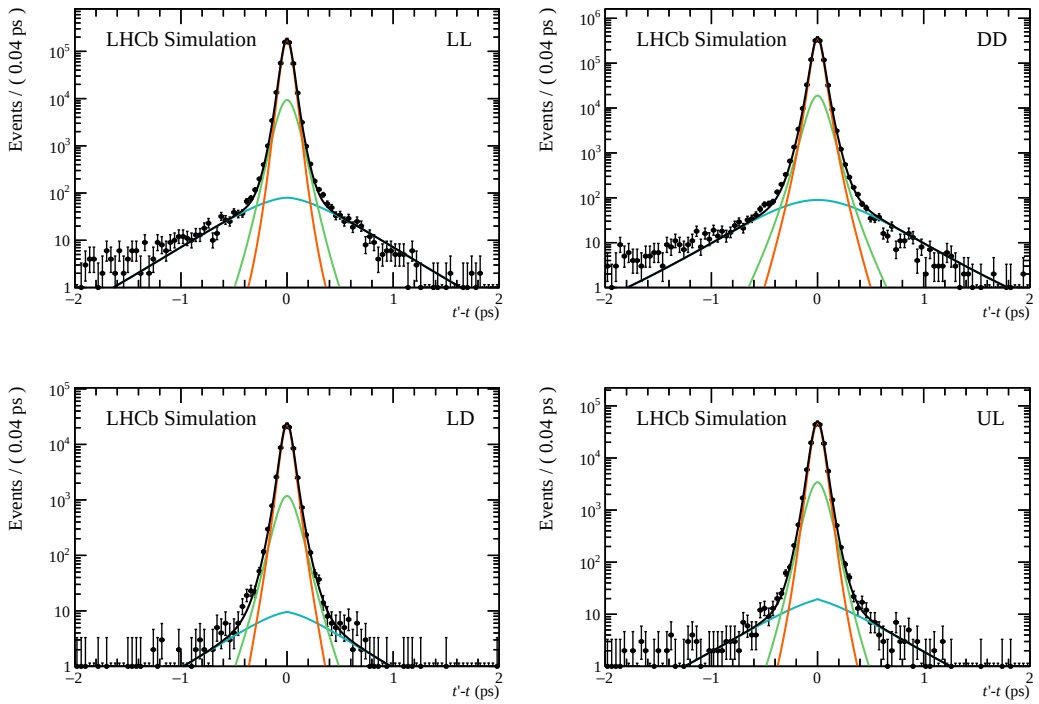

Figure A.14: Per-candidate fit to the decay time resolution on signal simulation for (upper-left) LL, (upper-right) DD, (lower-left) LD and (lower-right) UL $B^0 \to J/\psi(\to \mu\mu)K^0_S$ candidates with the sum of the three Gaussians (black) and the single Gaussians (coloured).

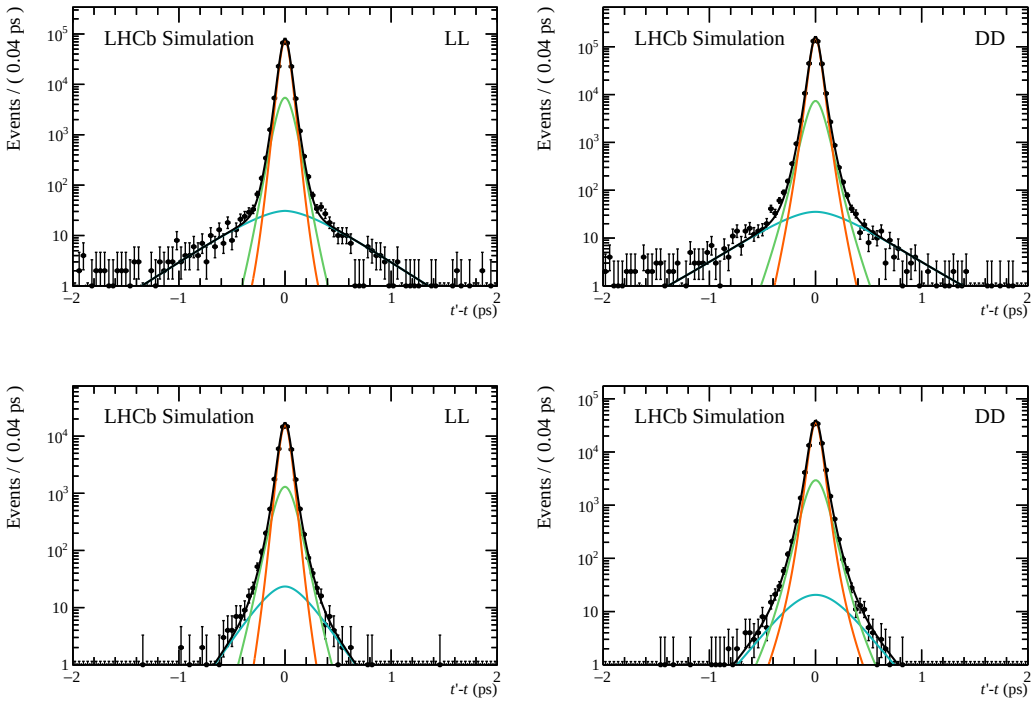

Figure A.15: Per-candidate fit to the decay time resolution on signal simulation for LL (left) and DD (right) for (top) $B^0 \to \psi(2S)K^0_S$ and (bottom) $B^0 \to J/\psi(\to ee)K^0_S$ candidates with the sum of the three Gaussians (black) and the single Gaussians (coloured).

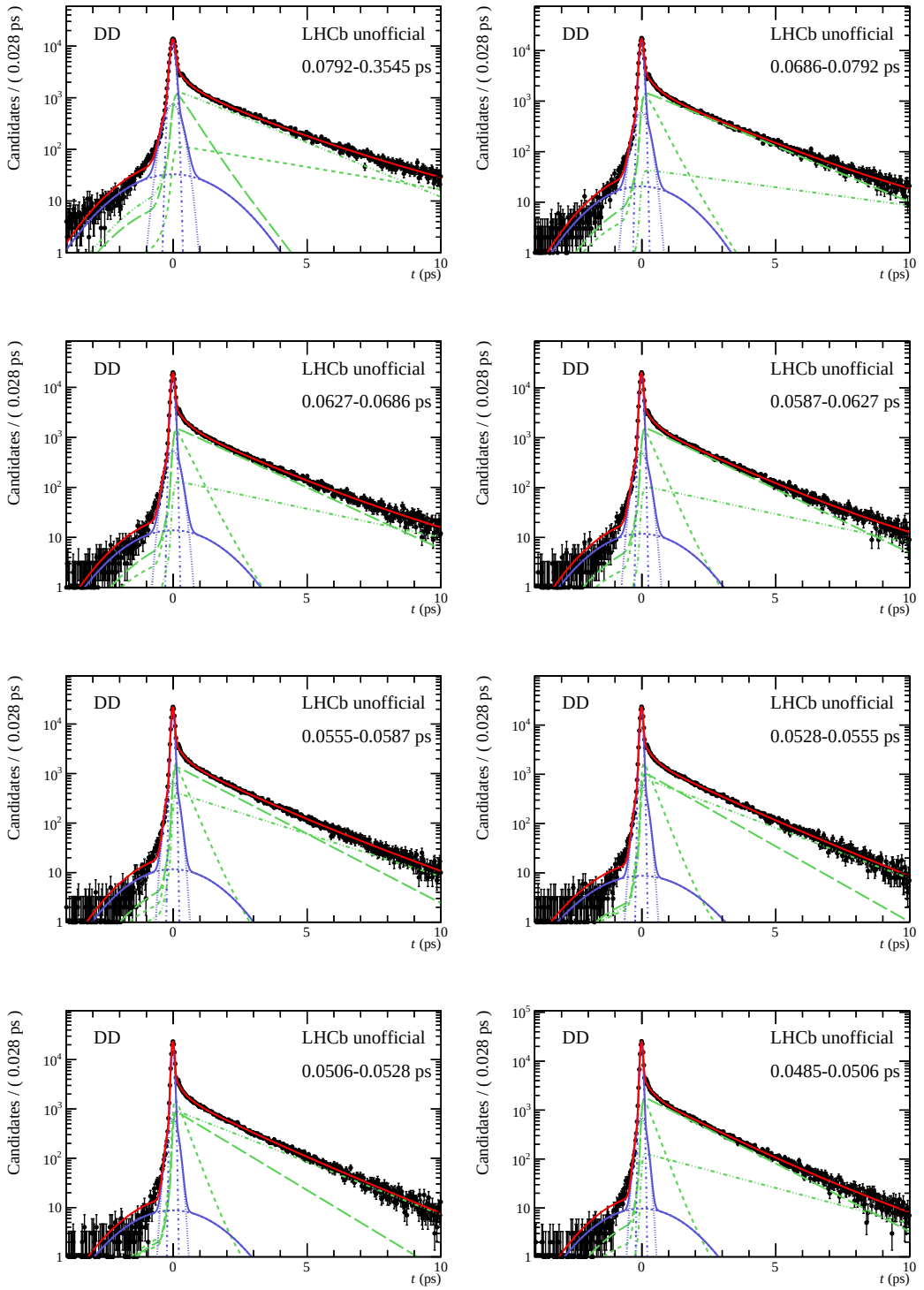

Figure A.16: Fit to the mean decay time resolution on prompt DD J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

