EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CERN-EP-20XX-ZZZ
LHCb-PAPER-20XX-YYY
November 20, 2018

Measurement of the CKM angle γ using $B_s^0 \to D_s K \pi\pi$ decays

LHCb collaboration[†]

Abstract

Guidelines for the preparation of LHCb documents are given. This is a "living" document that should reflect our current practice. It is expected that these guidelines are implemented for papers before they go into the first collaboration wide review. Please contact the Editorial Board chair if you have suggestions for modifications. This is the title page for journal publications (PAPER). For a CONF note or ANA note, switch to the appropriate template by uncommenting the corresponding line in the file `main.tex`.

Submitted to JHEP / Phys. Rev. D / Phys. Rev. Lett. / Phys. Lett. B / Eur. Phys. J. C / Nucl. Phys. B

[†]Authors are listed at the end of this paper.

1 Detector and simulation

The LHCb detector [1,2] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the pp interaction region [3], a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes [4, 5] placed downstream of the magnet. The polarity of the dipole magnet can be reversed, which is done periodically throughout the data-taking process to control systematic asymmetries. The tracking system provides a measurement of the momentum, p, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of $(15+29/p_{\mathrm{T}})$ μm, where p_{T} is the component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors [6]. The online event selection is performed by a trigger [7], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction.

At the hardware trigger stage, events are required to have a muon with high p_{T} or a hadron, photon or electron with high transverse energy in the calorimeters. For hadrons, the transverse energy threshold is 3.5 GeV. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from any primary pp interaction vertex. At least one charged particle must have a transverse momentum $p_{\mathrm{T}} > 1.6\,\mathrm{GeV}/c$ and be inconsistent with originating from a PV. A multivariate algorithm [8] is used for the identification of secondary vertices consistent with the decay of a b hadron.

Simulation is necessary to model the effects of the detector acceptance and to optimize the selection requirements. In the simulation, pp collisions are generated using Pythia [9] with a specific LHCb configuration [10]. Decays of hadrons are described by EvtGen [11], in which final-state radiation is generated using Photos [12]. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit [13] as described in Ref. [14].

2 Selection of signal candidates

In the first selection step, charged kaons and pions are reconstructed to form a D_s candidate in the decay modes $D_s \to KK\pi$, $D_s \to K\pi\pi$ and $D_s \to \pi\pi\pi$. These candidates are subsequently combined with a kaon and two pions or three pions from the secondary vertex to form $B_s^0 \to D_s K\pi\pi$ or $B_s^0 \to D_s\pi\pi\pi$ candidates. The resolution of the invariant mass of B_s^0 candidates, as well as the decay-time resolution, are improved using a kinematic fit [15] where the B_s^0 candidate is constrained to the PV for which it has the smallest IP signficance and the mass of the D_s is constrained to the world average.
Further kinematic vetoes and requirements on the particle identification (PID) information are used to distinguish the different D_s final states and isolate them from physical background, such as decays of Λ_c or $D^{\pm}$. The most abundant final state of the D_s meson, $D_s \to KK\pi$, is subdivided into $D_s \to \phi\pi$, $D_s \to K^{*0}K$ and $D_s \to (KK\pi)_{\text{non-res}}$, where the narrow ϕ and K^{*0} resonances allow for looser requirements on the PID variables for

those candidates. Additional kinematic selections are applied to the other three hadrons to suppress background from physical cross-feed, e.g. the $B_s^0 \to D_s D_s$ decay. The majority of criteria used in this analysis are guided by the selection procedures implemented in [16, 17]. To discriminate signal candidates and combinatorial background, a boosted decision tree (BDT) [18, 19] implemented in the TMVA toolkit [20] to separate signal from background is used. The decision tree is trained using background-subtacted $B_s^0 \to D_s\pi\pi\pi$ data as signal proxy, while the upper mass sideband of $B_s^0 \to D_s K\pi\pi$ candidates ($m_{B_s^0} > 5500\,\mathrm{MeV}/c^2$) is used as background proxy. Kinematic quantities of the B_s^0, the D_s and the reconstructed kaons and pions are used as discriminating variables for the training, while no information from the PID system is taken. The working point of the decision tree is chosen to optimize the significance of the $B_s^0 \to D_s K\pi\pi$ signal. After the full selection procedure is applied, approximately 1.5 % of events contain more than one signal candidate, off which all are used for the analysis.

