Measurement of the CKM Angle Using $B^{\pm}$ $D^{\pm}$ $K^{\mp}$ with $D$ $K^{S}$ $0^{+}$ $K^{0}$ $K^{+}$ Decays.

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>The LHCb collaboration, et al. “Measurement of the CKM Angle Using $B^{\pm}$ $D^{\pm}$ $K^{\mp}$ with $D$ $K^{S}$ $0^{+}$ $K^{0}$ $K^{+}$ Decays.” Journal of High Energy Physics, vol. 2018, no. 8, Aug. 2018.</th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="https://doi.org/10.1007/JHEP08(2018)176">https://doi.org/10.1007/JHEP08(2018)176</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Springer Berlin Heidelberg</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Mar 10 19:48:26 EDT 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/119154">http://hdl.handle.net/1721.1/119154</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a></td>
</tr>
</tbody>
</table>
Measurement of the CKM angle $\gamma$ using $B^{\pm} \rightarrow DK^{\pm}$ with $D \rightarrow K^0_S\pi^+\pi^-$, $K^0_SK^+K^-$ decays

The LHCb collaboration

E-mail: mikkel.bjoern@physics.ox.ac.uk

Abstract: A binned Dalitz plot analysis of $B^{\pm} \rightarrow DK^{\pm}$ decays, with $D \rightarrow K^0_S\pi^+\pi^-$ and $D \rightarrow K^0_SK^+K^-$, is used to perform a measurement of the CP-violating observables $x_\pm$ and $y_\pm$, which are sensitive to the Cabibbo-Kobayashi-Maskawa angle $\gamma$. The analysis is performed without assuming any $D$ decay model, through the use of information on the strong-phase variation over the Dalitz plot from the CLEO collaboration. Using a sample of proton-proton collision data collected with the LHCb experiment in 2015 and 2016, and corresponding to an integrated luminosity of 2.0 fb$^{-1}$, the values of the CP violation parameters are found to be $x_- = (9.0 \pm 1.7 \pm 0.7 \pm 0.4) \times 10^{-2}$, $y_- = (2.1 \pm 2.2 \pm 0.5 \pm 1.1) \times 10^{-2}$, $x_+ = (-7.7 \pm 1.9 \pm 0.7 \pm 0.4) \times 10^{-2}$, and $y_+ = (-1.0 \pm 1.9 \pm 0.4 \pm 0.9) \times 10^{-2}$. The first uncertainty is statistical, the second is systematic, and the third is due to the uncertainty on the strong-phase measurements. These values are used to obtain $\gamma = (87^{+11}_{-12})^\circ$, $r_B = 0.086^{+0.013}_{-0.014}$, and $\delta_B = (101 \pm 11)^\circ$, where $r_B$ is the ratio between the suppressed and favoured $B$-decay amplitudes and $\delta_B$ is the corresponding strong-interaction phase difference. This measurement is combined with the result obtained using 2011 and 2012 data collected with the LHCb experiment, to give $\gamma = (80^{+10}_{-9})^\circ$, $r_B = 0.080 \pm 0.011$, and $\delta_B = (110 \pm 10)^\circ$.

Keywords: B physics, CKM angle gamma, CP violation, Flavor physics, Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1806.01202
1 Introduction

The Standard Model (SM) description of CP violation [1, 2] can be tested by overconstraining the angles of the Unitarity Triangle. The Cabibbo-Kobayashi-Maskawa (CKM) angle \( \gamma \equiv \arg \left( V_{ud} V_{ub}^{*} / V_{cd} V_{cb}^{*} \right) \) is experimentally accessible through the interference between \( \bar{b} \rightarrow \bar{c}u\bar{s} \) and \( \bar{b} \rightarrow \bar{u}c\bar{s} \) transitions. It is the only CKM angle easily accessible in tree-level processes and it can be measured with negligible uncertainty from theory [3]. Hence, in the absence of new physics effects at tree level, a precision measurement of \( \gamma \) provides a SM benchmark that can be compared with other CKM-matrix observables more likely to be affected by physics beyond the SM. Such comparisons are currently limited by the uncertainty on direct measurements of \( \gamma \), which is about 5° [4] and is driven by the LHCb average.

The effects of interference between \( \bar{b} \rightarrow \bar{c}u\bar{s} \) and \( \bar{b} \rightarrow \bar{u}c\bar{s} \) transitions can be probed by studying CP-violating observables in \( B^{\pm} \rightarrow DK^{\pm} \) decays, where \( D \) represents a \( D^{0} \) or a \( \bar{D}^{0} \) meson reconstructed in a final state that is common to both [5–7]. This decay mode has been studied at LHCb with a wide range of \( D \)-meson final states to measure observables with sensitivity to \( \gamma \) [8–11]. In addition to these studies, other \( B \) decays have also been used with a variety of techniques to determine \( \gamma \) [12–15].
This paper presents a model-independent study of the decay mode $B^\pm \to DK^\mp$, using $D \to K_S^0\pi^+\pi^-$ and $D \to K_S^0K^+K^-$ decays (denoted $D \to K_S^0h^+h^-$ decays). The analysis utilises $pp$ collision data accumulated with LHCb in 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and corresponding to a total integrated luminosity of 2.0 fb$^{-1}$. The result is combined with the result obtained by LHCb with the same analysis technique, using data collected in 2011 and 2012 (Run 1) at centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV [9].

The sensitivity to $\gamma$ is obtained by comparing the distributions in the Dalitz plots of $D \to K_S^0h^+h^-$ decays from reconstructed $B^+$ and $B^-$ mesons [6, 7]. For this comparison, the variation of the strong-phase difference between $D^0$ and $\overline{D}^0$ decay amplitudes within the Dalitz plot needs to be known. An attractive, model-independent, approach makes use of direct measurements of the strong-phase variation over bins of the Dalitz plot [6, 16, 17]. The strong phase can be directly accessed by exploiting the quantum correlation of $D^0\overline{D}^0$ pairs from $\psi(3770)$ decays. Such measurements have been performed by the CLEO collaboration [18] and have been used by the LHCb [9] and Belle [19] collaborations to measure $\gamma$ in $B^\pm \to DK^\mp$ decays, and have also been used to study $B^0 \to DK^{*0}$ decays [20, 21]. An alternative method relies on amplitude models of $D \to K_S^0h^+h^-$ decays, determined from flavour-tagged $D \to K_S^0h^+h^-$ decays, to predict the strong-phase variation over the Dalitz plot. This method has been used for a variety of $B$ decays [22–28].

The separation of data into binned regions of the Dalitz plot leads to a loss of statistical sensitivity in comparison to using an amplitude model [16, 17]. However, the advantage of using the direct strong-phase measurements resides in the model-independent nature of the systematic uncertainties. Where the direct strong-phase measurements are used, there is only a systematic uncertainty associated with the finite precision of such measurements. Conversely, systematic uncertainties associated with determining a phase from an amplitude model are difficult to evaluate, as common approaches to amplitude-model building break the optical theorem [29]. Therefore, the loss in statistical precision is compensated by reliability in the evaluation of the systematic uncertainty, which is increasingly important as the overall precision on the CKM angle $\gamma$ improves.