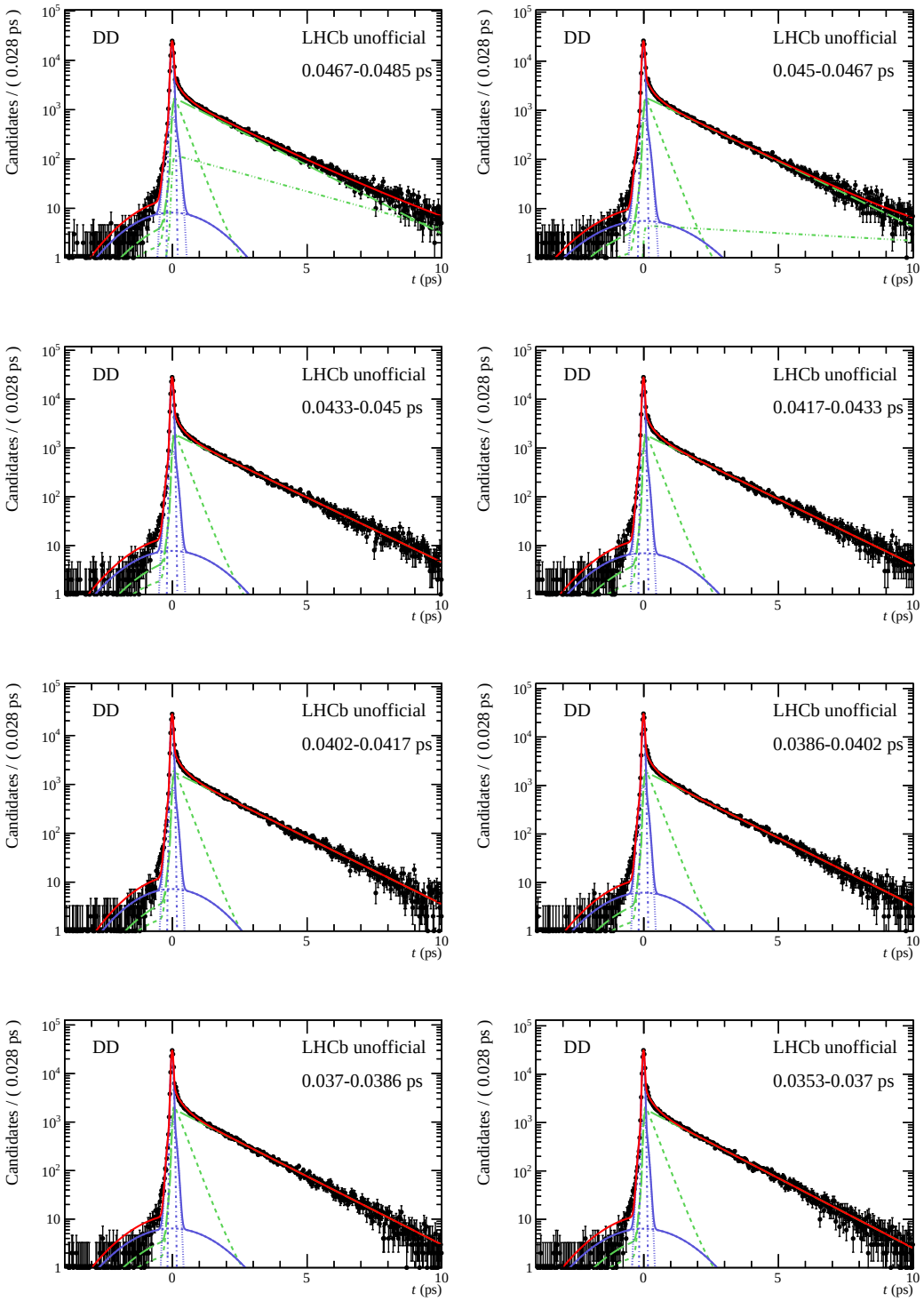

Figure A.17: Fit to the mean decay time resolution on prompt DD J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

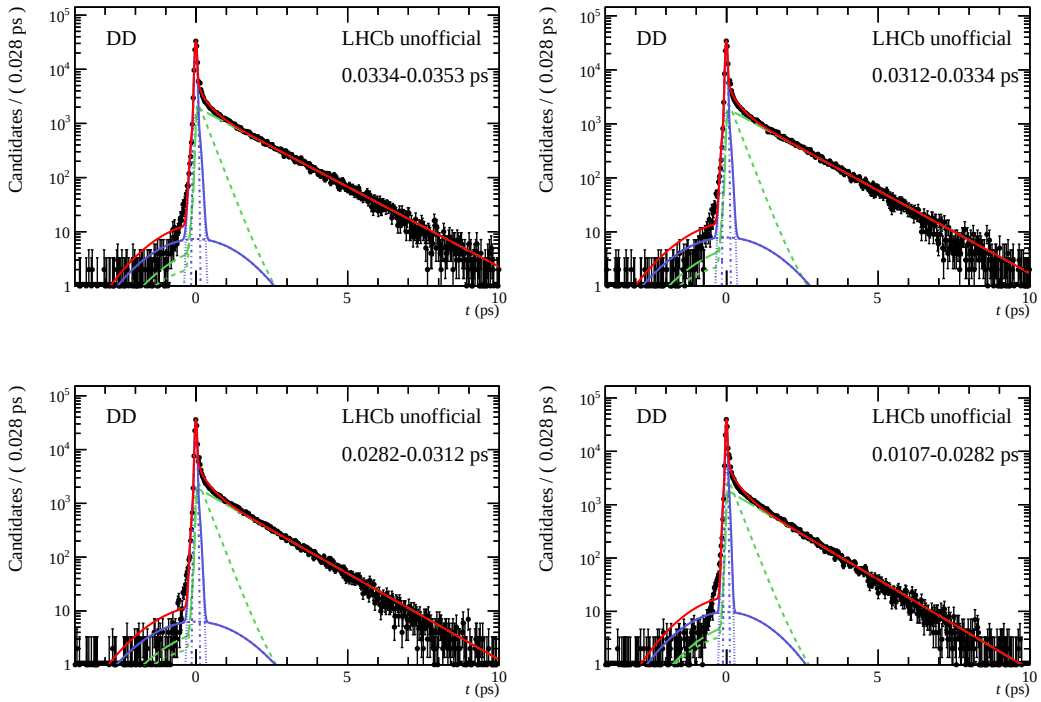

Figure A.18: Fit to the mean decay time resolution on prompt DD J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

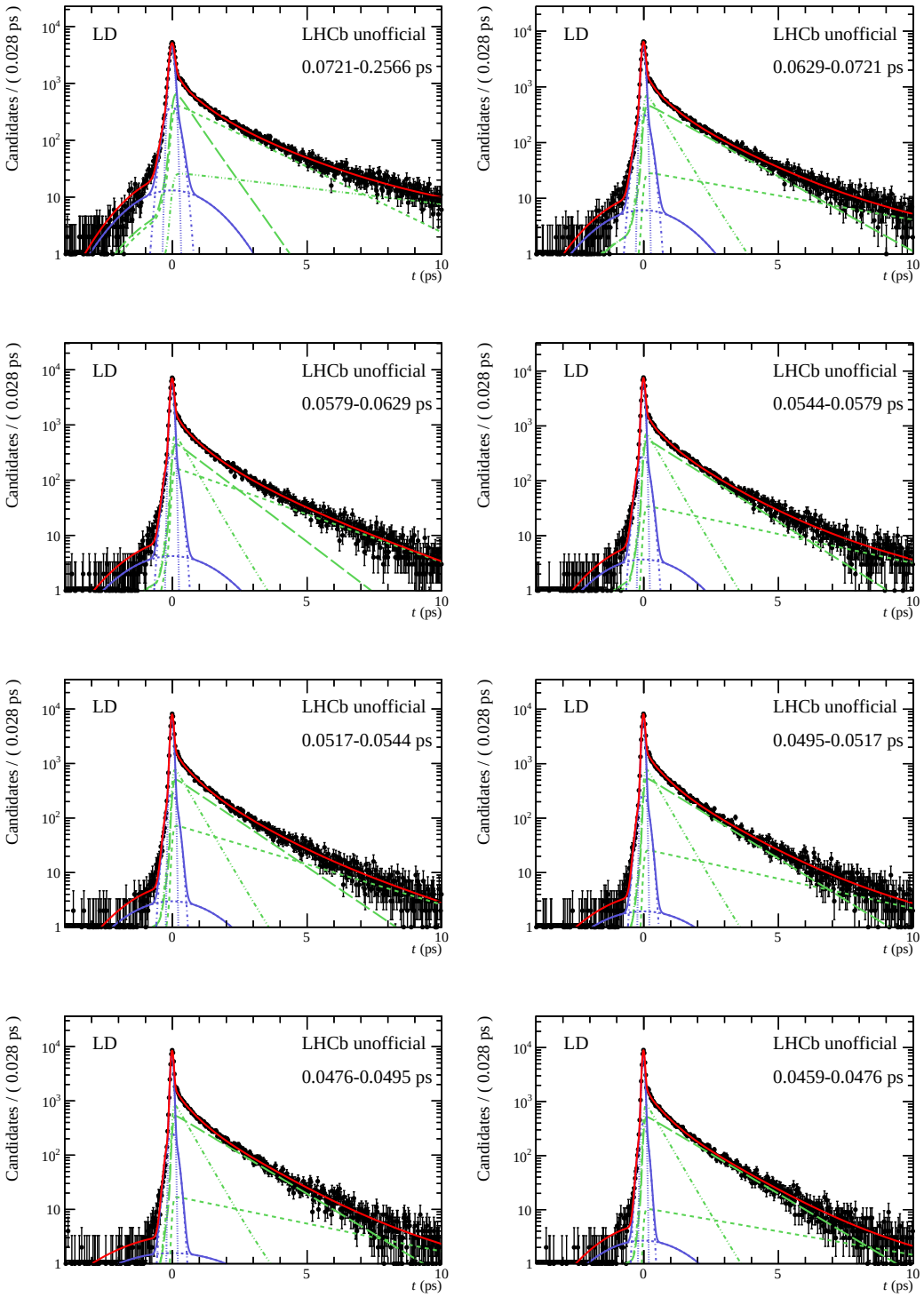

Figure A.19: Fit to the mean decay time resolution on prompt LD J/ψ data
132 in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

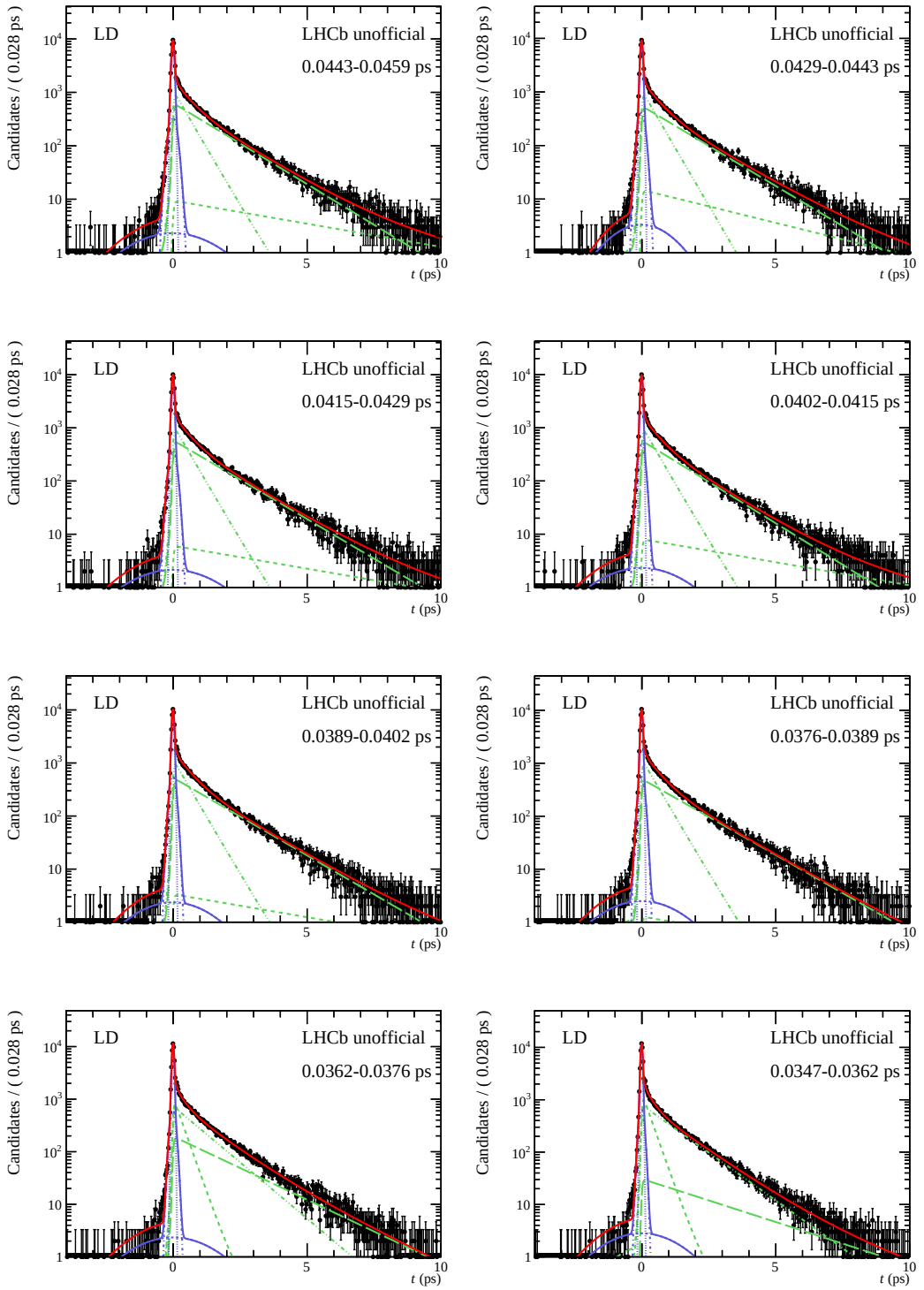

Figure A.20: Fit to the mean decay time resolution on prompt LD J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