3 Fit to invariant mass distribution of the $B_s^0 \to D_s h\pi\pi$ candidates

Probability density functions (PDFs) are used to describe the signal and background components of the invariant mass distributions of $B_s^0 D_s\pi\pi\pi$ and $B_s^0 \to D_s K\pi\pi$ candidates. They are obtained from a mixture of data-driven approaches and simulation, where the simulated distributions are corrected for kinematic differences between the simulation and data.
The shape of the signal candidates in the $B_s^0 \to D_s K\pi\pi$ and $B_s^0 \to D_s\pi\pi\pi$ distributions are modelled using a Johnsons's SU function [21], which results from a variable transformation of a normal distribution to allow for asymmetric tails. It provides a good description of the Gaussian signal peak, as well as reconstruction effects and radiative tails of the distribution. The shape of the Johnson's SU function is determined using simulation for both modes and subsequently fixed in the fit to data. To compensate small differences between the simulation and data, scale factors for the mean and width of the PDFs are introduced and floated during the fit. For the functional form of the combinatorial background, second order polynomials are used whose parameters are determined, for each D_s mode separately, in the fit to data. The partially reconstructed background component is described using an empirical description that is derived from simulation. In the fit to $B_s^0 D_s\pi\pi\pi$ data, all parameters are fixed to the ones obtained from simulation, except for a width parameter to account for small discrepancies between data and simulated samples. For the fit to $B_s^0 D_s K\pi\pi$ data, the shape is fixed to the one obtained from the control mode. A small fraction of $B_s^0 D_s\pi\pi\pi$ an $B_s^0 D_s{}^*\pi\pi\pi$ decays, where one of the pions is misidentified as a kaon, contaminate the $B_s^0 \to D_s K\pi\pi$ data sample. Simulated samples of the control mode is used to determine the shape of this background, where the mass hypothesis of one pion is changed to a kaon during the reconstruction process. The yield of this component is estimated from simulation and fixed in the fit to $B_s^0 \to D_s K\pi\pi$ data, taking into account the misidentification probability given the particle identification requirements imposed during the selection process.
Figure 1 shows the result of the fit to the invariant mass distributions of $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_s K\pi\pi$ candidates, where the data from Run I and II, as well as all D_s decay modes, are overlaid.

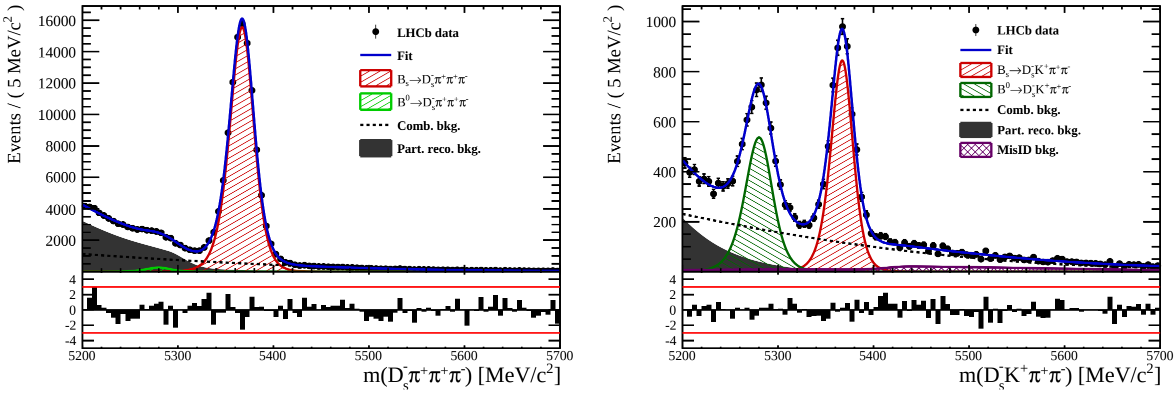

Figure 1: Invariant mass distribution of $B_s^0 \to D_s\pi\pi\pi$ (left) and $B_s^0 \to D_s K\pi\pi$ (right) candidates. The fit described in the text is overlaid.

4 Selection of signal candidates

In the first selection step, charged kaons and pions are reconstructed to form a D_s candidate in the decay modes $D_s \to KK\pi$, $D_s \to K\pi\pi$ and $D_s \to \pi\pi\pi$. These candidates are subsequently combined with a kaon and two pions or three pions from the secondary vertex to form $B_s^0 \to D_s K\pi\pi$ or $B_s^0 \to D_s\pi\pi\pi$ candidates. The resolution of the invariant mass of B_s^0 candidates, as well as the decay-time resolution, are improved using a kinematic fit [15] where the B_s^0 candidate is constrained to the PV for which it has the smallest IP signficance and the mass of the D_s is constrained to the world average.

Further kinematic vetoes and requirements on the particle identification (PID) information are used to distinguish the different D_s final states and isolate them from physical background, such as decays of Λ_c or $D^{\pm}$. The most abundant final state of the D_s meson, $D_s \to KK\pi$, is subdivided into $D_s \to \phi\pi$, $D_s \to K^{*0}K$ and $D_s \to (KK\pi)_{\text{non-res}}$, where the narrow ϕ and K^{*0} resonances allow for looser requirements on the PID variables for those candidates. Additional kinematic selections are applied to the other three hadrons to suppress background from physical cross-feed, e.g. the $B_s^0 \to D_s D_s$ decay. The majority of criteria used in this analysis are guided by the selection procedures implemented in [16,17].