2 Overview of the analysis

The amplitude of the decay $B^- \to DK^-$, $D \to K_S^0h^+h^-$ can be written as a sum of the favoured $B^- \to D^0K^-$ and suppressed $B^- \to \overline{D}^0K^-$ contributions as

$$A_B(m_2^2, m_4^2) \propto A_D(m_2^2, m_4^2) + r_B e^{i(\delta_B - \gamma)} A_{\overline{D}}(m_2^2, m_4^2),$$

(2.1)

where $m_2^2$ and $m_4^2$ are the squared invariant masses of the $K_S^0h^-$ and $K_S^0h^+$ particle combinations, respectively, that define the position of the decay in the Dalitz plot, $A_D(m_2^2, m_4^2)$ is the $D^0 \to K_S^0h^+h^-$ decay amplitude, and $A_{\overline{D}}(m_2^2, m_4^2)$ the $\overline{D}^0 \to K_S^0h^+h^-$ decay amplitude. The parameter $r_B$ is the ratio of the magnitudes of the $B^- \to D^0K^-$ and $B^- \to \overline{D}^0K^-$ amplitudes, while $\delta_B$ is their strong-phase difference. The equivalent expression for the charge-conjugated decay $B^+ \to DK^+$ is obtained by making the substitutions...
\[ \gamma \to -\gamma \text{ and } A_D(m^2, m^2_\pm) \leftrightarrow A_D(m^2_\pm, m^2). \]
Neglecting CP violation in charm decays, the charge-conjugated amplitudes satisfy the relation
\[ A_D(m^2, m^2_\pm) = A_D(m^2_\pm, m^2). \]

The D-decay Dalitz plot is partitioned into \( 2 \times N \) bins labelled from \( i = -N \) to \( i = +N \) (excluding zero), symmetric around \( m^- = m^+ \) such that if \( (m^2, m^2_\pm) \) is in bin \( i \) then \( (m^2_\pm, m^2) \) is in bin \( -i \). By convention, the positive values of \( i \) correspond to bins for which \( m^2_\pm > m^2 \). The strong-phase difference between the \( D^0 \) and \( \bar{D}^0 \)-decay amplitudes at a given point on the Dalitz plot is denoted as \( \delta_D(m^2, m^2_\pm) \). The cosine of \( \delta_D(m^2, m^2_\pm) \) weighted by the D-decay amplitude and averaged over bin \( i \) is written as \( c_i \) [6], and is given by
\[
 c_i \equiv \frac{\int dm^2_+ dm^2_- |A_D(m^2, m^2_\pm)||A_D(m^2_\pm, m^2_\pm)| \cos[\delta_D(m^2, m^2_\pm) - \delta_D(m^2_\pm, m^2_\pm)]}{\sqrt{\int dm^2_+ dm^2_- |A_D(m^2, m^2_\pm)|^2 \int dm^2_+ dm^2_- |A_D(m^2_\pm, m^2_\pm)|^2}}, \tag{2.2}
\]
where the integrals are evaluated over the phase space of bin \( i \). An analogous expression can be written for \( s_i \), which is the sine of the strong-phase difference weighted by the decay amplitude and averaged over the bin phase space. The values of \( c_i \) and \( s_i \) have been directly measured by the CLEO collaboration, exploiting quantum-correlated \( D^0\bar{D}^0 \) pairs produced at the \( \psi(3770) \) resonance [18].

The measurements of \( c_i \) and \( s_i \) are available in four different \( 2 \times 8 \) binning schemes for the \( D \to K^{0}_S\pi^+\pi^- \) decay. This analysis uses the ‘optimal binning’ scheme where the bins have been chosen to optimise the statistical sensitivity to \( \gamma \), as described in ref. [18]. The optimisation was performed assuming a strong-phase difference distribution as predicted by the BaBar model presented in ref. [23]. For the \( K^{0}_S K^+ K^- \) final state, three choices of binning schemes are available, containing \( 2 \times 2 \), \( 2 \times 3 \), and \( 2 \times 4 \) bins. The guiding model used to determine the bin boundaries is taken from the BaBar study described in ref. [24]. The \( 2 \times 2 \) binning scheme is chosen, due to the low signal yields in the \( D \to K^{0}_S K^+ K^- \) mode. The same choice of bins was used in the LHCb Run 1 analysis [9]. The measurements of \( c_i \) and \( s_i \) are not biased by the use of a specific amplitude model in defining the bin boundaries. The choice of the model only affects this analysis to the extent that a poor model description of the underlying decay would result in a reduced statistical sensitivity of the \( \gamma \) measurement. The binning choices for the two decay modes are shown in figure 1.

The physics parameters of interest, \( r_B \), \( \delta_B \), and \( \gamma \), are translated into four CP observables [22] that are measured in this analysis. These observables are defined as
\[
 x_\pm \equiv r_B \cos(\delta_B \pm \gamma) \quad \text{and} \quad y_\pm \equiv r_B \sin(\delta_B \pm \gamma). \tag{2.3}
\]
It follows from eq. (2.1) that the expected numbers of \( B^+ \) and \( B^- \) decays in bin \( i \), \( N_i^+ \) and \( N_i^- \), are given by
\[
 N_i^+ = h_{B+} \left[ F_{x,i} + (x_+^2 + y_+^2)F_{y,i} + 2\sqrt{F_{x,i}F_{y,i}(x_+s_{x,i} - y_+s_{y,i})} \right], \\
 N_i^- = h_{B-} \left[ F_{x,i} + (x_-^2 + y_-^2)F_{y,i} + 2\sqrt{F_{x,i}F_{y,i}(x_-s_{x,i} - y_-s_{y,i})} \right], \tag{2.4}
\]
where \( F_i \) are the fractions of decays in bin \( i \) of the \( D^0 \to K^{0}_S h^+ h^- \) Dalitz plot, and \( h_{B\pm} \) are normalisation factors, which can be different for \( B^+ \) and \( B^- \) due to production, detection,
Figure 1. Binning schemes for (left) \( D \to K_0^0 \pi^+ \pi^- \) decays and (right) \( D \to K_0^0 K^+ K^- \) decays. The diagonal line separates the positive and negative bins, where the positive bins are in the region in which \( m_2^2 > m_2^2 + m_+^2 \) is satisfied.

and \( CP \) asymmetries. In this measurement, the integrated yields are not used to provide information on \( x_\pm \) and \( y_\pm \), and so the analysis is insensitive to such effects. From eq. (2.4) it is seen that studying the distribution of candidates over the \( D \to K_0^0 h^+ h^- \) Dalitz plot gives access to the \( x_\pm \) and \( y_\pm \) observables. The detector and selection requirements placed on the data lead to a non-uniform efficiency over the Dalitz plot, which affects the \( F_i \) parameters. The efficiency profile for the signal candidates is denoted as \( \eta(m_2^2, m_+^2) \).

The parameters \( F_i \) can then be expressed as

\[
F_i = \frac{\int \int dm_2^2 dm_+^2 |A_D(m_2^2, m_+^2)|^2 \eta(m_2^2, m_+^2)}{\sum_j \int \int dm_2^2 dm_+^2 |A_D(m_2^2, m_+^2)|^2 \eta(m_2^2, m_+^2)}. \tag{2.5}
\]

The values of \( F_i \) are determined from the control decay mode \( \bar{B} \to D^{*\pm} \mu^+ \nu_\mu X \), where the \( D^{*\pm} \) meson decays to \( D^0 \pi^- \) and the \( D^0 \) meson decays to either the \( K_0^0 \pi^+ \pi^- \) or \( K_0^0 K^+ K^- \) final state. The symbol \( X \) indicates other particles which may be produced in the decay but are not reconstructed. Samples of simulated events are used to correct for the small differences in efficiency arising through unavoidable differences in selecting \( \bar{B} \to D^{*\pm} \mu^+ \nu_\mu X \) and \( B^{\pm} \to D K^{\pm} \) decays, as discussed further in section 5.

In addition to \( B^{\pm} \to D K^{\pm} \) and \( \bar{B} \to D^{*\pm} \mu^+ \nu_\mu X \) candidates, \( B^{\pm} \to D \pi^{\pm} \) decays are selected. These provide an important control sample that is used to constrain the invariant-mass shape of the \( B^{\pm} \to D K^{\pm} \) signal, as well as to determine the yield of \( B^{\pm} \to D \pi^{\pm} \) decays misidentified as \( B^{\pm} \to D K^{\pm} \) candidates. Note that this channel is not optimal for determining the values of \( F_i \) as the small level of \( CP \) violation in the decay leads to a significant systematic uncertainty, as was reported in ref. [30]. This uncertainty is eliminated when using the flavour-specific semileptonic decay, in favour of a smaller systematic uncertainty associated with efficiency differences.