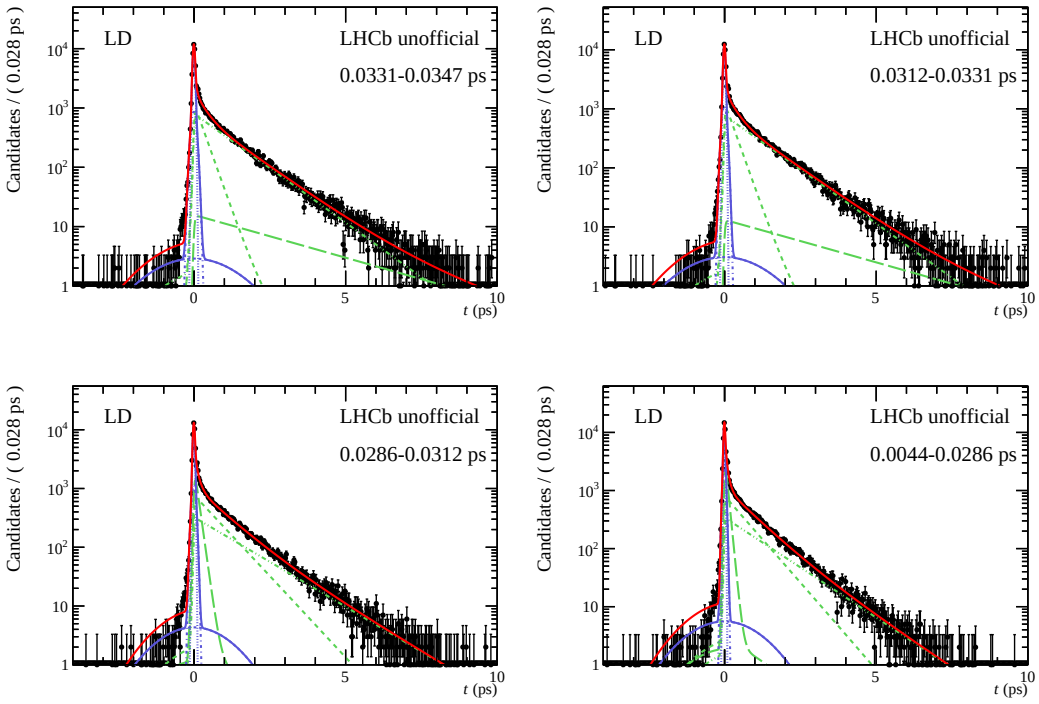

Figure A.21: Fit to the mean decay time resolution on prompt LD J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

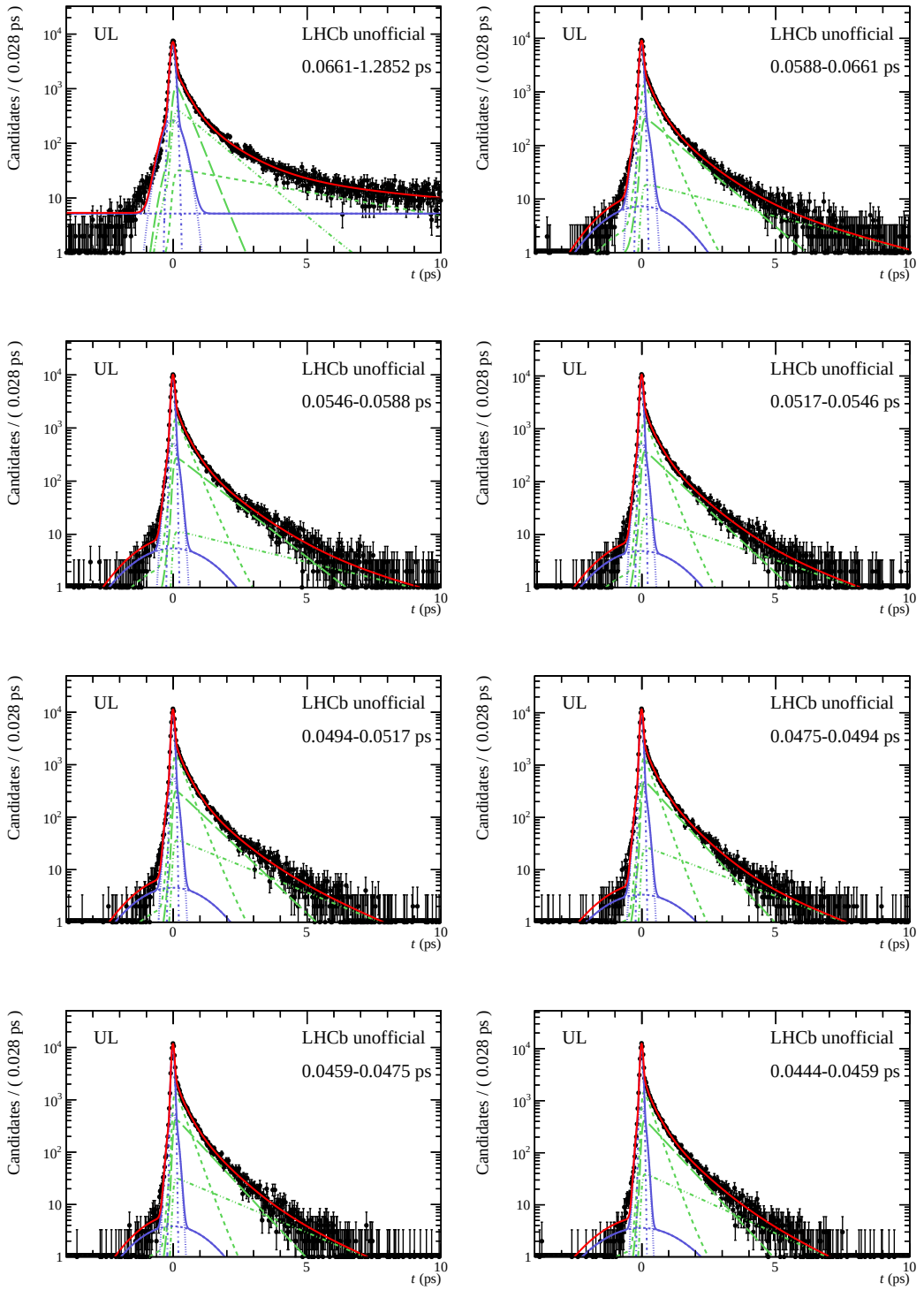

Figure A.22: Fit to the mean decay time resolution on prompt UL J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

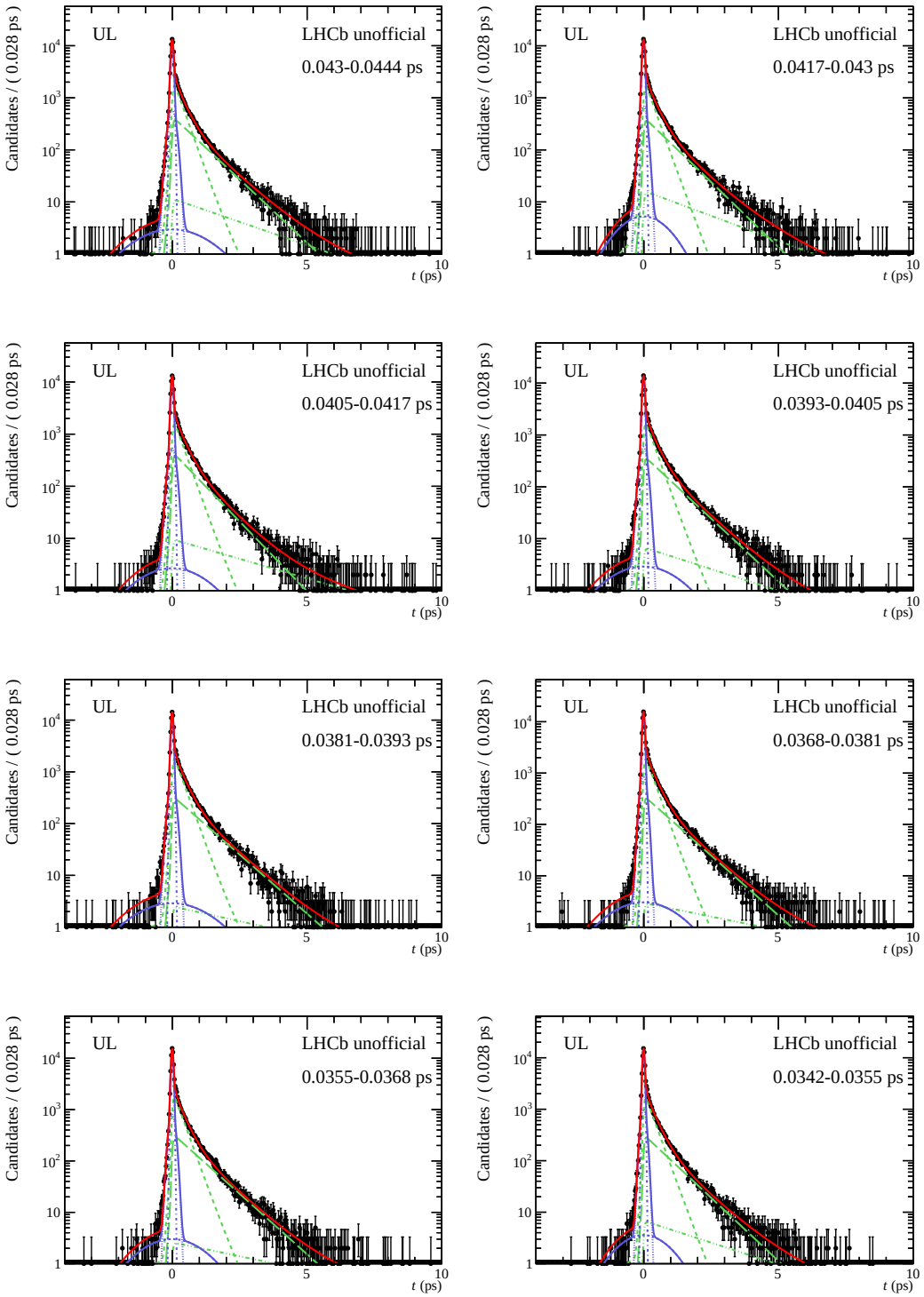

Figure A.23: Fit to the mean decay time resolution on prompt UL J/ψ data
136 in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