To discriminate signal candidates and combinatorial background, a boosted decision tree (BDT) [18, 19] implemented in the TMVA toolkit [20] to separate signal from background is used. The decision tree is trained using background-subtacted $B_s^0 \to D_s\pi\pi\pi$ data as signal proxy, while the upper mass sideband of $B_s^0 \to D_s K\pi\pi$ candidates ($m_{B_s^0} > 5500\,\text{MeV}/c^2$) is used as background proxy. Kinematic quantities of the B_s^0, the D_s and the reconstructed kaons and pions are used as discriminating variables for the training, while no information from the PID system is taken. The working point of the decision tree is chosen to optimize the significance of the $B_s^0 \to D_s K\pi\pi$ signal. After the full selection procedure is applied, approximately 1.5 % of events contain more than one signal candidate, off which all are used for the analysis.

5 Flavour tagging

To identify the initial flavour state of the B_s^0 meson, a number of flavour tagging algorithms are used that either exploit the pair-wise production of b quarks and determine the flavour of the non-signal b-hadron produced in the event (opposite site, OS) or use particles produced in the fragmentation of the signal candidate $B_s^0/\overline{B}_s^0$ (same side, SS). The same side kaon tagger searches for the charge of an additional kaon that accompanies the fragmentation of the signal B_s^0 or $\overline{B}_s^0$ candidate [22]. For the opposite site tagger [23], five different tagging algorithms are chosen: the algorithms that use the charge of an electron or a muon from semi-leptonic B decays, the tagger that uses the charge of a kaon from a b $\to$ c $\to$ s decay chain, the algorithm which reconstructs opposite-side charm hadrons from a number of c-decays and the algorithm that determines the $B_s^0/\overline{B}_s^0$ candidate flavour from the charge of a secondary vertex, reconstructed from the OS b-decay product. All five taggers are then combined into a single OS tagger.
Every tagging algorithm is prone to misidentify the signal candidate at a certain mistag rate ω = (wrong tags)/(all tags). This might be caused by particle misidentification, flavour oscillation of the neutral opposite site B-meson or by the selection of tracks from the underlying event. An imperfect determination of the B_s^0 production flavour dilutes the observed CP asymmetry by $D_{tag} = 1 - 2\omega$. Therefore, the statistical precision with which the CP asymmetry can be measured scales as the inverse square root of the effective tagging efficiency:

$$\epsilon_{eff} = \epsilon_{tag}(1-2\omega)^2, \tag{1}$$

where ϵ_{tag} is the fraction of tagged signal candidates.
For each $B_s^0/\overline{B}_s^0$ candidate, the tagging algorithms provide a prediction for the mistag probability η based on the output of neural networks that take various variables, such as the kinematics of the tagging particles, as input. These are trained on either simulated or sWeighted samples of flavour specific control channels ($B_s^0 \to D_s^- \pi^+$ (SS algorithm) and $B^+ \to J/\psi K^+$ (OS algorithms)) and are optimized for highest ϵ_{eff} on data. Utilizing flavour-specific final states, the estimated mistag η of each tagger has to be calibrated to match the actual mistag probability ω. For the calibration, a linear model

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - \langle\eta\rangle), \tag{2}$$

is used where $\langle\eta\rangle$ is the average estimated mistag probability. A perfectly calibrated tagger would lead to $\omega(\eta) = \eta$ and one would expect $p_1 = 1$ and $p_0 = \langle\eta\rangle$. Due to the different interaction cross-sections of oppositely charged kaons, the tagging calibration parameters depend on the initial state flavour of the B_s^0. Therefore, the flavour asymmetry parameters Δp_0, Δp_1 and $\Delta\epsilon_{tag}$ are introduced and defined as the difference of the corresponding values for B_s^0 and $\overline{B}_s^0$ mesons.

The OS electron, muon, kaon, charm and the secondary vertex charge tagging algorithms are individually calibrated and then combined into a singleOS tagger. We choose the flavour specific decay $B_s \to D_s\pi\pi\pi$ as calibration mode due to the portability to the signal mode since its similarity wit the $B_s \to D_s K\pi\pi$ decay. The calibration is performed separately for Run-I and Run-II data, while the OS-c tagger is not included for Run-I data since the statistics is too low. Tables 1 and 2 list the measured tagging performances. The combined OS and SS-Kaon taggers are calibrated simultaneously by means of a fit to the decay-time distribution of background-subtracted $B_s \to D_s\pi\pi\pi$ candidates,

Table 1: The flavour tagging performances for the used OS taggers for Run-I data.

Tagger	ϵ	ω	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$
OS μ	$(8.713 \pm 0.206)\%$	$(28.893 \pm 0.180(\text{stat}) \pm 2.291(\text{cal}))\%$	$(1.553 \pm 0.045(\text{stat}) \pm 0.337(\text{cal}))\%$
OS e	$(3.201 \pm 0.129)\%$	$(28.792 \pm 0.363(\text{stat}) \pm 3.611(\text{cal}))\%$	$(0.576 \pm 0.030(\text{stat}) \pm 0.196(\text{cal}))\%$
OS K	$(32.230 \pm 0.342)\%$	$(38.451 \pm 0.093(\text{stat}) \pm 1.145(\text{cal}))\%$	$(1.719 \pm 0.033(\text{stat}) \pm 0.341(\text{cal}))\%$
Vertex Charge	$(21.855 \pm 0.302)\%$	$(35.712 \pm 0.091(\text{stat}) \pm 1.474(\text{cal}))\%$	$(1.785 \pm 0.033(\text{stat}) \pm 0.368(\text{cal}))\%$

Table 2: The flavour tagging performances for the used OS taggers for Run-II data.