The effect of \( D^0 - \bar{D}^0 \) mixing was ignored in the above discussion. If the parameters \( F_i \) are obtained from \( \bar{B} \to D^{*\pm} \mu^+ \nu_\mu X \), where the \( D^{*} \) decays to \( D^0 \pi^- \), \( D^0 - \bar{D}^0 \) mixing has been shown to lead to a bias of approximately 0.2° in the \( \gamma \) determination [31], which is
negligible for the current analysis. The effects of $CP$ violation in the neutral kaon system and of the different nuclear interaction cross-sections for $K^0$ and $\bar{K}^0$ mesons are discussed in section 7, where a systematic uncertainty is assigned.

3 Detector and simulation

The LHCb detector \cite{32, 33} 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, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about $4\,\text{Tm}$, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The polarity of the dipole magnet is reversed periodically throughout data-taking. The tracking system provides a measurement of momentum, $p$, of charged particles with 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 $(15+29/\sqrt{p_T})\,\mu\text{m}$, 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. Photons, electrons, and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The online event selection is performed by a trigger, 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 transverse momentum $p_T > 1.6$ GeV/$c$ and be inconsistent with originating from a PV. A multivariate algorithm \cite{34} is used for the identification of secondary vertices consistent with the decay of a $b$ hadron. Small changes in the trigger thresholds were made throughout both years of data taking.

In the simulation, $pp$ collisions are generated using PYTHIA 8 \cite{35, 36} with a specific LHCb configuration \cite{37}. Decays of hadronic particles are described by EvtGen \cite{38}, in which final-state radiation is generated using PHOTOS \cite{39}. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit \cite{40, 41} as described in ref. \cite{42}.

4 Event selection and fit to the invariant-mass spectrum for $B^\pm \rightarrow DK^\pm$ and $B^\pm \rightarrow D\pi^\pm$ decays

Decays of $K^0_S$ mesons to the $\pi^+\pi^-$ final state are reconstructed in two categories, the first containing $K^0_S$ mesons that decay early enough for the pions to be reconstructed
in the vertex detector and the second containing $K_S^0$ mesons that decay later such that track segments of the pions cannot be formed in the vertex detector. These categories are referred to as long and downstream, respectively. The candidates in the long category have better mass, momentum and vertex resolution than those in the downstream category. Hereinafter, $B^\pm$ candidates are denoted long or downstream depending on which category of $K_S^0$ candidate they contain.

For many of the quantities used in the selection and analysis of the data, a kinematic fit \cite{43} is imposed on the full $B^\pm$ decay chain. Depending on the quantity being calculated, the $D$ and $K_S^0$ candidates may be constrained to have their known masses \cite{44}, as described below. The fit also constrains the $B^\pm$ candidate momentum vector to point towards the associated PV, defined as the PV for which the candidate has the smallest IP significance. These constraints improve the resolution of the calculated quantities, and thus help improve separation between signal and background decays. Furthermore, it improves the resolution on the Dalitz plot coordinates and ensures that all candidates lie within the kinematically allowed $D \rightarrow K_S^0 h^+ h^-$ phase space.

The $D$ ($K_S^0$) candidates are required to be within 25 MeV/$c^2$ (15 MeV/$c^2$) of their known mass \cite{44}. These requirements are placed on masses obtained using kinematic fits in which all constraints are applied except for that on the mass under consideration. Combinatorial background is primarily suppressed through the use of a boosted decision tree (BDT) multivariate classifier \cite{45, 46}. The BDT is trained on simulated signal events and background taken from the high $B$ mass sideband (5800–7000 MeV/$c^2$). Separate BDTs are trained for the long and downstream categories.

Each BDT uses the same set of variables: the $\chi^2$ of the kinematic fit of the whole decay chain; $p$ and $p_T$ of the companion, $D$, and $B^\pm$ after the kinematic refit (here and in the following, companion refers to the final state $\pi^\pm$ or $K^\pm$ meson produced in the $B^\pm \rightarrow Dh^\pm$ decay); the vertex quality of the $K_S^0$, $D$, and $B^\pm$ candidates; the distance of closest approach between tracks forming the $D$ and $B^\pm$ vertices; the cosine of the angle between the momentum vector and the vector between the production and decay vertices of a given particle, for each of the $K_S^0$, $D$, and $B^\pm$ candidates; the minimum and maximum $\chi^2$IP of the pions from both the $D$ and $K_S^0$ decays, where $\chi^2$IP is defined as the difference in $\chi^2$ of the PV fit with and without the considered particle; the $\chi^2$IP for the companion, $K_S^0$, $D$, and $B^\pm$ candidates; the $B^\pm$ flight-distance significance; the radial distance from the beamline to the $D$ and $B^\pm$-candidate vertices; and a $B^\pm$ isolation variable, which is designed to ensure the $B^\pm$ candidate is well isolated from other tracks in the event. The $B^\pm$ isolation variable is the asymmetry between the $p_T$ of the signal candidate and the sum of the $p_T$ of other tracks in the event that lie within a distance of 1.5 in $\eta$–$\phi$ space, where $\phi$ is the azimuthal angle measured in radians. Candidates in the data samples that have a BDT output value below a threshold are rejected. An optimal threshold value is determined for each of the two BDTs, using a series of pseudoeperiments to obtain the values that provide the best sensitivity to $x_\pm$ and $y_\pm$. Across all $B^\pm \rightarrow DK^\pm$ channels this requirement is found to reject 99.1% of the combinatorial background in the high $B$ mass sideband that survives all other requirements, while having an efficiency of 92.4% on simulated $B^\pm \rightarrow DK^\pm$ signal samples.
Figure 2. Dalitz plots of long and downstream (left) $B^+ \to DK^+$ and (right) $B^- \to DK^-$ candidates for (top) $D \to K_s^0 \pi^+ \pi^-$ and (bottom) $D \to K_s^0 K^+ K^-$ decays in which the reconstructed invariant mass of the $B^\pm$ candidate is in a region of $\pm 25 \text{ MeV}/c^2$ around the $B^\pm$ mass. The narrow region is chosen to obtain high purity, as no background subtraction has been made. The Dalitz coordinates are calculated using the results of a kinematic fit in which the $D$ and $K^0_s$ masses are constrained to their known values. The blue lines show the kinematic boundaries of the decays.

Particle identification (PID) requirements are placed on the companion to separate $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ candidates, and on the charged decay products of the $D$ meson to remove cross-feed between different $D \to K_s^0 h^+ h^-$ decays. To ensure good control of the PID performance it is required that information from the RICH detectors is present. To remove background from $D \to \pi^+ \pi^- \pi^+ \pi^-$ or $D \to \pi^+ \pi^- K^+ K^-$ decays, long $K_s^0$ candidates are required to have travelled a significant distance from the $D$ vertex. This requirement is not necessary for downstream candidates. Similarly, the $D$ decay vertex is required be significantly displaced from the $B^\pm$ decay vertex in order to remove charmless $B^\pm$ decays.

The Dalitz plots for $B^\pm \to DK^\pm$ candidates in a narrow region of $\pm 25 \text{ MeV}/c^2$ around the $B^\pm$ mass are shown in figure 2, for both $D \to K_s^0 h^+ h^-$ final states samples. Separate
plots are shown for $B^+$ and $B^-$ decays. The Dalitz coordinates are calculated from the kinematic fit with all mass constraints applied.