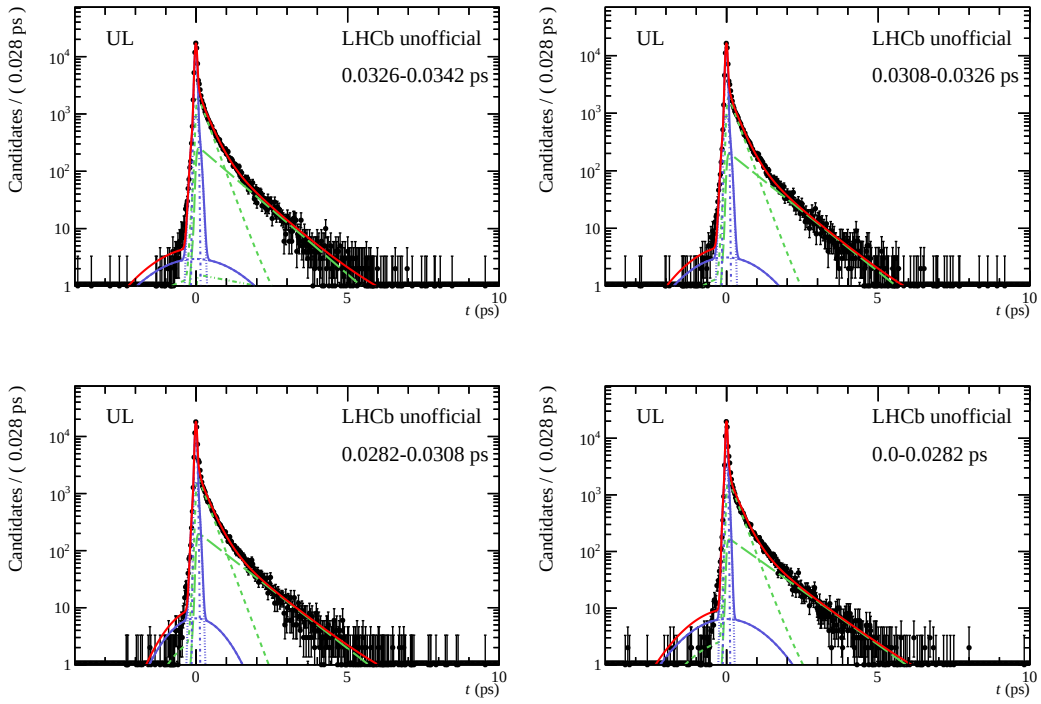

Figure A.24: Fit to the mean decay time resolution on prompt UL J/ψ data in different bins of the decay-time resolution estimate with the sum of the three Gaussians (purple, solid) and the single Gaussians (purple, dashed) with three exponential functions (green, dashed) and the overall fit in red.

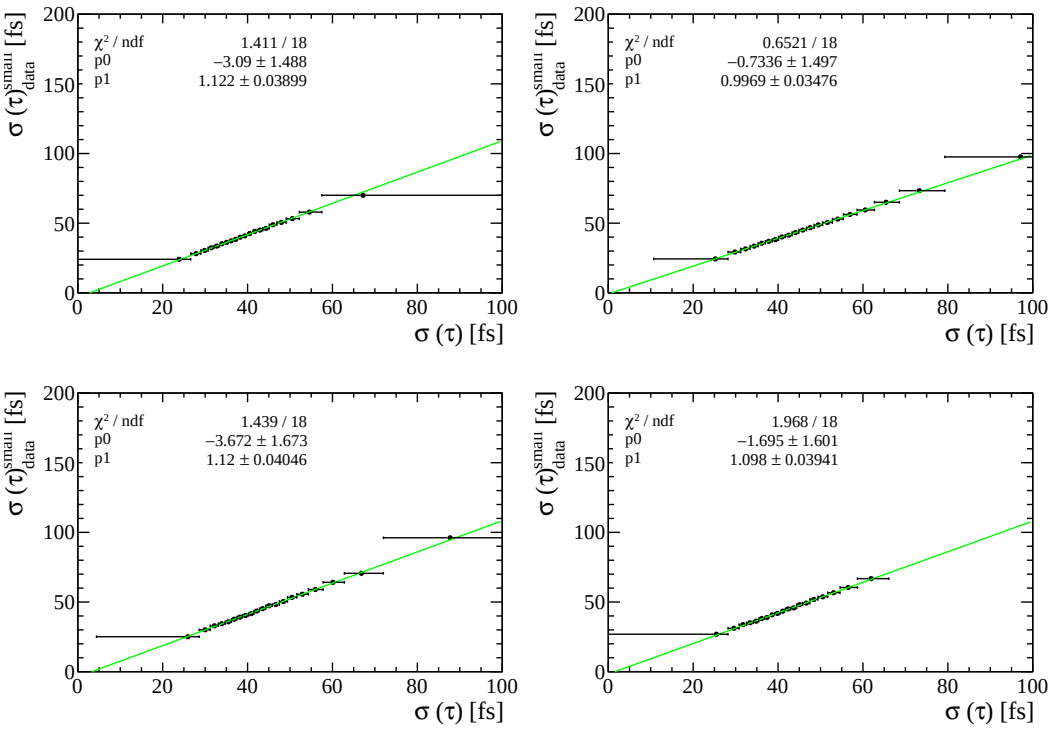

Figure A.25: The width of the smallest Gaussian for different bins of the decay-time resolution estimate for the different track types. With a linear fit to the data points.

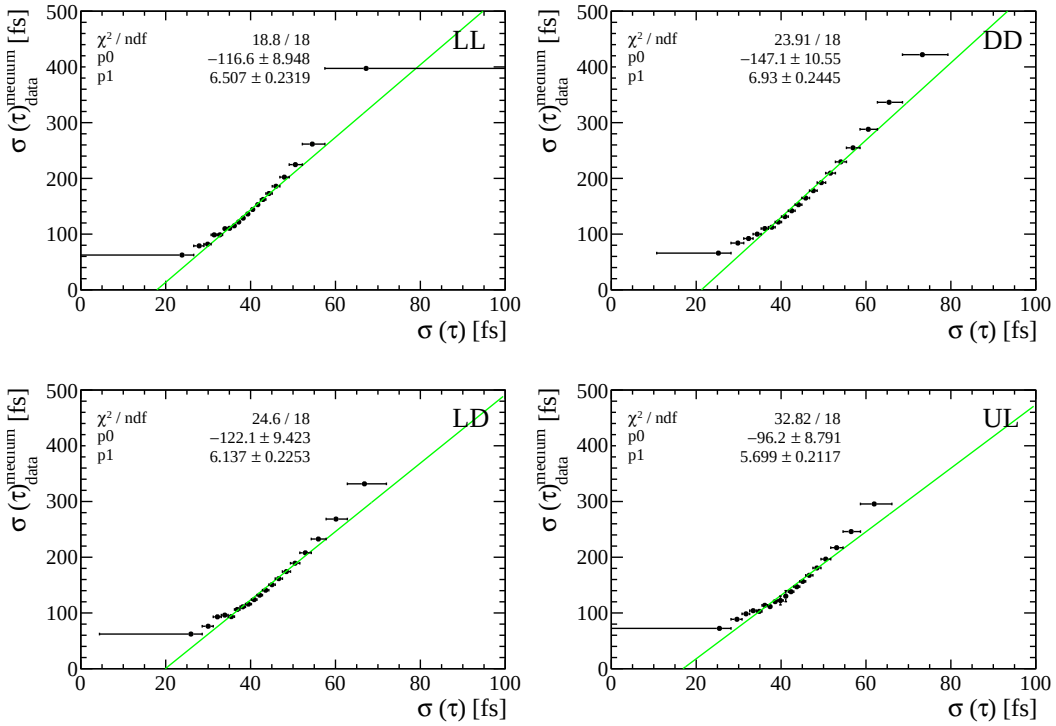

Figure A.26: The width of the medium Gaussian for different bins of the decay-time resolution estimate for the different track types. With a linear fit to the data points.

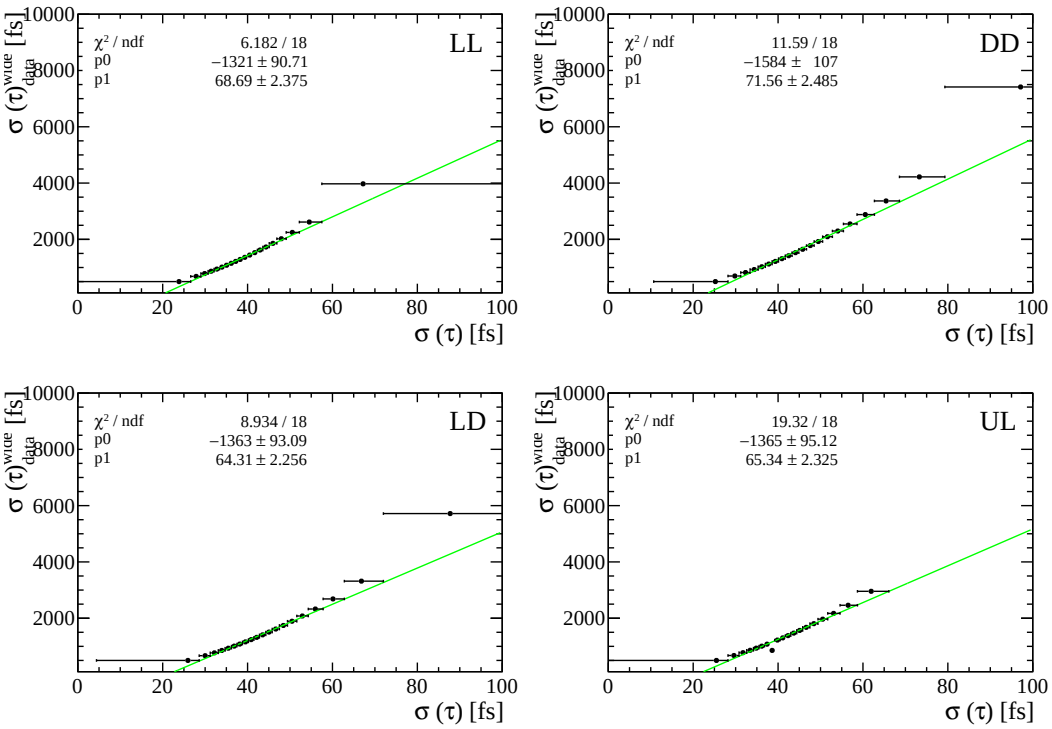

Figure A.27: The width of the wide Gaussian for different bins of the decay-time resolution estimate for the different track types. With a linear fit to the data points.

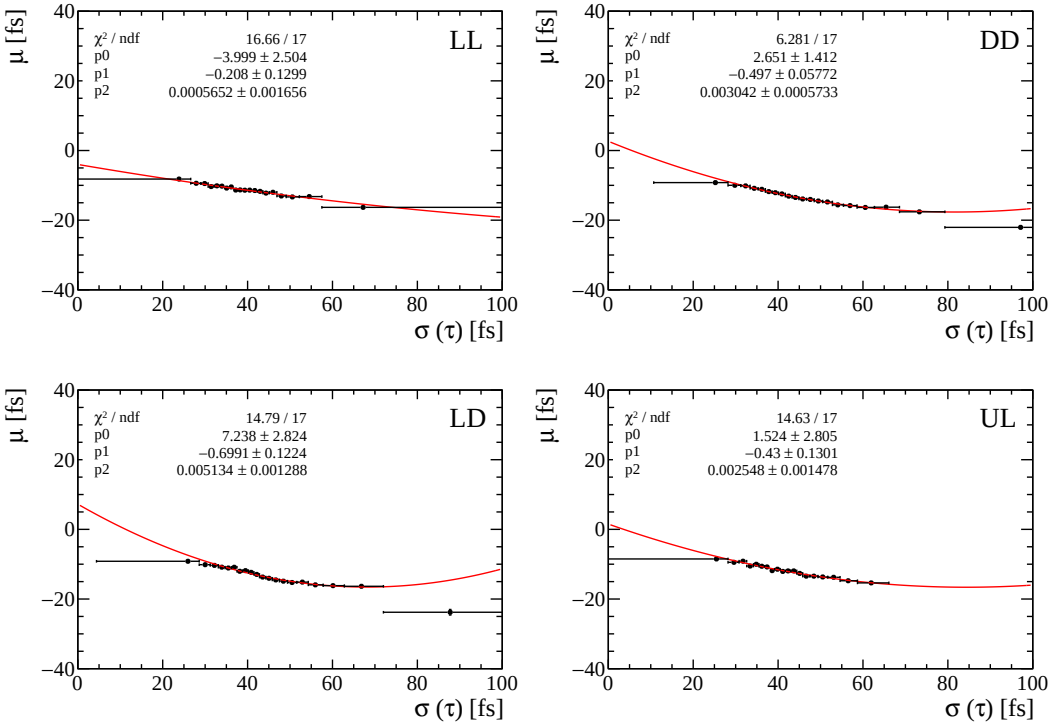

Figure A.28: The bias of the Gaussians for different bins of the decay-time resolution estimate for the different track types. With a quadratic fit to the data points.