Tagger	ϵ	ω	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$
OS μ	$(9.664 \pm 0.151)\%$	$(30.911 \pm 0.115(\text{stat}) \pm 1.369(\text{cal}))\%$	$(1.409 \pm 0.028(\text{stat}) \pm 0.202(\text{cal}))\%$
OS e	$(4.590 \pm 0.107)\%$	$(33.577 \pm 0.140(\text{stat}) \pm 2.007(\text{cal}))\%$	$(0.495 \pm 0.014(\text{stat}) \pm 0.121(\text{cal}))\%$
OS K	$(20.185 \pm 0.205)\%$	$(36.918 \pm 0.071(\text{stat}) \pm 0.969(\text{cal}))\%$	$(1.382 \pm 0.021(\text{stat}) \pm 0.205(\text{cal}))\%$
Vertex Charge	$(20.597 \pm 0.207)\%$	$(34.751 \pm 0.075(\text{stat}) \pm 0.961(\text{cal}))\%$	$(1.916 \pm 0.027(\text{stat}) \pm 0.242(\text{cal}))\%$
OS c	$(5.500 \pm 0.116)\%$	$(32.581 \pm 0.092(\text{stat}) \pm 1.848(\text{cal}))\%$	$(0.668 \pm 0.016(\text{stat}) \pm 0.142(\text{cal}))\%$

as discussed in Sec. 8. In this fit, the predicted mistag probabilities η_{OS} and η_{SS} are included as per-event observables, effectively giving a larger weight to the events that have a lower mistag probability. The tagger responses are combined into a single response on an event-by-event basis during the fit. Tables 3 and 4 report the tagging performances for the OS and SS combination considering three mutually exclusive categories of tagged events: OS only, SS only and both OS and SS. The tagging calibration parameters are listed in Table **??**.

Table 3: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-I data.

$B_s \to D_s\pi\pi\pi$	$\epsilon_{tag}[\%]$	$\langle\omega\rangle[\%]$	$\epsilon_{eff}[\%]$
Only OS	14.74 ± 0.11	39.09 ± 0.80	1.25 ± 0.16
Only SS	35.38 ± 0.18	44.26 ± 0.62	1.05 ± 0.18
Both OS-SS	33.04 ± 0.30	37.33 ± 0.73	3.41 ± 0.33
Combined	83.16 ± 0.37	40.59 ± 0.70	5.71 ± 0.40

Table 4: The flavour tagging performances for only OS tagged, only SS tagged and both OS and SS tagged events for Run-II data.

$B_s \to D_s\pi\pi\pi$	$\epsilon_{tag}[\%]$	$\langle\omega\rangle[\%]$	$\epsilon_{eff}[\%]$
Only OS	11.78 ± 0.05	37.01 ± 0.51	1.15 ± 0.07
Only SS	41.28 ± 0.10	42.65 ± 0.35	1.79 ± 0.12
Both OS-SS	28.62 ± 0.15	35.35 ± 0.40	3.63 ± 0.16
Combined	81.68 ± 0.19	39.28 ± 0.40	6.57 ± 0.21

6 Decay-time resolution

The CP-violating parameters measured in the time-dependent fit are prone to dilution due to the fast B_s^0-$\overline{B}_s^0$ oscillation frequency, which is of the same order as the average decay-time resolution of the LHCb detector of $\mathcal{O}(50\,\mathrm{fs}^{-1})$ [2]. Therefore, it is crucial to correctly describe the decay-time resolution in order to accurately measure time-dependent CP violation. In particular, the parameters related to the amplitudes of the sine and cosine terms in Equation xXx are highly corelated to the chosen resolution model. Since the time resolution depends on the particular event, especially the decay time itself, the sensitivity on the CP parameters can be significantly improved by using an event-dependent model rather than an average resolution. For this purpose, the signal PDF is convolved with a Gaussian resolution function that has a different width for each candidate, obtained from the global kinematic fit to the B_s^0 vertex and the D_s mass. To ensure the correct application, the per-candidate decay-time uncertainty σ_t has to be calibrated to match the effective decay-time resolution observed in data, $\sigma(\sigma_t)$.

For data taken during Run I, a study of simulated $B_s^0 \to D_s K\pi\pi$ events is used to confirm the portability of the calibration relation determined in the closely related analysis of $B_s^0 \to D_s K$ decays [17]. The spread of the difference between the reconstructed and true decay time, $\Delta t = t - t_{\mathrm{true}}$, follows the shape of a double Gaussian distribution and is a direct measure of the effective decay-time resolution for simulated events. The resulting two Gaussian widths are combined to calculate the dilution $\mathcal{D}$, which describes the effective damping of the CP amplitudes due to the finite time resolution:

$$\mathcal{D} = f_1 e^{-\sigma_1^2 \Delta m_s{}^2/2} + (1 - f_1) e^{-\sigma_2^2 \Delta m_s{}^2/2}, \tag{3}$$

where σ_1 and σ_2 are the widths of the Gaussians, f_1 is the relative fraction of events described by the first Gaussian relative to the second and Δm_s is the oscillation frequency of B_s^0 mesons. An effective single Gaussian width is calculated from the dilution as,

$$\sigma_{eff} = \sqrt{(-2/\Delta m_s{}^2) \ln \mathcal{D}}, \tag{4}$$

which converts the resolution into a single-Gaussian function with an effective resolution that causes the same damping effect on the magnitude of the B_s oscillation. The calibration relation is found to be portable between the $B_s^0 \to D_s K$ and $B_s^0 \to D_s K\pi\pi$ decay channels and thus it is used for data taken in Run I.