In order to determine the parameterisation of the signal and background components that are used in the fit of partitioned regions of the Dalitz plot described in section 6, an extended maximum likelihood fit to the invariant-mass distributions of the $B^{\pm}$ candidates is performed, in which the $B^+$ and $B^-$ candidates in all of the Dalitz bins are combined. The invariant mass of each $B^{\pm}$ candidate is calculated using the results of a kinematic fit in which the $D$ and $K^0_S$ masses are constrained to their known values. The sample is split into $B^{\pm} \rightarrow DK^{\pm}$ and $B^{\pm} \rightarrow D\pi^{\pm}$ candidates, by $D$ decay mode and by $K^0_S$ category. In order to allow sharing of some parameters, the fit is performed simultaneously for all of the above categories. The projections of the fit and the invariant-mass distributions of the selected $B^{\pm}$ candidates are shown in figures 3 and 4 for $D \rightarrow K^0_S\pi^+\pi^-$ and $D \rightarrow K^0_SK^+K^-$ candidates, respectively. The fit range is between 5080 MeV/c$^2$ and 5800 MeV/c$^2$ in the $B^{\pm}$ candidate invariant mass.
Figure 4. Invariant-mass distributions of (left) $B^\pm \to DK^\pm$ and (right) $B^\pm \to D\pi^\pm$ candidates, with $D \to K^0_S K^+ K^-$, shown separately for the (top) long and (bottom) downstream $K^0_S$ categories. Fit results, including the signal component and background components due to misidentified companions, partially reconstructed decays and combinatorial background, are also shown.

The peaks corresponding to actual $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ candidates are fitted with a sum of two Crystal Ball [47] functions, which are parameterised as

$$\text{CB}(m, \mu, \sigma, \alpha, n) \propto \begin{cases} 
\exp \left[ -\frac{1}{2} \left( \frac{m - \mu}{\sigma} \right)^2 \right] & \text{if } (m - \mu)/\sigma > -\alpha \\
A \left( B + \frac{m - \mu}{\sigma} \right)^{-n} & \text{otherwise},
\end{cases} \quad (4.1)$$

where $\alpha > 0$, and

$$A = \left( \frac{n}{\alpha} \right)^n \exp[-\alpha^2/2], \quad (4.2)$$

$$B = \frac{n}{\alpha} - \alpha. \quad (4.3)$$

The sum is implemented such that the Crystal Ball functions have tails pointing in either direction. They share a common width, $\sigma$, and mean, $\mu$. In practice, the signal probability
density function (PDF) is defined as

\[ f_{\text{signal}}(m, \mu, \sigma, \alpha_L, n_L, \alpha_R, n_R, f_{\text{CB}}) = f_{\text{CB}} \cdot \text{CB}(m, \mu, \sigma, \alpha_L, n_L) + (1 - f_{\text{CB}}) \cdot \text{CB}(m, \mu, -\sigma, \alpha_R, n_R). \] (4.4)

The tail parameters, \( n_{L,R} \) and \( \alpha_{L,R} \), are fixed from simulation, while the other parameters are left as free parameters in the fit. Separate tail parameters, \( f_{\text{CB}} \), and \( \sigma \) are used for long and downstream candidates. Different widths are used for the \( B^\pm \to DK^\pm \) and \( B^\pm \to D\pi^\pm \) channels, with their ratio \( r_\sigma = \sigma_{DK}/\sigma_{D\pi} \) shared between all categories. The mean is shared among all categories. The yield of \( B^\pm \to D\pi^\pm \) decays in each \( K_S^0 \) and \( D \)-meson decay category, \( N_{\text{cat}}(B^\pm \to D\pi^\pm) \), is determined in the fit. Instead of fitting the yield of \( B^\pm \to DK^\pm \) decays directly in each category, it is determined from the \( B^\pm \to D\pi^\pm \) yield in the corresponding category and the ratio

\[ R = \frac{N_{\text{cat}}(B^\pm \to DK^\pm)}{\varepsilon_{\text{PID}}(B^\pm \to DK^\pm)} / \frac{N_{\text{cat}}(B^\pm \to D\pi^\pm)}{\varepsilon_{\text{PID}}(B^\pm \to D\pi^\pm)}, \] (4.5)

which is a free parameter in the fit. The category-dependent PID efficiencies, \( \varepsilon_{\text{PID}}(B^\pm \to Dh^\pm) \), are taken into account, so that a single \( R \) parameter can be shared between all categories in the fit. How these efficiencies are obtained is described below. As the parameter \( R \) is efficiency corrected, it is equal to the ratio of branching fractions between the \( B^\pm \to DK^\pm \) and \( B^\pm \to D\pi^\pm \) decay modes. The measured ratio is found to be \( R = (7.66 \pm 0.14) \% \), where the uncertainty is statistical only, and this is consistent with the expected value of \( (7.8 \pm 0.4) \% \) [44].

The background consists of random track combinations, partially reconstructed \( B \) decays, and \( B^\pm \to Dh^\pm \) decays in which the companion has been misidentified. The random track combinations are modelled by an exponential PDF. The slopes of the exponentials are free parameters in the fit to the data. These slopes are independent for each of the \( B^\pm \to D\pi^\pm \) categories, while they are shared for the \( B^\pm \to DK^\pm \) categories to improve the stability of the fit. When these slopes are allowed to be independent, the fit returns results that are statistically compatible.

In the \( B^\pm \to DK^\pm \) sample there is a clear contribution from \( B^\pm \to D\pi^\pm \) decays in which the companion particle is misidentified as a kaon by the RICH system. The rate for \( B^\pm \to DK^\pm \) decays to be misidentified and placed in the \( B^\pm \to D\pi^\pm \) sample is much lower due to the reduced branching fraction. Nevertheless, this contribution is still accounted for in the fit. The yields of these backgrounds are fixed in the fit, using knowledge of misidentification efficiencies and the fitted yields of reconstructed decays with the correct particle hypothesis. The misidentification efficiencies are obtained from large samples of \( D^{\pm \to \bar{D}^0\pi^\pm}, \bar{D}^0 \to K^\mp\pi^\pm \) decays. These decays are selected using only kinematic variables in order to provide pure samples of \( K^\pm \) and \( \pi^\pm \) that are unbiased in the PID variables. The PID efficiency is parameterised as a function of the companion momentum and pseudorapidity, and the charged-particle multiplicity in the event. The calibration sample is weighted so that the distribution of these variables matches that of the candidates in the signal region of the \( B^\pm \) sample, thereby ensuring that the measured PID performance
is representative for the samples used in this measurement. The efficiency to identify a kaon correctly is found to be approximately 86\%, while that for a pion is approximately 97\%. The PDFs of the backgrounds due to misidentified companion particles are determined using data. As an example, consider the case of true $B^\pm \rightarrow D\pi^\pm$ decays misidentified as $B^\pm \rightarrow DK^\pm$ candidates. The sPlot method [48] is used on the $B^\pm \rightarrow D\pi^\pm$ sample in order to isolate the contribution from the signal decays. The $B^\pm$ invariant mass is then calculated using the kaon mass hypothesis for the companion pion, and weighting by PID efficiencies in order to properly reproduce the kinematic properties of pions misidentified as kaons in the signal $B^\pm \rightarrow DK^\pm$ sample. The weighted distribution is fitted with a sum of two Crystal Ball shapes. The fitted parameters are subsequently fixed in the fit to the $B^\pm$ invariant-mass spectrum, with the procedure applied separately for long and downstream candidates. An analogous approach is used to determine the shape of the misidentified $B^\pm \rightarrow DK^\pm$ contribution in the $B^\pm \rightarrow D\pi^\pm$ sample.