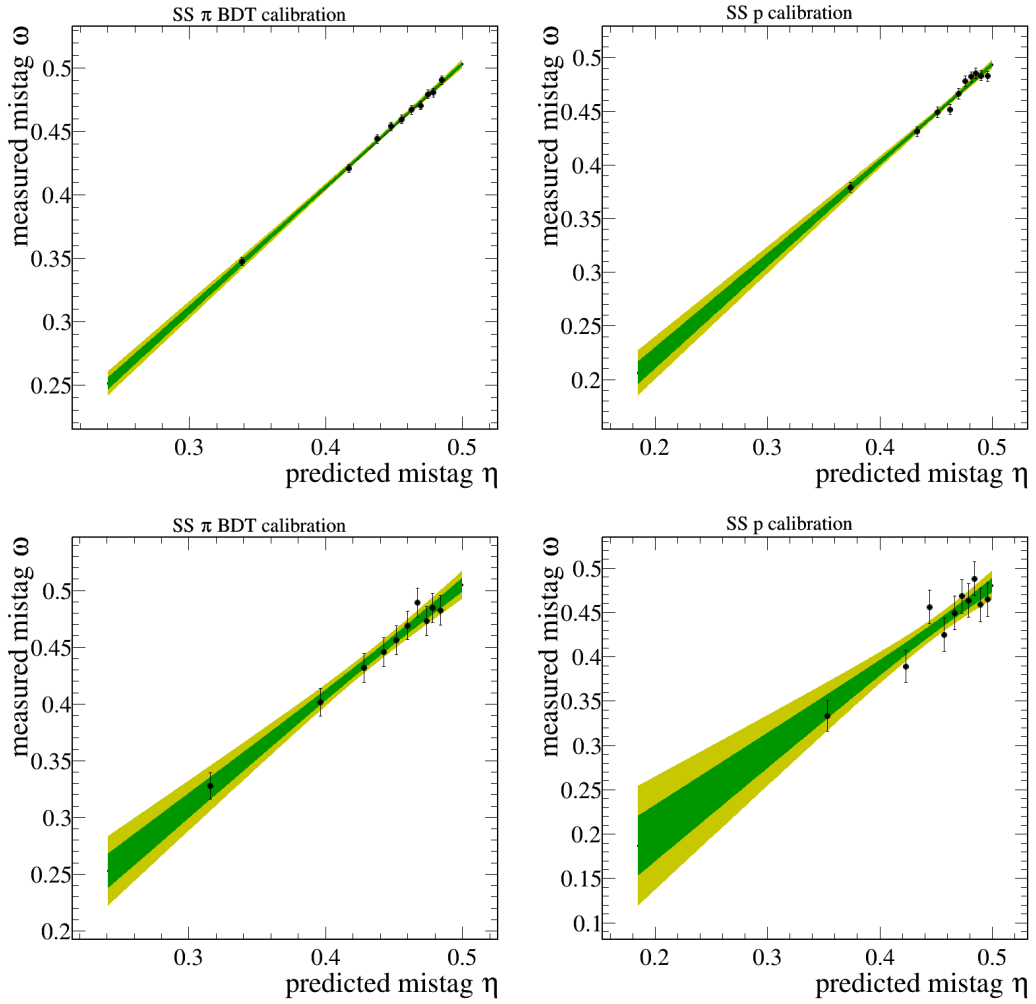

Figure A.29: Flavour tagging calibration for (top) $B^0 \to \psi(2S)K^0_S$ and (bottom) $B^0 \to J/\psi(\to ee)K^0_S$ of the SS single tagger (left) Pion and (right) Proton.

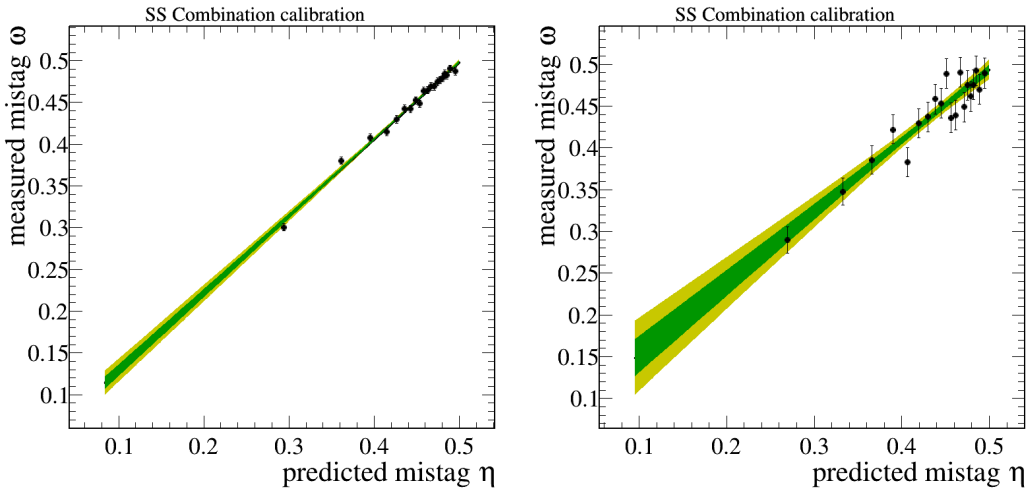

Figure A.30: Flavour tagging calibration for the calibration of all SS taggers combines for (left) $B^0 \to \psi(2S)K^0_S$ and (right) $B^0 \to J/\psi(\to ee)K^0_S$.

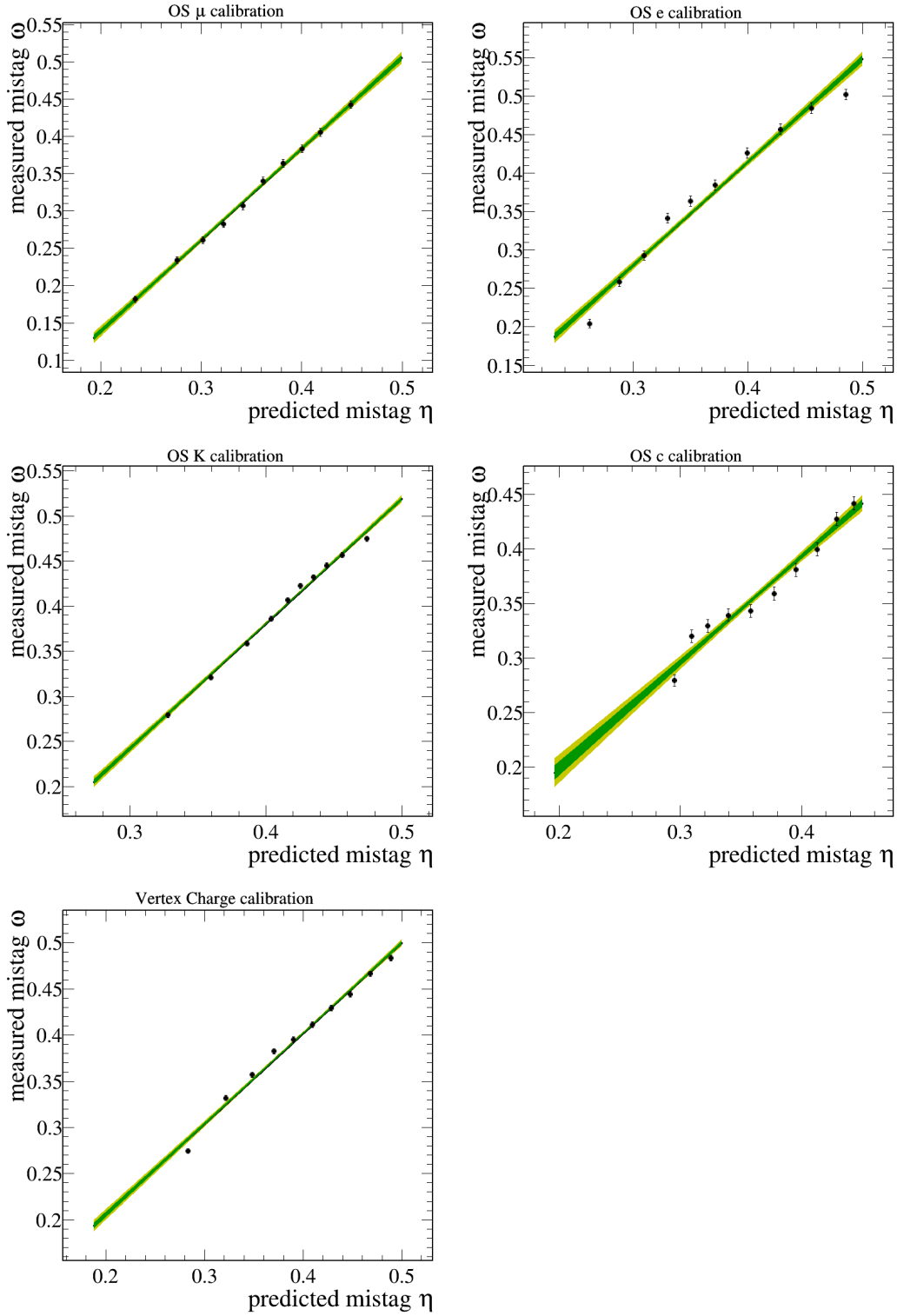

Figure A.31: Flavour tagging calibration for $B^0 \to \psi(2S)K^0_S$ of the OS single tagger (upper-left) Muon, (upper-right) Electron, (middle-left) Kaon, (middle-right) Charm and (lower-left) vertex charge.