For data taken during Run II, the calibration is performed using a sample of prompt D_s mesons, combined with a kaon and two pions originating from the primary vertex to form 'fake' B_s^0 candidates with a lifetime of $t = 0$ by construction. The spread of observed decay times of the 'fake' candidates is described by a double Gaussian function, where only negative decay times are used to determine the effective resolution to avoid uncertainties introduced by physical backgrounds. Following the same approach used for data taken during Run I, the effective resolution is calculated from the dilution $\mathcal{D}$.

7 Decay-time acceptance

The decay-time distribution of the B_s^0 mesons is distorted due to the geometry of the LHCb detector and the applied selections, described in Section 4. In particular, any requirement on the flight distance, the impact parameter or the direction angle (DIRA) of the B_s^0 mesons leads to a decay-time dependent efficiency $\epsilon(t)$. This acceptance effect in the $B_s^0 \to D_s K\pi\pi$ decay-time distribution is strongly correlated wih the CP parameters. However, for the flavour-specific control channel $B_s^0 \to D_s \pi\pi\pi$, the acceptance can be measured since all CP-violating parameters are fixed to zero or unity. Using Γ_s as input, the parameters of the acceptance shape, as well as Δm_s, is measured using a time-dependent fit to the background-subtracted decay-time distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates. To correct small differences between the signal and the control sample, the fit is performed simultaneously to the decay-time distributions of simulated $B_s^0 \to D_s\pi\pi\pi$, $B_s^0 \to D_sK\pi\pi$, as well as to $B^0 \to D_sK\pi\pi$ data candidates. For all samples, the acceptance is parametrized using segments of cubic b-splines, which are implemented into the decay-time PDF in an analytic way [24]. The decay-time distribution of background-subtracted $B_s^0 \to D_s\pi\pi\pi$ data candidates, as well as the time-dependent fit to determine the accpetance shape, is shown in Figure 2.

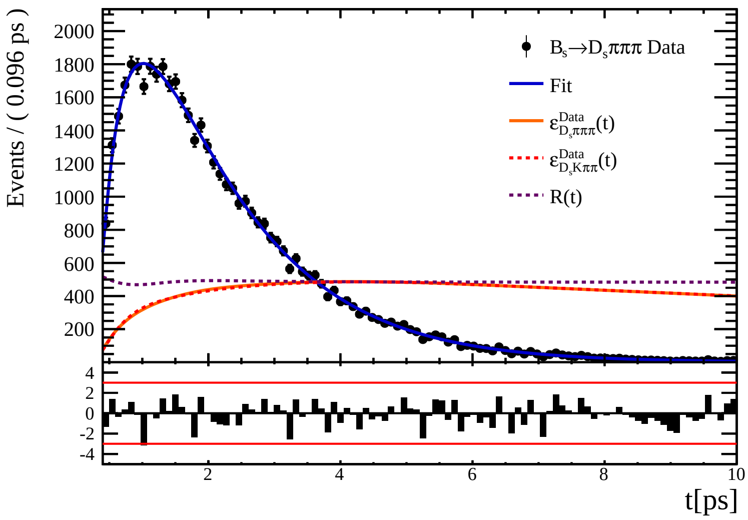

Figure 2: Decay-time distribution of background-subtracted $B_s^0 \to D_s\pi\pi\pi$ data. The fit to determine the shape of the time-dependent efficiency is overlaid, where the acceptance function is shown in an arbitrary scale.

8 Decay-time fit to $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ candidates

The sFit technique [25] is used to statistically subtract the background from the $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ data samples. During the fit procedure, Γ_s and $\Delta\Gamma_s$ are fixed to the corresponding HFLAV [26] world average. The B_s^0 production asymmetry A_p, defined as the relative difference in the production cross sections $\frac{\sigma(\bar{B}_s^0)-\sigma(B_s^0)}{\sigma(\bar{B}_s^0)+\sigma(B_s^0)}$, contributes with a factor $(1 \pm A_p)$ to the signal PDF, where the sign depends on the flavour of the b meson. For data recorded during Run I, A_p is taken from [27]. The PDFs used for the fits to the $B_s^0 \to D_s\pi\pi\pi$ and $B_s^0 \to D_sK\pi\pi$ candidates are convolved with a Gaussian function representing the per-candidate decay-time resolution and multiplied by the decay-time acceptance described in Sections 6 and 7, respectively.