Partially reconstructed $b$-hadron decays contaminate the sample predominantly at invariant masses smaller than that of the signal peak. These decays contain an unconstructed pion or photon, which predominantly comes from an intermediate resonance. There are contributions from $B^0 \rightarrow D^{*\pm}h^\mp$ and $B^\pm \rightarrow D^{*0}h^\pm$ decays in all channels (denoted as $B \rightarrow D^{*h}$ decays), while $B^\pm \rightarrow D\rho^\pm$ and $B^\pm \rightarrow DK^{*\pm}$ decays contribute to the $B^\pm \rightarrow D\pi^\pm$ and $B^\pm \rightarrow DK^\pm$ channels, respectively. In the $B^\pm \rightarrow DK^\pm$ channels there is also a contribution from $B^0_s \rightarrow \bar{D}^0\pi^+K^- (\bar{B}^0_s \rightarrow D^0\pi^-K^+)$ decays where the charged pion is not reconstructed. The invariant-mass distributions of these backgrounds depend on the spin and mass of the missing particle, as described in ref. [49]. The shape of the background from $B^0_s$ decays is based on the results of ref. [50]. Additionally, each of the above backgrounds of $B^\pm \rightarrow D\pi^\pm$ decays can contribute in the $B^\pm \rightarrow DK^\pm$ channels if the pion is misidentified. The inverse contribution is negligible and is neglected. The shapes for the decays in which a pion is misidentified as a kaon are parameterised with semi-empirical PDFs formed from sums of Gaussian and error functions. The parameters of these backgrounds are fixed to the results of fits to data from two-body $D$ decays [49], where they were obtained with a much larger data sample. However, the width of the resolution function and a shift along the $B$ mass are allowed to differ in order to accommodate small differences between the $D$ decay modes.

In each of the $B^\pm \rightarrow D\pi^\pm$ channels, the total yield of the partially reconstructed background is fitted independently. The relative amount of each $B \rightarrow D^*\pi^\mp$ mode is fixed from efficiencies obtained from simulation and known branching fractions, while the fraction of $B^\pm \rightarrow D\rho^\pm$ decays is left free. In the $B^\pm \rightarrow DK^\pm$ channels, the yield of the $B^0_s \rightarrow \bar{D}^0\pi^+K^-$ background is fixed relative to the corresponding $B^\pm \rightarrow D\pi^\mp$ yield, using efficiencies from simulation and the known branching fraction. The total yield of the remaining partially reconstructed backgrounds is expressed via a single fraction, $R_{DK/D\pi^\pm}^{low}$, relative the $B^\pm \rightarrow D\pi^\pm$ yields. It is free in the fit, and common to all channels after taking into account the different particle-identification efficiencies. The relative amount of each $B \rightarrow D^*K^\mp$ mode is fixed using efficiencies from simulation and known branching fractions, while the fraction of $B^\pm \rightarrow DK^{*\mp}$ decays is fixed using the results of ref. [49]. The yields of the partially reconstructed modes with a companion pion misidentified as
a kaon are fixed via the known PID efficiencies, based on the fitted yield of the partially reconstructed backgrounds in the corresponding $B^{\pm} \to D\pi^{\pm}$ channel.

In the $B^{\pm} \to D\pi^{\pm}$ channels, a total signal yield of approximately 56 100 (7750) is found in the signal region 5249–5319 MeV/$c^2$ of the $D \to K_s^0\pi^+\pi^-$ ($D \to K_s^0K^+K^-$) channel, 31\% (32\%) of which is in the long $K_s^0$ category. The purity in the signal region is found to be 98.4\% (97.7\%), with the dominant background being combinatorial. In the $B^{\pm} \to DK^{\pm}$ channels, a total signal yield of approximately 3900 (530) is found in the signal region of the $D \to K_s^0\pi^+\pi^-$ ($D \to K_s^0K^+K^-$) channel, again finding 31\% (32\%) of the candidates in the long $K_s^0$ category. The purity in the signal region is found to be 81\%. The dominant background is from misidentified $B^{\pm} \to D\pi^{\pm}$ decays, which accounts for 66\% of the background in the signal region. Equal amounts of combinatorial background and partially reconstructed decays, predominantly including a misidentified companion pion, make up the remaining background.

5 Event selection and yield determination for $\bar{B} \to D^{*\mp}\mu^{\mp}\nu_{\mu}$ decays

A sample of $\bar{B} \to D^{*\mp}\mu^{\mp}\nu_{\mu}X$, $D^{\pm} \to \bar{D}\pi^{\pm}$, $\bar{D} \to K_s^0h^+h^-$ decays is used to determine the quantities $F_i$, defined in eq. (2.5), as the expected fractions of $D^0$ decays falling into the $i$th Dalitz plot bin, taking into account the efficiency profile of the signal decay. The semileptonic decay of the $B$ meson and the strong-interaction decay of the $D^{*\pm}$ meson allow the flavour of the $D^0$ meson to be determined from the charges of the muon and the soft pion from the $D^{*\pm}$ decay. This particular decay chain, involving a flavour-tagged $D^0$ decay, is chosen due to its high yield, low background level, and low mistag probability. The selection requirements are chosen to minimise changes to the efficiency profile with respect to those associated with the $B^{\pm} \to DK^{\pm}$ channels.

The selection is identical to that applied in ref. [9], except for a tighter requirement on the significance of the $D^0$ flight distance that helps to suppress backgrounds from charmless $B$ decays. To improve the resolution of the distribution of candidates across the Dalitz plot, the $B$-decay chain is refitted [43] with the $D^0$ and $K_s^0$ candidates constrained to their known masses. An additional fit, in which only the $K_s^0$ mass is constrained, is performed to improve the $D^0$ and $D^{*\pm}$ mass resolution in the invariant-mass fit used to determine signal yields.

The invariant mass of the $D^0$ candidate, $m(K_s^0h^+h^-)$, and the invariant-mass difference, $\Delta m \equiv m(K_s^0h^+h^-) - m(K_s^0h^+h^-)$, are fitted simultaneously to determine the signal yields. This two-dimensional parameterisation allows the yield of selected candidates to be measured in three categories: true $D^{*\pm}$ candidates (signal), candidates containing a true $D^0$ meson but random soft pion (RSP) and candidates formed from random track combinations that fall within the fit range (combinatorial background). Background contributions from real $D^{*\pm}$ decays paired with a random $\mu$ are determined to be negligible by selecting pairs of $D^{*\pm}$ mesons and $\mu^{\pm}$ with the same charge.

An example projection of $m(K_s^0\pi^+\pi^-)$ and $\Delta m$ is shown in figure 5. The result of a two-dimensional extended, unbinned, maximum likelihood fit is superimposed. The fit is performed simultaneously for the two $D^0$ final states and the two $K_s^0$ categories with some
parameters allowed to be independent between categories. Candidates selected from data recorded in 2015 and 2016 are fitted separately, in order to accommodate different trigger threshold settings that result in slightly different Dalitz plot efficiency profiles. The fit region is defined by $1830 < m(K_{S}^{0}h^{+}h^{-}) < 1910 \text{MeV/c}^2$ and $139.5 < \Delta m < 153.0 \text{MeV/c}^2$. Within this $m(K_{S}^{0}h^{+}h^{-})$ range, the $\Delta m$ resolution does not vary significantly.

The signal is parameterised in $\Delta m$ with a sum of two Crystal Ball functions, as for the $B^{\pm} \to Dh^{\pm}$ signal. The mean, $\mu$, is shared between all categories, while the other parameters are different for long and downstream candidates. The tail parameters are fixed from simulation. The combinatorial and RSP backgrounds are both parameterised with an empirical model given by

$$f(\Delta m; \Delta m_0, x, p_1, p_2) = \left[1 - \exp\left(-\frac{\Delta m - \Delta m_0}{x}\right)\right]\left(\frac{\Delta m}{\Delta m_0}\right)^{p_1} + p_2 \left(\frac{\Delta m}{\Delta m_0} - 1\right)$$  \hspace{1cm} (5.1)$$

for $\Delta m - \Delta m_0 > 0$ and $f(\Delta m) = 0$ otherwise, where $\Delta m_0$, $x$, $p_1$, and $p_2$ are free parameters. The parameter $\Delta m_0$, which describes the kinematic threshold for a $D^{*\pm} \to \overline{D}^{0}\pi^{\pm}$ decay, is shared in all data categories and for both the combinatorial and RSP shapes. The remaining parameters are determined separately for $D \to K_{S}^{0}\pi^{+}\pi^{-}$ and $D \to K_{S}^{0}K^{+}K^{-}$ candidates.