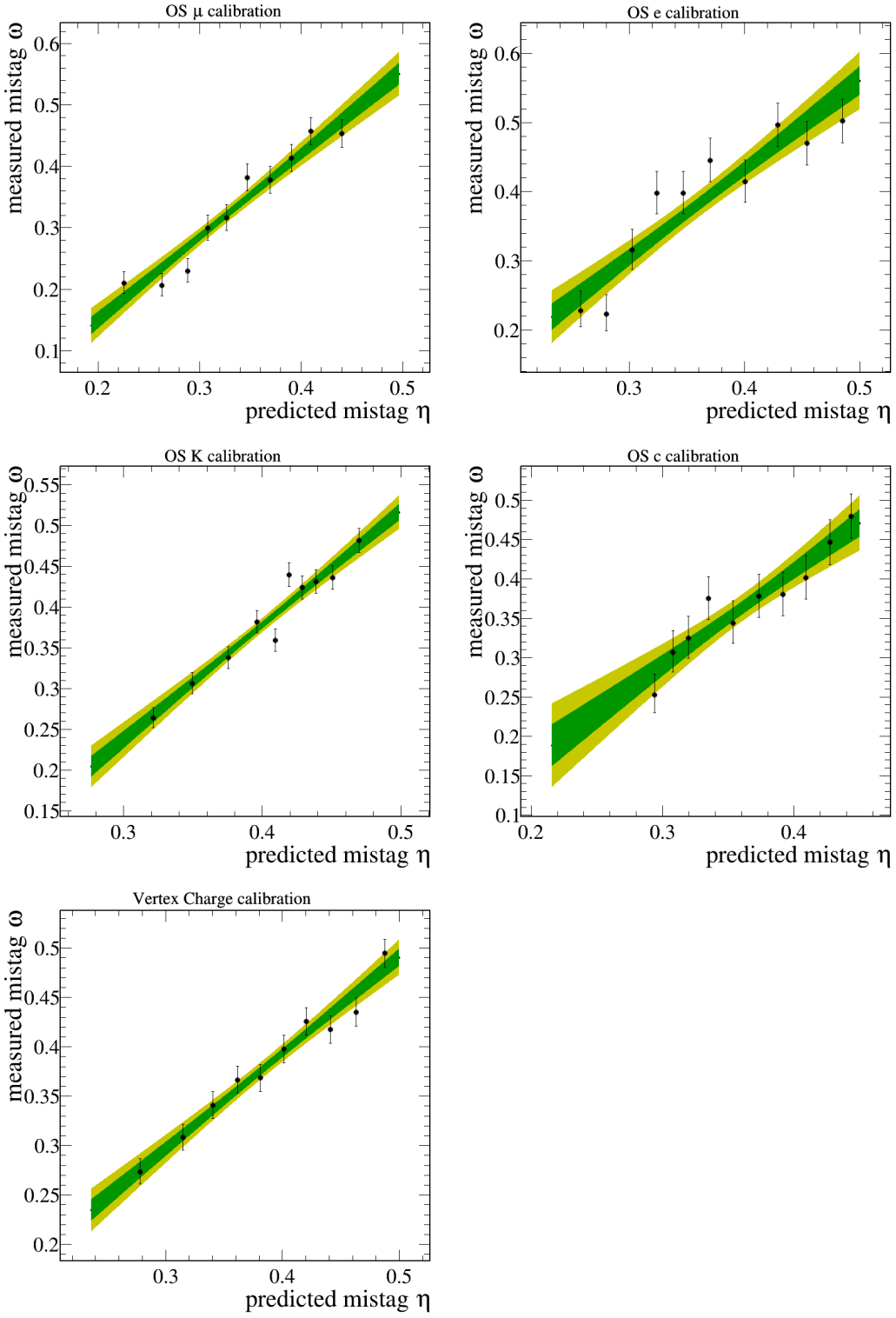

Figure A.32: Flavour tagging calibration for $B^0 \to J/\psi(\to ee)K_S^0$ of the OS single tagger (upper-left) Muon, (upper-right) Electron, (middle-left) Kaon, (middle-right) Charm and (lower-left) vertex charge.

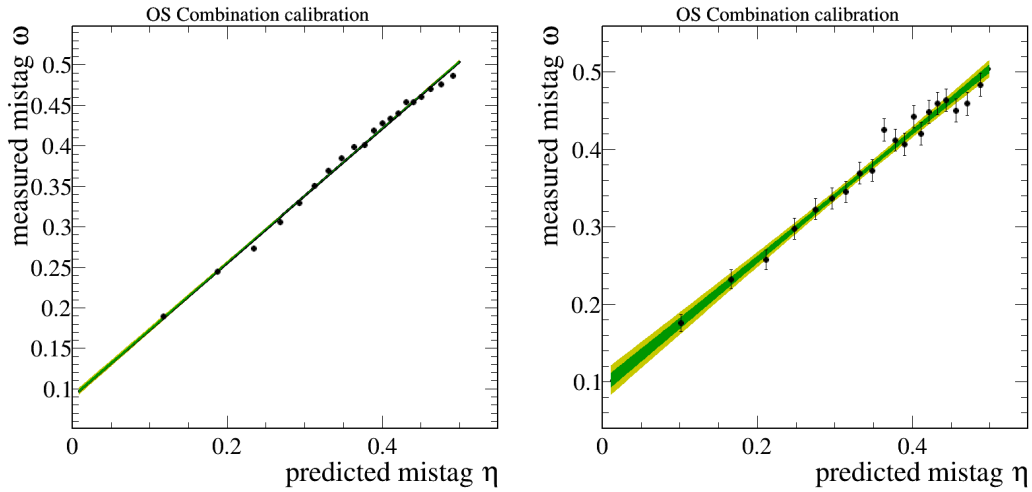

Figure A.33: Flavour tagging calibration for the calibration of all OS taggers combines for (left) $B^0 \to \psi(2S)K^0_S$ and (right) $B^0 \to J/\psi(\to ee)K^0_S$.

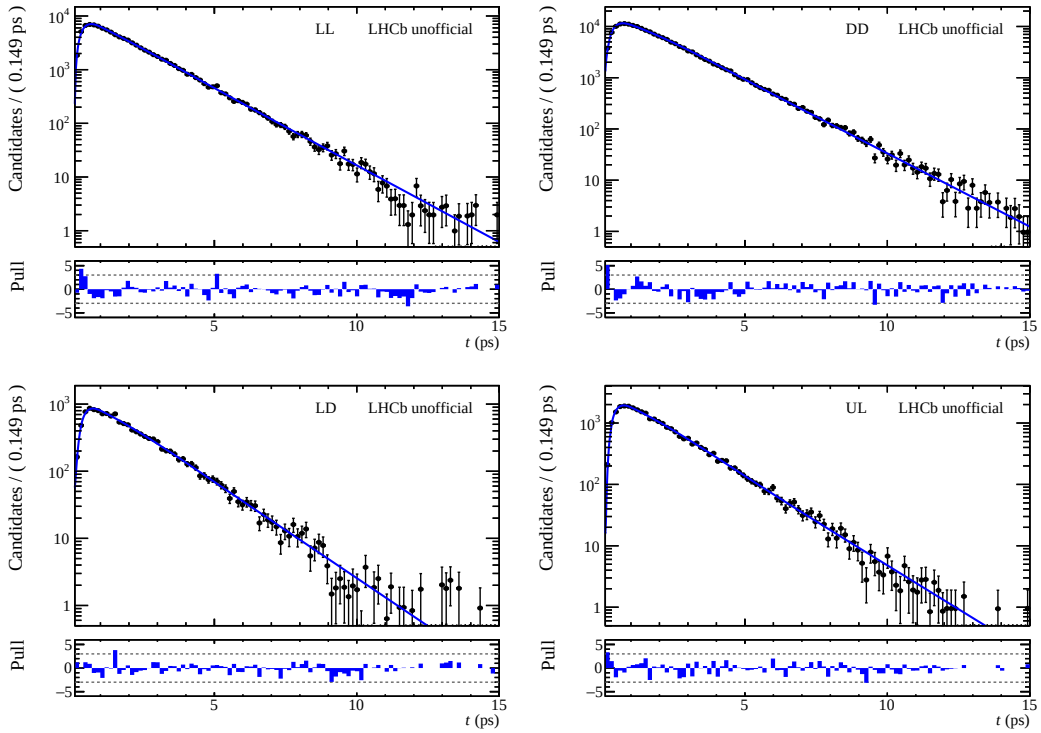

Figure A.34: Decay-time projection of the blinded *CP* fit in $B^0 \to J/\psi K^0_S$ decays for (upper-left) LL, (upper-right) DD, (lower-left) LD, and (lower-right) UL reconstructed candidates.

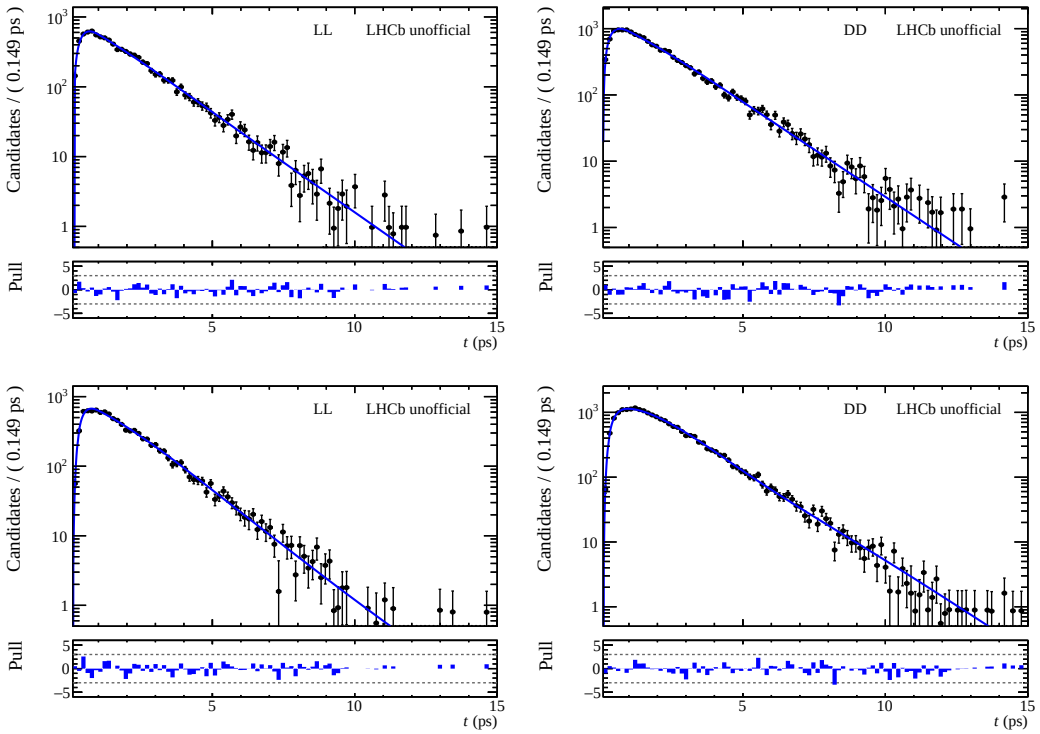

Figure A.35: Decay-time projection of the blinded *CP* fit in (top) $B^0 \to \psi(2S)K^0_S$ and (bottom) $B^0 \to J/\psi(e^+e^-)K^0_S$ decays for (left) LL and (right) DD reconstructed candidates.

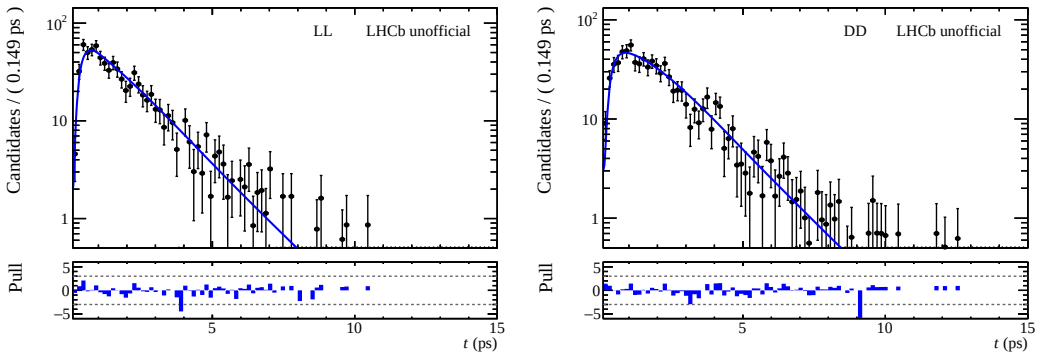

Figure A.36: Decay-time projection of the blinded *CP* fit in $B^0_s \to J/\psi\, K^0_S$ decays for (left) LL and (right) DD reconstructed candidates.