Since the decay $B_s^0 \to D_s\pi\pi\pi$ is flavour specific, the *CP* coefficients defined in Equation xXx can be fixed to $C = 1$ and $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$. In the fit, the calibration parameters for the OS and SS taging algorithms, the B_s^0 production asymmetry for Run II data, as well as the B_s^0 oscillation frequency Δm_s, are measured. The fit to the decay-time distribution is shown in Figure 3 and the mixing frequency is measured to be

$$\Delta m_s = (xx.xx \pm 0.0084 \pm 0.0058)\,\mathrm{ps}^{-1}, \tag{5}$$

where the uncertainties are statistical and systematic, respectively.

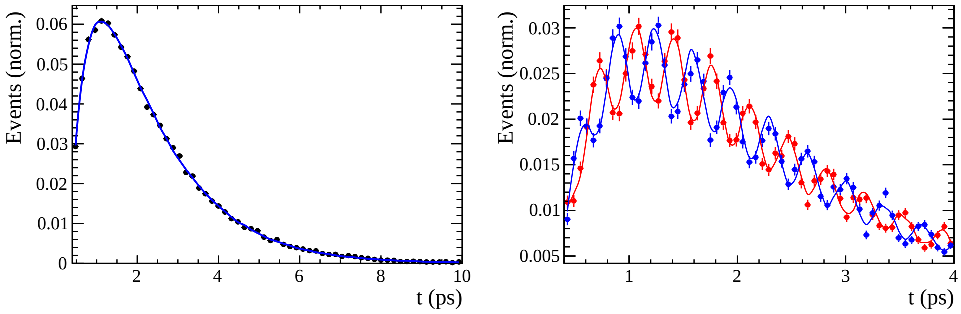

Figure 3: Left: Flavour averaged decay-time distribution of $B_s^0 \to D_s\pi\pi\pi$ candidates. Right: Tagged decay-time distribution of mixed (red) and unmixed (blue) signal candidates.

The fit to $B_s^0 \to D_s K\pi\pi$ data is sensitive to a possible charge asymmetry of the kaon, introduced by its charge-dependent nuclear cross-section. Therefore, the detection asymmetry A_{det} is introduced and multiplied as $(1 \pm A_{det})$, where the sign depends on the charge of the kaon, to the signal PDF. It is determined using a data-driven technique described in [28]. The tagging calibration parameters are taken from the fit to the control sample and included in the fit using Gaussian-constrains. The measured *CP* coefficients are reported in Table 5 and the fit projection is shown in Figure 4.

Table 5: *CP* coefficients determined from a fit to the $B_s \to D_s K\pi\pi$ decay-time distribution. The uncertainties are statistical and systematic, respectively.

Fit Parameter	Value
C	x.xx ± 0.12 ± 0.02
D	x.xx ± 0.32 ± 0.08
$\bar{D}$	x.xx ± 0.30 ± 0.08
S	x.xx ± 0.17 ± 0.04
$\bar{S}$	x.xx ± 0.17 ± 0.04

9 Systematic uncertainties

Systematic uncertainties derive from the modelling of the background in the invariant B_s^0 mass distribution, the detection and production asymmetries A_{det} and A_p, the limited

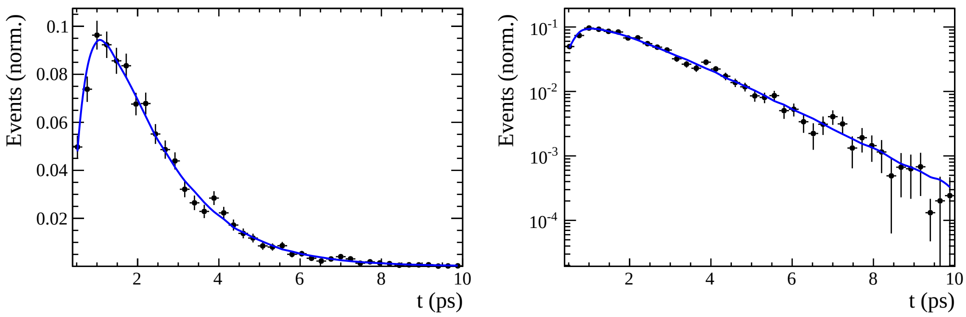

Figure 4: Decay-time distribution of $B_s^0 \to D_s K\pi\pi$ signal candidates with the fit projection overlaid in (left) regular and (right) logarithmic scale.

knowledge of the decay-time acceptance and resolution, as well as from the uncertainty on the LHCb length and momentum scale, which directly translates in an uncertainty on Δm_s. For the time-dependent amplitude fit, additional sources of systematic uncertainties arise from the description of the phase-space acceptance, the modelling of resonance shapes and the explicit choice of amplitudes used in the fit. The systematic uncertainties on the measured observables are summarized in Table 6 for the phase-space integrated decay-time fit and in Table 7 for the full time-dependent amplitude fit to $B_s^0 \to D_s K\pi\pi$ data. The individual contributions are discussed below.

Since the choice of signal and background models for the description of the invariant mass spectrum of B_s^0 candidates is not unique, several alternative parametrizations are tested. For each case new signal weights are obtained and the sFit procedure is repeated. The sample variance of the obtained differences to the nominal fit value are assigned as systematic uncertainty due to the background subtraction.