In the $m(K_{S}^{0}h^{+}h^{-})$ fit, all of the parameters in the signal and RSP PDFs are constrained to be the same as both describe a true $D^{0}$ candidate. These are also fitted with a sum of two Crystal Ball functions, with the tail parameters fixed from simulation. The parameters are fitted separately for the $D \to K_{S}^{0}\pi^{+}\pi^{-}$ and $D \to K_{S}^{0}K^{+}K^{-}$ shapes, due to the different phase space available in the $D^{0}$ decay. The combinatorial background is parameterised by an exponential function in $m(K_{S}^{0}h^{+}h^{-})$.

A total signal yield of approximately $113000 \ (15000)$ $D \to K_{S}^{0}\pi^{+}\pi^{-} \ (D \to K_{S}^{0}K^{+}K^{-})$ decays is obtained. This is approximately 25 times larger than the $B^{\pm} \to DK^{\pm}$ yield. In the range surrounding the signal peaks, defined as $1840$–$1890 \ (1850$–$1880)$ MeV/c$^2$ in $m(K_{S}^{0}\pi^{+}\pi^{-}) \ (m(K_{S}^{0}K^{+}K^{-}))$ and $143.9$–$146.9 \text{MeV/c}^2$ in $\Delta m$, the background components account for 2–5% of the yield depending on the category.

Figure 5. Result of the simultaneous fit to $\overline{B} \to D^{*\pm}\mu^{\mp}\nu_{\mu}$, $D^{\pm} \to \overline{D}^{0} \to K_{S}^{0}\pi^{+}\pi^{-}\pi^{\pm}$ decays with downstream $K_{S}^{0}$ candidates, in 2016 data. The projections of the fit result are shown for (left) $m(K_{S}^{0}\pi^{+}\pi^{-})$ and (right) $\Delta m$. The (blue) total fit PDF is the sum of components describing (solid red) signal, (dashed black) combinatorial background and (dotted green) random soft pion background.
Figure 6. Example efficiency profiles of (left) $B^\pm \to D\pi^\pm$ and (right) $\overline{B} \to D^{\ast\pm}\mu^\mp\overline{\nu}_\mu X$ decays in the simulation. The top (bottom) plots are for $D \to K^0_{\pi^\pm}\pi^-$ ($D \to K^0_{\pi^\pm}\pi^+\pi^-$) decays. These plots refer to downstream $K^0$ candidates under 2016 data taking conditions. The normalisation is chosen so that the average over the Dalitz plot is unity.

The two-dimensional fit in $m(K^0_{\pi^+}\pi^-)$ and $\Delta m$ of the $\overline{B} \to D^{\ast\pm}\mu^\mp\overline{\nu}_\mu X$ decay is repeated in each Dalitz plot bin with all of the PDF parameters fixed, resulting in a raw control-mode yield, $R_i$, for each bin $i$. The measured $R_i$ are not equivalent to the $F_i$ fractions required to determine the $CP$ parameters due to unavoidable differences from selection criteria in the efficiency profiles of the signal and control modes. Examples of the efficiency profiles from simulation of the downstream candidates in 2016 data are shown in figure 6. For each Dalitz plot bin $i$ a correction factor $\xi_i$ is determined to account for these efficiency differences, defined as

$$\xi_i = \frac{\int_i dm^2_{\pi^\pm} dm^2_{\pi^-} |A_D(m^2_{\pi^-}, m^2_{\pi^\pm})|^2 \eta(m^2_{\pi^-}, m^2_{\pi^\pm})_{D\pi}}{\int_i dm^2_{\pi^\pm} dm^2_{\pi^-} |A_D(m^2_{\pi^-}, m^2_{\pi^\pm})|^2 \eta(m^2_{\pi^-}, m^2_{\pi^\pm})_{D\mu}}$$

(5.2)

where $\eta(m^2_{\pi^-}, m^2_{\pi^\pm})_{D\pi}$ and $\eta(m^2_{\pi^-}, m^2_{\pi^\pm})_{D\mu}$ are the efficiency profiles of the $B^\pm \to D\pi^\pm$ and $\overline{B} \to D^{\ast\pm}\mu^\mp\overline{\nu}_\mu X$ decays, respectively, and are determined from simulation. The $B^\pm \to$
$D\pi^\pm$ decay mode is used rather than $B^\pm \to DK^\pm$ as the simulation is more easily compared to the data, due to the larger decay rate and the smaller interference between $B^\pm \to D^0\pi^\pm$ and $B^\pm \to D^0\pi^\mp$ decays, compared to in the $B^\pm \to DK^\pm$ decay mode. It is verified using simulation that the efficiency profiles of the $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ decays are the same. The simulated events are generated with a flat distribution across the $D \to K^0_S h^+h^-$ phase space; hence the distribution of simulated events after triggering, reconstruction and selection is directly proportional to the efficiency profile. The amplitude models used to determine the Dalitz plot intensity for the correction factor are those from ref. [23] and ref. [24] for the $D \to K^0_S \pi^+\pi^-$ and $D \to K^0_S K^+K^-$ decays, respectively. The amplitude models provide a description of the intensity distribution over the Dalitz plot and introduce no significant model dependence into the analysis. The $F_i$ values can be determined via the relation $F_i = h^i\xi_iR_i$, where $h^i$ is a normalisation factor such that the sum of all $F_i$ is unity. The $F_i$ values are determined separately for each year of data taking and $K^0_S$ category and are then combined in the fractions observed in the $B^\pm \to D\pi^\pm$ signal region in data. This method of determining the $F_i$ parameters is preferable to using solely the amplitude models and $B^\pm \to D\pi^\pm$ simulated events, since the method is data-driven. The amplitude models and simulation data enter the correction factor as a ratio, and thus imperfections in the simulation and the model cancel at first order. The average correction factor over all bins is approximately 2% from unity and the largest correction factor is within 7%. Uncertainties on these correction factors are driven by the size of the simulation samples and are of a similar size as the corrections themselves.

6 Dalitz plot fit to determine the CP-violating parameters $x_\pm$ and $y_\pm$

The Dalitz plot fit is used to measure the CP-violating parameters $x_\pm$ and $y_\pm$, as introduced in section 2. Following eq. (2.4), these parameters are determined from the populations of the $B^+$ and $B^-$ Dalitz plot bins, given the external information of the $c_i$ and $s_i$ parameters from CLEO-c data and the values of $F_i$ from the semileptonic control decay modes. Although the absolute numbers of $B^+$ and $B^-$ decays integrated over the $D$ Dalitz plot have some dependence on $x_\pm$ and $y_\pm$, the sensitivity gained compared to using just the relations in eq. (2.4) is negligible [51] given the available sample size. Consequently, as stated previously, the integrated yields are not used to provide information on $x_\pm$ and $y_\pm$ and the analysis is insensitive to $B$ meson production and detection asymmetries.