Bibliography

[1] S. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[2] A. Salam and J. C. Ward. "Electromagnetic and weak interactions". In: *Phys. Lett.* 13 (1964), pp. 168–171. DOI: 10.1016/0031-9163(64)90711-5.

[3] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[4] G. Aad et al. "Measurement of the branching ratio $\Gamma(\Lambda_b^0 \to \psi(2S)\Lambda^0)/\Gamma(\Lambda_b^0 \to J/\psi\Lambda^0)$ with the ATLAS detector". In: *Phys. Lett.* B751 (2015), pp. 63–80. DOI: 10.1016/j.physletb.2015.10.009. arXiv: 1507.08202 [hep-ex].

[5] T. Gutsche et al. "Polarization effects in the cascade decay $\Lambda_b^0 \to (\to p\pi^-) + J/\psi(\to \ell^+\ell^-)$ in the covariant confined quark model". In: *Phys. Rev.* D88 (2013), p. 114018. DOI: 10.1103/PhysRevD.88.114018. arXiv: 1309.7879 [hep-ph].

[6] T. Gutsche et al. "Towards an assessment of the ATLAS data on the branching ratio $\Gamma(\Lambda_b^0 \to \psi(2S)\Lambda)/\Gamma(\Lambda_b^0 \to J/\psi\Lambda)$". In: *Phys. Rev.* D92 (2015), p. 114008. DOI: 10.1103/PhysRevD.92.114008. arXiv: 1510.02266 [hep-ph].

[7] R. Aaij et al. "Measurement of the ratio of branching fractions of the decays $\Lambda_b^0 \to \psi(2S)\Lambda$ and $\Lambda_b^0 \to J/\psi\Lambda$". In: *JHEP* 03 (2019), p. 126. DOI: 10.1007/JHEP03(2019)126. arXiv: 1902.02092 [hep-ex].

[8] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531.

[9] M. Kobayashi and T. Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[10] B. Aubert et al. "Observation of CP violation in the B^0 meson system". In: *Phys. Rev. Lett.* 87 (2001), p. 091801. DOI: 10.1103/PhysRevLett.87.091801. arXiv: hep-ex/0107013.

[11] K. Abe et al. "Observation of large CP violation in the neutral B meson system". In: *Phys. Rev. Lett.* 87 (2001), p. 091802. DOI: 10.1103/PhysRevLett.87.091802. arXiv: hep-ex/0107061.

[12] Y. S. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018". In: (2019). updated results and plots available at https://hflav.web.cern.ch/. arXiv: 1909.12524 [hep-ex].

[13] R. Aaij et al. "Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays". In: *Phys. Rev. Lett.* 122 (2019), p. 191801. DOI: 10.1103/PhysRevLett.122.191801. arXiv: 1903.09252 [hep-ex].

[14] R. Aaij et al. "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays". In: *JHEP* 08 (2017), p. 055. DOI: 10.1007/JHEP08(2017)055. arXiv: 1705.05802 [hep-ex].

[15] N. Aghanim et al. "Planck 2018 results. VI. Cosmological parameters". In: *Astron. Astrophys.* 641 (2020), A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[16] P. A. Zyla et al. "Review of particle physics". In: *Prog. Theor. Exp. Phys.* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[17] R. Aaij et al. "Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays". In: *Phys. Rev. Lett.* 118 (2017), p. 022003. DOI: 10.1103/PhysRevLett.118.022003. arXiv: 1606.07895 [hep-ex].

[18] R. Aaij et al. "Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays". In: *Phys. Rev.* D95 (2017), p. 012002. DOI: 10.1103/PhysRevD.95.012002. arXiv: 1606.07898 [hep-ex].

[19] R. Aaij et al. "Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays". In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. arXiv: 1507.03414 [hep-ex].

[20] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[21] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[22] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[23] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. Taylor, pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[24] P. A. Dirac. "Quantum theory of emission and absorption of radiation". In: *Proc. Roy. Soc. Lond. A* 114 (1927), p. 243. DOI: 10.1098/rspa.1927.0039.

[25] A. I. Vainshtein, V. I. Zakharov, and M. A. Shifman. "A Possible mechanism for the Delta T = 1/2 rule in nonleptonic decays of strange particles". In: *JETP Lett.* 22 (1975), pp. 55–56.

[26] J. Ellis et al. "The phenomenology of the next left-handed quarks". In: *Nuclear Physics B* 131.2 (1977), pp. 285–307. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(77)90374-1. URL: https://www.sciencedirect.com/science/article/pii/0550321377903741.

[27] B. B. Brandt et al. "Form factors in lattice QCD". In: *Eur. Phys. J. ST* 198 (2011), pp. 79–94. DOI: 10.1140/epjst/e2011-01484-1. arXiv: 1106.1554 [hep-lat].

[28] D. Melikhov. "Advances in QCD sum rule calculations". In: *AIP Conf. Proc.* 1701.1 (2016). Ed. by A. Andrianov et al., p. 020014. DOI: 10.1063/1.4938603. arXiv: 1501.06319 [hep-ph].

[29] H. David Politzer and M. B. Wise. "Effective field theory approach to processes involving both light and heavy fields". In: *Physics Letters B* 208.3 (1988), pp. 504–507. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(88)90656-9. URL: https://www.sciencedirect.com/science/article/pii/0370269388906569.

[30] C. W. Bauer et al. "An effective field theory for collinear and soft gluons: heavy to light decays". In: *Phys. Rev. D* 63 (2001), p. 114020. DOI: 10.1103/PhysRevD.63.114020. arXiv: hep-ph/0011336.

[31] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 51 (1983), p. 1945. DOI: 10.1103/PhysRevLett.51.1945.

[32] J. Charles et al. "Current status of the standard model CKM fit and constraints on $\Delta F = 2$ new physics". In: *Phys. Rev.* D91 (2015). updated results and plots available at http://ckmfitter.in2p3.fr/, p. 073007. DOI: 10.1103/PhysRevD.91.073007. arXiv: 1501.05013 [hep-ph].

[33] A. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Sov. Phys. Usp.* 34.5 (1991), pp. 392–393. DOI: 10.1070/PU1991v034n05ABEH002497.

[34] T. Inami and C. Lim. "Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_{\mathrm{L}}^0 \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$". In: *Prog. Theor. Phys.* 65 (1981). [Erratum: Prog.Theor.Phys. 65, 1772 (1981)], p. 297. DOI: 10.1143/PTP.65.297.

[35] E. Schrödinger. "Quantisierung als Eigenwertproblem". In: *Annalen Phys.* 384.4 (1926), pp. 361–376. DOI: 10.1002/andp.19263840404.

[36] R. Aaij et al. "Measurement of the *CP* asymmetry in B_s^0–$\overline{B}_s^0$ mixing". In: *Phys. Rev. Lett.* 117 (2016), p. 061803. DOI: 10.1103/PhysRevLett.117.061803. arXiv: 1605.09768 [hep-ex].

[37] T. Jubb et al. "On the ultimate precision of meson mixing observables". In: *Nucl. Phys. B* 915 (2017), pp. 431–453. DOI: 10.1016/j.nuclphysb.2016.12.020. arXiv: 1603.07770 [hep-ph].

[38] R. Aaij et al. "Measurement of the semileptonic *CP* asymmetry in B^0–$\overline{B}^0$ mixing". In: *Phys. Rev. Lett.* 114 (2015), p. 041601. DOI: 10.1103/PhysRevLett.114.041601. arXiv: 1409.8586 [hep-ex].

[39] Y. S. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018". In: (Sept. 2019). arXiv: 1909.12524 [hep-ex].

[40] K. De Bruyn and R. Fleischer. "A Roadmap to Control Penguin Effects in $B_d^0 \to J/\psi K_S^0$ and $B_s^0 \to J/\psi\phi$". In: *JHEP* 03 (2015), p. 145. DOI: 10.1007/JHEP03(2015)145. arXiv: 1412.6834 [hep-ph].

[41] M. Z. Barel et al. "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". In: (Oct. 2020). arXiv: 2010.14423 [hep-ph].

[42] M. Jung. "Determining weak phases from $B \to J/\psi P$ decays". In: *Phys. Rev. D* 86 (2012), p. 053008. DOI: 10.1103/PhysRevD.86.053008. arXiv: 1206.2050 [hep-ph].

[43] P. Frings, U. Nierste, and M. Wiebusch. "Penguin contributions to *CP* phases in $B_{d,s}$ decays to charmonium". In: *Phys. Rev. Lett.* 115.6 (2015), p. 061802. DOI: 10.1103/PhysRevLett.115.061802. arXiv: 1503.00859 [hep-ph].

[44] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[45] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[46] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[47] A. A. Alves Jr. et al. "The LHCb detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[48] E. Mobs. "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019". In: (July 2019). General Photo. URL: https://cds.cern.ch/record/2684277.

[49] R. Aaij et al. "LHCb detector performance". In: *Int. J. Mod. Phys.* A30 (2015), p. 1530022. DOI: 10.1142/S0217751X15300227. arXiv: 1412.6352 [hep-ex].

[50] R. Aaij et al. "Performance of the LHCb Vertex Locator". In: *JINST* 9 (2014), P09007. DOI: 10.1088/1748-0221/9/09/P09007. arXiv: 1405.7808 [physics.ins-det].

[51] M. Kucharczyk, P. Morawski, and M. Witek. "Primary Vertex Reconstruction at LHCb". Geneva, Sept. 2014. URL: http://cds.cern.ch/record/1756296.

[52] C. Abellan Beteta et al. "Monitoring radiation damage in the LHCb Tracker Turicensis". In: (2018). arXiv: 1809.05063 [physics.ins-det].

[53] R. Arink et al. "Performance of the LHCb Outer Tracker". In: *JINST* 9 (2014), P01002. DOI: 10.1088/1748-0221/9/01/P01002. arXiv: 1311.3893 [physics.ins-det].

[54] P. d'Argent et al. "Improved performance of the LHCb Outer Tracker in LHC Run 2". In: *JINST* 9 (2017), P11016. DOI: 10.1088/1748-0221/12/11/P11016. arXiv: 1708.00819 [physics.ins-det].

[55] A. A. Alves Jr. et al. "Performance of the LHCb muon system". In: *JINST* 8 (2013), P02022. DOI: 10.1088/1748-0221/8/02/P02022. arXiv: 1211.1346 [physics.ins-det].

[56] R. Aaij et al. "The LHCb trigger and its performance in 2011". In: *JINST* 8 (2013), P04022. DOI: 10.1088/1748-0221/8/04/P04022. arXiv: 1211.3055 [hep-ex].

[57] M. Adinolfi et al. "Performance of the LHCb RICH detector at the LHC". In: *Eur. Phys. J.* C73 (2013), p. 2431. DOI: 10.1140/epjc/s10052-013-2431-9. arXiv: 1211.6759 [physics.ins-det].