The fit procedure is validated using a large set of pseudoexperiments, which are generated with the central values of the *CP* parameters reported in Table 5. Subsequently, they are processed by the nominal fit procedure and the values obtained by the fits are compared to the generated ones. For each parameter, a distribution is formed by normalizing the differences between fitted and generated values to the uncertainties measured in the nominal fit. The mean and the width of the distribution is added in quadrature and assigned as systematic uncertainty due to a fit bias for the respective parameter.

The systematic uncertainty related to the decay-time acceptance, as well as Γ_s and $\Delta\Gamma_s$ are studied with the same set of pseudoexperiments. They are fit with the nominal model and a model in which the acceptance parameters together with Γ_s and $\Delta\Gamma_s$ are randomized within their uncertainties. Distributions are calculated by dividing the difference between the obtained values of the nominal fit and the fit using randomly shifted acceptance parameters by the uncertainty in the nominal fit. The bias in the mean of this distribution is added to its width, in quadrature, in order to arrive at the final systematic uncertainty for each parameter.

This procedure is repeated, varying the production, detection asymmetries and Δm_s within their respective uncertainties instead of the acceptance parametrization.

To study systematic effects originating from the scaling of the decay-time error estimate, two alternative decay-time resolution models are tested. Due to the high correlation between the decay-time resolution and the tagging calibration, their systematic uncertainty needs to be studied simultaneously. First, the decay-time dependent fit to

$B_s^0 \to D_s\pi\pi\pi$ data is repeated using a alternative decay-time error scaling function. In this fit, new tagging calibration parameters are obtained and subsequently used with Gaussian-constrains in the fit to $B_s^0 D_s K\pi\pi$ data. The largest change in the central value of each CP observable is assigned as the systematic uncertainty due to the decay-time resolution and flavour tagging for the respective parameter.
A possible systematic effect is studied by repeating the sFit, randomly keeping only one candidate in events where multiple candidates are found. No shift in the nominal fit values is observed.
The uncertainty on the LHCb length scale is estimated to be at most 0.020%, which translates directly in an uncertainty on Δm_s of 0.020% with other parameters being unaffected.

DESCRIPTION OF AMPLITUDE SYSTEMATICS HERE

Table 6: Systematic uncertainties on the fit parameters of the phase-space integrated fit to $B_s \to D_s K\pi\pi$ data in units of statistical standard deviations.

Fit Parameter	Fit bias	Acceptance	Resolution	Δm_s	Asymmetries	Background	Total
C	0.06	0.04	0.15	0.06	0.02	0.06	0.19
D	0.02	0.22	0.01	0.02	0.04	0.10	0.25
$\bar{D}$	0.04	0.22	0.02	0.02	0.04	0.13	0.26
S	0.01	0.02	0.07	0.21	0.03	0.06	0.23
$\bar{S}$	0.07	0.02	0.05	0.22	0.02	0.03	0.24

Table 7: Systematic uncertainties on the fit parameters of the full time-dependent amplitude fit to $B_s \to D_s K\pi\pi$ data in units of statistical standard deviations.