A simultaneous fit is performed on the $B^\pm \to Dh^\pm$ data, split into the two $B$ charges, the two $K^0_S$ categories, the $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ candidates, and the two $D \to K^0_S h^+h^-$ final states. The invariant mass of each $B^\pm$ candidate is calculated using the results of a kinematic fit in which the $D$ and $K^0_S$ masses are constrained to their known values. Each category is then divided into the Dalitz plot bins shown in figure 1, where there are 16 bins for $D \to K^0_S \pi^+\pi^-$ and 4 bins for $D \to K^0_S K^+K^-$. The $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ samples are fitted simultaneously because the yield of $B^\pm \to D\pi^\pm$ signal in each Dalitz plot bin is used to determine the yield of misidentified candidates in the corresponding $B^\pm \to DK^\pm$ Dalitz plot bin. The PDF parameters for both the signal and background invariant-mass distributions are fixed to the values determined in the invariant-
mass fit described in section 4. The $B^\pm$ mass range is reduced to 5150–5800 GeV/$c^2$ to avoid the need of a detailed description of the shape of the partially reconstructed background. The yields of signal candidates for each bin in the $B^\pm \to D\pi^\pm$ sample are free parameters. In each of the $B^\pm \to DK^\pm$ channels, the total yield integrated over the Dalitz plot is a free parameter. The fractional yields in each bin are defined using the expressions for the Dalitz plot distribution in terms of $x_\pm$, $y_\pm$, $F_i$, $c_i$, and $s_i$ in eq. (2.4), where the $x_\pm$ and $y_\pm$ parameters are free and the values of $F_i$ are Gaussian-constrained within their uncertainties. The values of $c_i$ and $s_i$ are fixed to their central values, which is taken into account as a source of systematic uncertainty. The yields of the component due to $B^\pm \to D\pi^\pm$ decays, where the companion has been misidentified as a kaon, are fixed in each $B^\pm \to DK^\pm$ bin, relative to the yield in the corresponding $B^\pm \to D\pi^\pm$ bin, using the known PID efficiencies. A component for misidentified $B^\pm \to DK^\pm$ decays in the $B^\pm \to DK^\pm$ channels is not included, as it is found to contribute less than 0.5% of the yield in the signal region in the global fit described in section 4. The total yield of the partially reconstructed $B^\pm$ and $B^0$ backgrounds is fitted in each bin, using the same shape in all bins, with the fractions of each component taken from the global fit. The total yield of the $B^0_s \to D^0\pi^+K^-$ ($D^0_s \to D^0\pi^-K^+$) background is fixed in each channel, using the results of the global fit. The yield in each bin is then fixed from the $F_i$ parameters, using the known Dalitz distribution of $\bar{D}^0(D^0) \to K^0_S h^+h^-$ decays. The separate treatment of the partially reconstructed background from $B^0_s$ decays is necessary due to the significantly different Dalitz distribution, arising because only a $D^0$ meson is produced along with a $K^-$ meson, while for the remaining modes, the $D$ meson is either a $D^0$ meson or an admixture where the $\bar{D}^0$ component is $r_B$-suppressed. The yield of the combinatorial background in each bin is a free parameter. In bins in which an auxiliary fit determines the yield of the partially reconstructed or combinatorial background to be negligible, the corresponding yields are set to zero to facilitate the calculation of the covariance matrix [52, 53].

A large ensemble of pseudoexperiments is performed to validate the fit procedure. In each pseudoexperiment the numbers and distributions of signal and background candidates are generated according to the expected distribution in data, and the full fit procedure is then executed. The input values for $x_\pm$ and $y_\pm$ correspond to $\gamma = 70^\circ$, $r_B = 0.1$, and $\delta_B = 130^\circ$. The uncertainties determined by the fit to data are consistent with the size of the uncertainties determined by the pseudoexperiments. Small biases are observed in the central values and are due to the low event yields in some of the bins. These biases are observed to decrease in simulated experiments of larger size. The central values are corrected for the biases and a systematic uncertainty is assigned, as described in section 7.

The $CP$ parameters obtained from the fit are

$$
\begin{align*}
x_- &= (9.0 \pm 1.7) \times 10^{-2}, \\
y_- &= (2.1 \pm 2.2) \times 10^{-2}, \\
x_+ &= (-7.7 \pm 1.9) \times 10^{-2}, \\
y_+ &= (-1.0 \pm 1.9) \times 10^{-2},
\end{align*}
$$

where the uncertainties are statistical only. The correlation matrix is shown in table 1. The total $B^\pm \to DK^\pm$ yields in the signal region, where the invariant mass of the $B$ candidate
Table 1. Statistical correlation matrix for the fit to data.

<table>
<thead>
<tr>
<th></th>
<th>$x_-$</th>
<th>$y_-$</th>
<th>$x_+$</th>
<th>$y_+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_-$</td>
<td>1</td>
<td>-0.21</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>$y_-$</td>
<td>1</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$x_+$</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$y_+$</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Fit results for the total $B^\pm \rightarrow DK^\pm$ yields in the signal region, where the invariant mass of the $B$ candidate is in the interval 5249–5319 MeV/$c^2$, integrated over the Dalitz plots.

<table>
<thead>
<tr>
<th></th>
<th>$D \rightarrow K_S^0 \pi^+ \pi^-$</th>
<th>$D \rightarrow K_S^0 K^+ K^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Downstream</td>
<td>602 ± 26</td>
<td>1315 ± 39</td>
</tr>
<tr>
<td>Long Downstream</td>
<td>606 ± 26</td>
<td>1334 ± 39</td>
</tr>
<tr>
<td>Long Downstream</td>
<td>92 ± 10</td>
<td>189 ± 15</td>
</tr>
<tr>
<td>Long Downstream</td>
<td>82 ± 10</td>
<td>193 ± 15</td>
</tr>
</tbody>
</table>

Figure 7. Confidence levels at 68.2%, 95.5% and 99.7% probability for $(x_+, y_+)$ and $(x_-, y_-)$ as measured in $B^\pm \rightarrow DK^\pm$ decays (statistical uncertainties only). The parameters $(x_+, y_+)$ relate to $B^+$ decays and $(x_-, y_-)$ refer to $B^-$ decays. The black dots show the central values obtained in the fit.

is in the interval 5249–5319 MeV/$c^2$, are shown in table 2.

The measured values of $(x_\pm, y_\pm)$ from the fit to data are displayed in figure 7, along with their likelihood contours, corresponding to statistical uncertainties only. The systematic uncertainties are discussed in the next section. The two vectors defined by the coordinates $(x_-, y_-)$ and $(x_+, y_+)$ are not consistent with zero magnitude and they have a non-zero opening angle. Therefore the data sample exhibits the expected features of $CP$ violation. The opening angle is equal to $2\gamma$, as illustrated in figure 7.

In order to assess the goodness of fit, and to demonstrate that the equations in $(x_\pm, y_\pm)$ provide a good description of data, an alternative fit is performed where the $B^\pm \rightarrow DK^\pm$ yields are measured independently in each bin. In figure 8 (left) the obtained yields are compared with the yields predicted from the values of $(x_\pm, y_\pm)$ obtained in the default
Figure 8. (Left) Comparison of total signal yields from the direct fit (points) to those calculated from the central values of $x_{\pm}$ and $y_{\pm}$ (solid line). The yields are given for the effective bin: $+i$ for $B^+$ and $-i$ for $B^-$, and summed over $B$ charge and $K^0_S$ decay category. (Right) Comparison of the difference between the $B^+$ and $B^-$ yield obtained in the direct fit for each effective bin (points), the prediction from the central values of $x_{\pm}$ and $y_{\pm}$ (solid line), and the prediction assuming no CP violation (dotted line).

The yields from the direct fit agree with the prediction with a $p$-value of 0.33. In figure 8 (right) the difference $N_{B^+} - N_{B^-}$ in each bin is calculated using the results of the direct fit of the $B^{\pm} \to DK^{\mp}$ yields. This distribution is compared to that predicted by the central $(x_{\pm}, y_{\pm})$ values. The measured yield differences are compatible with the prediction with a $p$-value of 0.58. In addition, data are fitted with the assumption of no CP violation by enforcing $x_+ = x_- \equiv x_0$ and $y_+ = y_- \equiv y_0$. The obtained $x_0$ and $y_0$ values are used to determine the predicted values of $N_{B^+} - N_{B^-}$, which are also shown in figure 8 (right). This prediction is not zero because the $B$ meson production and various detection effects can induce a global asymmetry in the measured yields. The comparison of the data to this hypothesis yields a $p$-value of $1 \times 10^{-6}$, which strongly disfavours the CP-conserving hypothesis.