[58] P. Cherenkov. "Visible luminescence of pure liquids under the influence of γ-radiation". In: *Dokl. Akad. Nauk SSSR* 2.8 (1934), pp. 451–454. DOI: 10.3367/UFNr.0093.196710n.0385.

[59] R. Aaij et al. "Allen: A high level trigger on GPUs for LHCb". In: *Comput. Softw. Big Sci.* 4.1 (2020), p. 7. DOI: 10.1007/s41781-020-00039-7. arXiv: 1912.09161 [physics.ins-det].

[60] T. Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[61] I. Belyaev et al. "Handling of the generation of primary events in Gauss, the LHCb simulation framework". In: *J. Phys. Conf. Ser.* 331 (2011), p. 032047. DOI: 10.1088/1742-6596/331/3/032047.

[62] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[63] P. Golonka and Z. Was. "PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays". In: *Eur.Phys.J.* C45 (2006), pp. 97–107. DOI: 10.1140/epjc/s2005-02396-4. arXiv: hep-ph/0506026 [hep-ph].

[64] J. Allison et al. "Geant4 developments and applications". In: *IEEE Trans.Nucl.Sci.* 53 (2006), p. 270. DOI: 10.1109/TNS.2006.869826.

[65] S. Agostinelli et al. "Geant4: A simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[66] M. Clemencic et al. "The LHCb simulation application, Gauss: Design, evolution and experience". In: *J. Phys. Conf. Ser.* 331 (2011), p. 032023. DOI: 10.1088/1742-6596/331/3/032023.

[67] V. Blobel and E. Lohrmann. "Statistische und numerische Methoden der Datenanalyse". Teubner Studienbücher Physik. Stuttgart: Teubner, 1998. URL: https://cds.cern.ch/record/437773.

[68] M. Pivk and F. R. Le Diberder. "SPlot: A Statistical tool to unfold data distributions". In: *Nucl. Instrum. Meth.* A555 (2005), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106. arXiv: physics/0402083 [physics.data-an].

[69] Y. Xie. "sFit: a method for background subtraction in maximum likelihood fit". In: (2009). arXiv: 0905.0724 [physics.data-an].

[70] C. Langenbruch. "Parameter uncertainties in weighted unbinned maximum likelihood fits". In: (Nov. 2019). arXiv: 1911.01303 [physics.data-an].

[71] P. Mackowiak. "Measurement of the relative branching ratio $\mathcal{B}(\Lambda_b^0 \to \psi(2S)\Lambda)$". Master's thesis, Technische Universität Dortmund. 2016.

[72] M. Tanabashi et al. "Review of particle physics". In: *Phys. Rev.* D98 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[73] W. D. Hulsbergen. "Decay chain fitting with a Kalman filter". In: *Nucl. Instrum. Meth.* A552 (2005), pp. 566–575. DOI: 10.1016/j.nima.2005.06.078. arXiv: physics/0503191 [physics].

[74] R. Aaij et al. "Measurements of the $\Lambda_b^0 \to J/\psi\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s}$ =7 TeV". In: *Phys. Lett.* B724 (2013), p. 27. DOI: 10.1016/j.physletb.2013.05.041. arXiv: 1302.5578 [hep-ex].

[75] D. Martinez Santos and F. Dupertuis. "Mass distributions marginalised over per-event errors". In: *Nucl. Instrum. Meth.* A764 (2014), pp. 150–155. DOI: 10.1016/j.nima.2014.06.081. arXiv: 1312.5000 [hep-ex].

[76] K. S. Cranmer. "Kernel estimation in high-energy physics". In: *Comput. Phys. Commun.* 136 (2001), pp. 198–207. DOI: 10.1016/S0010-4655(00)00243-5. arXiv: hep-ex/0011057 [hep-ex].

[77] T. Skwarnicki. "A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances". DESY-F31-86-02. PhD thesis. Institute of Nuclear Physics, Krakow, 1986.

[78] S. Tolk et al. *Data driven trigger efficiency determination at LHCb*. Tech. rep. CERN-LHCb-PUB-2014-039. Geneva: CERN, May 2014. URL: https://cds.cern.ch/record/1701134.

[79] A. Poluektov. "Kernel density estimation of a multidimensional efficiency profile". In: *JINST* 10.02 (2015), P02011. DOI: 10.1088/1748-0221/10/02/P02011. arXiv: 1411.5528 [physics.data-an].

[80] R. Aaij et al. "Measurement of the time-dependent *CP* asymmetry in $B^0 \to J/\psi K^0_S$ decays". In: *Phys. Lett.* B721 (2013), p. 24. DOI: 10.1016/j.physletb.2013.02.054. arXiv: 1211.6093 [hep-ex].

[81] R. Aaij et al. "Measurement of *CP* violation in $B^0 \to J/\psi K^0_S$ decays". In: *Phys. Rev. Lett.* 115 (2015), p. 031601. DOI: 10.1103/PhysRevLett.115.031601. arXiv: 1503.07089 [hep-ex].

[82] R. Aaij et al. "Measurement of *CP* violation in $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S)K^0_S$ decays". In: *JHEP* 11 (2017), p. 170. DOI: 10.1007/JHEP11(2017)170. arXiv: 1709.03944 [hep-ex].

[83] V. Jevtić. "Measurement of the *CP*-Violation Parameter $\sin(2\beta)$ in $B^0 \to J/\psi K^0_S$ Decays with Run II Data at LHCb". Master's thesis, Technische Universität Dortmund. 2018.

[84] J. H. Friedman. "Stochastic gradient boosting". In: *Comput. Stat. Data Anal.* 38 (2002), pp. 367–378. DOI: 10.1016/S0167-9473(01)00065-2.

[85] C. M. Bishop. "Pattern Recognition and Machine Learning (Information Science and Statistics)". Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 0387310738.

[86] G. Knott. "Interpolating Cubic Splines". Jan. 2000. DOI: 10.1007/978-1-4612-1320-8.

[87] T. M. Karbach, G. Raven, and M. Schiller. "Decay time integrals in neutral meson mixing and their efficient evaluation". In: (July 2014). arXiv: 1407.0748 [physics.data-an].

[88] R. Aaij et al. "Measurements of the B^+, B^0, B^0_s meson and Λ^0_b baryon lifetimes". In: *JHEP* 04 (2014), p. 114. DOI: 10.1007/JHEP04(2014)114. arXiv: 1402.2554 [hep-ex].

[89] M. Schellenberg. "Analyse der VELO-Rekonstruktionseffizienz im Zerfall $B^0 \to J/\psi K^0_S$ am LHCb-Experiment". B.Sc. thesis, Technische Universität Dortmund. 2013.

[90] D. Fazzini. "Flavour Tagging in the LHCb experiment". In: *Proceedings, 6th Large Hadron Collider Physics Conference (LHCP 2018): Bologna, Italy, June 4-9, 2018*. Vol. LHCP2018. 2018, p. 230. DOI: 10.22323/1.321.0230.

[91] R. Aaij et al. "Measurement of *CP* violation in $B^0 \to D^+D^-$ decays". In: *Phys. Rev. Lett.* 117 (2016), p. 261801. DOI: 10.1103/PhysRevLett.117.261801. arXiv: 1608.06620 [hep-ex].

[92] R. Aaij et al. "Updated measurement of time-dependent *CP*-violating observables in $B_s^0 \to J/\psi K^+K^-$ decays". In: *Eur. Phys. J.* C79 (2019), p. 706. DOI: 10.1140/epjc/s10052-019-7159-8. arXiv: 1906.08356 [hep-ex].

[93] R. E. Bellman. "Dynamic Programming". USA: Dover Publications, Inc., 2003. ISBN: 0486428095.

[94] A. Rogozhnikov. "Reweighting with Boosted Decision Trees". In: *J. Phys. Conf. Ser.* 762.1 (2016). https://github.com/arogozhnikov/hep_ml. DOI: 10.1088/1742-6596/762/1/012036. arXiv: 1608.05806 [physics.data-an].

[95] J. Wimberly. "EspressoPerformanceMonitor". https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor. 2020.

[96] R. Aaij et al. "Precise determination of the B_s^0–$\overline{B}_s^0$ oscillation frequency". in preparation.

[97] M. Chaichian and A. Fridman. "On a possibility for measuring effects of *CP* violation at pp colliders". In: *Phys. Lett. B* 298 (1993), pp. 218–223. DOI: 10.1016/0370-2693(93)91733-4.

[98] E. Norrbin and R. Vogt. "Bottom production asymmetries at the LHC". In: *5th Workshop on Electronics for the LHC Experiments (LEB 99)*. Mar. 2000. arXiv: hep-ph/0003056.

[99] R. Aaij et al. "Measurement of B^0, B_s^0, B^+ and Λ_b^0 production asymmetries in 7 and 8 TeV proton-proton collisions". In: *Phys. Lett.* B774 (2017), p. 139. DOI: 10.1016/j.physletb.2017.09.023. arXiv: 1703.08464 [hep-ex].

[100] R. Aaij et al. "Measurement of the time-dependent *CP* asymmetries in $B_s^0 \to J/\psi K_S^0$". In: *JHEP* 06 (2015), p. 131. DOI: 10.1007/JHEP06(2015)131. arXiv: 1503.07055 [hep-ex].

[101] T. Abe et al. "Belle II Technical Design Report". In: (Nov. 2010). arXiv: 1011.0352 [physics.ins-det].

[102] R. Aaij et al. "Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era". ISBN 978-92-9083-494-6. Geneva, Aug. 2018. DOI: 10347/15157. URL: https://cds.cern.ch/record/2636441.

[103] R. Aaij et al. "The LHCb Trigger and its Performance in 2011". In: *JINST* 8 (2013), P04022. DOI: 10.1088/1748-0221/8/04/P04022. arXiv: 1211.3055 [hep-ex].

[104] J. Albrecht et al. "Performance of the LHCb High Level Trigger in 2012". In: *J. Phys. Conf. Ser.* 513 (2014), p. 012001. DOI: 10.1088/1742-6596/513/1/012001. arXiv: 1310.8544 [hep-ex].

Acknowledgements

This dissertation would not have been possible without a large number of people. First of all, I want to thank Prof. Dr. Bernhard Spaan for the opportunity to do my bachelor thesis, master thesis and now my dissertation within his group. This did not only allowed me to analyse LHCb data and work in an internal collaboration but to gain further hands-on experience with detector quality assurance and commissioning. I am very grateful for all these opportunities.
Further, I want to thank Prof. Dr. Kevin Kröninger for reviewing this thesis.

Also I want to thank the Federal Ministry of Education and Research for the support. Further, I want to thank the whole Experimentelle Physik V working group for the very harmonic every day work life and discussions, which even progressed during a pandemic with virtual coffee/tea meetings and gaming nights. Special thanks go to all my office mates throughout the years (in historical order): Kevin, Philipp, Titus, Ramon, Vanessa, Holger, Martin, Vukan, and Gerwin. Without you, the daily work would have been most likely not that efficient and certainly less fun. Additionally, I want to thank Antje, Gerwin, Louis, Philipp, Quentin, Vukan, and especially Sophie for taking the time to proof-read at multiple stages of the writing process.

As a work in modern particle physics is always a team afford, I want to thank all involved members of the LHCb collaboration, especially the B to Charmonia working group for frequent input and discussions. Furthermore, I want to thank Vukan and Gerwin for the close and fruitful cooperation in the $B^0 \to \psi K^0_S$ measurement.

A special thank goes to my family and Franziska, for their unconditional support during the whole study and especially during the writing of this dissertation.