Fit Parameter	Fit bias	Time-Acc.	Resolution	Δm_s	Asymmetries	Background	Lineshapes	Resonances m,Γ	Form-Factors	Phsp-Acc.	Amp. Model	Total
$B_s \to D_s\,(K_1(1270) \to K^*(892)\,\pi)$ Mag	0.10	0.01	0.04	0.01	0.00	0.13	0.48	0.24	0.52	0.06		0.77
$B_s \to D_s\,(K_1(1270) \to K^*(892)\,\pi)$ Phase	0.07	0.01	0.04	0.01	0.01	0.08	0.35	0.28	0.34	0.12		0.58
$B_s \to D_s\,(K_1(1270) \to K_0^*(1430)\,\pi)$ Mag	0.04	0.01	0.01	0.00	0.00	0.24	1.44	0.11	0.17	0.04		1.47
$B_s \to D_s\,(K_1(1270) \to K_0^*(1430)\,\pi)$ Phase	0.04	0.01	0.02	0.01	0.00	0.19	5.83	0.19	0.61	0.09		5.87
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Mag$(b \to c)$	0.13	0.03	0.16	0.06	0.02	0.34	1.32	0.37	0.78	0.19		1.64
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Phase$(b \to c)$	0.14	0.02	0.09	0.02	0.01	0.18	0.54	0.26	0.40	0.08		0.77
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Mag$(b \to u)$	0.10	0.04	0.05	0.12	0.04	0.32	0.35	0.22	0.73	0.16		0.93
$B_s \to D_s\,(K_1(1400) \to K^*(892)\,\pi)$ Phase$(b \to u)$	0.02	0.04	0.04	0.10	0.03	0.08	0.79	0.21	0.31	0.08		0.89
$B_s \to D_s\,(K^*(1410) \to K^*(892)\,\pi)$ Mag$(b \to c)$	0.08	0.03	0.08	0.08	0.03	0.18	0.61	0.25	0.75	0.28		1.06
$B_s \to D_s\,(K^*(1410) \to K^*(892)\,\pi)$ Phase$(b \to c)$	0.35	0.01	0.06	0.01	0.01	0.13	0.60	0.19	0.68	0.08		1.00
$B_s \to D_s\,(K^*(1410) \to K\,\rho(770))$ Mag	0.35	0.01	0.02	0.01	0.00	0.18	0.59	0.12	0.34	0.06		0.79
$B_s \to D_s\,(K^*(1410) \to K\,\rho(770))$ Phase	0.18	0.00	0.01	0.01	0.00	0.24	0.34	0.09	0.21	0.06		0.51
$B_s \to D_s\,(K(1460) \to K^*(892)\,\pi)$ Mag$(b \to u)$	0.14	0.03	0.05	0.05	0.02	0.37	0.43	0.27	0.60	0.12		0.89
$B_s \to D_s\,(K(1460) \to K^*(892)\,\pi)$ Phase$(b \to u)$	0.13	0.04	0.11	0.07	0.03	0.21	0.84	0.49	0.46	0.06		1.11
$B_s \to (D_s\,\pi)_P\; K^*(892)$ Mag$(b \to c)$	0.03	0.02	0.06	0.02	0.01	0.24	0.95	0.11	0.55	0.13		1.14
$B_s \to (D_s\,\pi)_P\; K^*(892)$ Phase$(b \to c)$	0.20	0.01	0.13	0.02	0.01	0.51	1.10	0.18	0.52	0.26		1.38
$B_s \to (D_s\,\pi)_P\; K^*(892)$ Mag$(b \to u)$	0.14	0.04	0.07	0.06	0.02	0.11	0.78	0.24	0.54	0.17		1.01
$B_s \to (D_s\,\pi)_P\; K^*(892)$ Phase$(b \to u)$	0.24	0.05	0.19	0.06	0.03	0.47	1.54	0.28	0.59	0.17		1.77
$B_s \to (D_s\,K)_P\; \rho(770)$ Mag$(b \to u)$	0.35	0.04	0.02	0.05	0.02	0.25	0.75	0.31	0.60	0.06		1.10
$B_s \to (D_s\,K)_P\; \rho(770)$ Phase$(b \to u)$	0.12	0.03	0.05	0.06	0.02	0.68	0.50	0.38	0.66	0.08		1.14
$m_{K_1(1400)}$	0.09	0.01	0.08	0.01	0.00	0.14	0.21	0.13	0.37	0.09	0.72	0.87
$\Gamma_{K_1(1400)}$	0.01	0.01	0.01	0.02	0.01	0.14	0.46	0.13	0.44	0.10	0.62	0.91
$m_{K^*(1410)}$	0.05	0.01	0.02	0.01	0.00	0.08	0.26	0.04	1.29	0.12	0.67	1.49
$\Gamma_{K^*(1410)}$	0.25	0.00	0.02	0.01	0.00	0.14	0.15	0.04	1.40	0.07	0.72	1.61
r	0.11	0.05	0.09	0.12	0.03	0.47	0.74	0.12	0.26	0.12	0.79	1.23
δ	0.19	0.04	0.07	0.10	0.05	0.10	0.29	0.03	0.11	0.02	0.52	0.66
$\gamma - 2\beta_s$	0.10	0.06	0.12	0.06	0.02	0.12	0.27	0.03	0.10	0.03	0.39	0.53

References

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

[2] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[3] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[4] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[5] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **9** (2017) P11016, `arXiv:1708.00819`.

[6] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[7] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[8] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, `arXiv:1210.6861`.

[9] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`; T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[10] I. Belyaev *et al.*, *Handling of the generation of primary events in* Gauss*, the LHCb simulation framework*, Nuclear Science Symposium Conference Record (NSS/MIC) **IEEE** (2010) 1155.

[11] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[12] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, `arXiv:hep-ph/0506026`.

[13] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[14] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[15] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nuclear Instruments and Methods in Physics Research A **552** (2005) 566, `arXiv:physics/0503191`.

[16] LHCb collaboration, R. Aaij *et al.*, *First observation of the decays $\overline{B}{}^0_{(s)} \to D_s^+ K^- \pi^+ \pi^-$ and $\overline{B}{}^0_s \to D_{s1}(2536)^+ \pi^-$*, Phys. Rev. **D86** (2012) 112005, `arXiv:1211.1541`.

[17] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays*, JHEP **03** (2018) 059, `arXiv:1712.07428`.

[18] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[19] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.

[20] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[21] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949) 149.

[22] LHCb, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, `arXiv:1602.07252`.

[23] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, `arXiv:1202.4979`.

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

[25] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[26] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, `arXiv:1412.7515`, updated results and plots available at `http://www.slac.stanford.edu/xorg/hfag/`.

[27] LHCb, 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*, Phys. Lett. **B774** (2017) 139, `arXiv:1703.08464`.

[28] A. Davis *et al.*, *Measurement of the instrumental asymmetry for $K^- \pi^+$-pairs at LHCb in Run 2*, Tech. Rep. LHCb-PUB-2018-004. CERN-LHCb-PUB-2018-004, CERN, Geneva, Mar, 2018.