7 Systematic uncertainties

Systematic uncertainties on the measurements of the $x_{\pm}$ and $y_{\pm}$ parameters are evaluated and are presented in table 3. The source of each systematic uncertainty is described below. The systematic uncertainties are generally determined from an ensemble of pseudoexperiments where the simulated data are generated in an alternative configuration and fitted with the default method. The mean shifts in the fitted values of $x_{\pm}$ and $y_{\pm}$ in comparison to their input values are taken as the systematic uncertainty.

The limited precision on $(c_i, s_i)$ coming from the CLEO measurement induces uncertainties on $x_{\pm}$ and $y_{\pm}$ [18]. These uncertainties are evaluated by fitting the data multiple times, each with different $(c_i, s_i)$ values sampled according to their experimental uncertainties and correlations. The resulting widths in the distributions of $x_{\pm}$ and $y_{\pm}$ values are assigned as the systematic uncertainties. Values of $(0.4-1.1) \times 10^{-2}$ are found for the fit to the full sample. The uncertainties are similar to, but different from, those reported in
Table 3. Summary of uncertainties for the parameters $x_\pm$ and $y_\pm$. The various sources of systematic uncertainties are described in the main text. All entries are given in multiples of $10^{-2}$.

<table>
<thead>
<tr>
<th>Source</th>
<th>$x_-$</th>
<th>$y_-$</th>
<th>$x_+$</th>
<th>$y_+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical</td>
<td>1.7</td>
<td>2.2</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Efficiencies corrections</td>
<td>0.6</td>
<td>0.2</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>Mass fit PDFs</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Different mis-ID shape over Dalitz plot</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Different low mass shape over Dalitz plot</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Uncertainty on $B_s^0 \to D^{0}\pi^+K^-$ yield</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Bias correction</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Bin migration</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$K^0$ CP violation and material interaction</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Total experimental systematic uncertainty</td>
<td>0.7</td>
<td>0.5</td>
<td>0.7</td>
<td>0.4</td>
</tr>
</tbody>
</table>

ref. [9]. This is as expected since it is found from simulation studies that the $(c_i, s_i)$-related uncertainty depends on the particular sample under study. It is found that the uncertainties do become constant when simulated samples with very high signal yields are studied. The uncertainties arising from the CLEO measurements are kept separate from the other experimental uncertainties.

A systematic uncertainty arises from imperfect modeling in the simulation used to derive the efficiency correction for the determination of the $F_i$ parameters. As the simulation enters the correction in a ratio, it is expected that imperfections cancel to first order. To determine the residual systematic uncertainty associated with this correction, an additional set of correction factors is calculated and used to evaluate an alternative set of $F_i$ parameters. To determine this additional factor, a new rectangular binning scheme is used, which is shown in figure 9. The bin-to-bin efficiency variation in this rectangular scheme is significantly larger than for the default partitioning and is more sensitive to imperfections in the simulated data efficiency profile. The yields of the $B^\pm \to D\pi^\pm$ and $B^\pm \to D^{*\pm}\mu^\mp\nu_\mu X$ decays in each bin of the rectangular scheme are compared to the predictions from the amplitude model and the simulated data efficiency profile. The usage of the rectangular binning also helps to dilute the small level of $CP$ violation in $B^\pm \to D\pi^\pm$ such that differences from this comparison will come primarily from efficiency effects. The alternative correction factors $\xi_i^{\text{alt}}$ are calculated as

$$\xi_i^{\text{alt}} = \frac{\int dm_2^2 dm_2^2 \eta(m_2^2, m_2^2) |A_D(m_2^2, m_2^2)|^2 C_D(m_2^2, m_2^2)}{\int dm_2^2 dm_2^2 \eta(m_2^2, m_2^2) |A_D(m_2^2, m_2^2)|^2 C_D(m_2^2, m_2^2)}, \quad (7.1)$$

where the $C(m_2^2, m_2^2)$ terms are the ratios between the predicted and observed data yields in the rectangular bins. Many pseudoexperiments are performed, in which the data are generated according to the alternative $F_i$ parameters and then fitted with the default $F_i$ parameters. The overall shift in the fitted values of the $CP$ parameters in comparison to
Figure 9. Rectangular binning schemes for (left) $D \to K^0\pi^+\pi^-$ decays and (right) $D \to K^0_s K^+ K^-$ decays. The diagonal line separates the positive and negative bins, where the positive bins are in the region in which $m_2^+ > m_2^-$ is satisfied.

Their input values is taken as the systematic uncertainty, yielding $0.6 \times 10^{-2}$ for $x_\pm$ and $0.1(0.2) \times 10^{-2}$ for $y_+ (y_-)$.

Various effects are considered to assign an uncertainty for the imperfections in the description of the invariant-mass spectrum. For the PDF used to fit the signal, the parameters of the PDF used in the binned fit are varied according to the uncertainties obtained in the global fit. An alternative shape is also tested. The global fit is repeated with the mean and width of the shape used to describe the background due to misidentified companions allowed to vary freely. The results are used to generate data sets with an alternative PDF, and fit them using the default setup. The description of the partially reconstructed background is changed to a shape obtained from a fit of the PDF to simulated decays. The slope of the exponential used to fit the combinatorial background is also fluctuated according to the uncertainty obtained in the global fit. The contributions from each change are summed in quadrature and are $0.2 \times 10^{-2}$ for each of the $x_\pm$ parameters and $0.3 \times 10^{-2}$ for each of the $y_\pm$ parameters.

Two systematic uncertainties associated with the misidentified $B^\pm \to D \pi^\pm$ background in the $B^\pm \to DK^\pm$ sample are considered. First, the uncertainties on the particle misidentification probabilities are found to have a negligible effect on the measured values of $x_\pm$ and $y_\pm$. Second, it is possible that the invariant-mass distribution of the misidentified background (the mis-ID shape) is not uniform over the Dalitz plot, as assumed in the fit. This can occur through kinematic correlations between the reconstruction efficiency across the Dalitz plot of the $D$ decay and the momentum of the companion pion from the $B^\pm$ decay. Alternative mass shapes are constructed by repeating the procedure used to obtain the default shape for each Dalitz bin individually. The alternative shapes are used when generating data sets for pseudoexperiments, and the fits then performed assuming a single shape, as in the fit to data. The resulting uncertainty is at most $0.2 \times 10^{-2}$ for all $CP$ parameters.

In the fit to the data, the relative contributions of the partially reconstructed $B^\pm$
and $B^0$ backgrounds are kept the same in each Dalitz bin. This is a simplification as some partially reconstructed backgrounds will be distributed as $D^0(\bar{D}^0) \rightarrow K^0_S h^+ h^-$ for reconstructed $B^- (B^+)$ candidates, while partially reconstructed $B^\pm \rightarrow D^{(*)} K^{(*)\pm}$ decays will be distributed as a $D^0 - \bar{D}^0$ admixture depending on the relevant CP violation parameters. Pseudoexperiments are generated, where the $D$-decay Dalitz plot distribution for $B^\pm \rightarrow D^* K^+$ is based on the CP parameters reported in ref. [54] and those for $B^\pm \rightarrow D K^{*\pm}$ are taken from ref. [55]. The generated samples are fitted with the standard method. The resulting uncertainty is at most $0.2 \times 10^{-2}$ for all CP parameters.

The total yield of the $B^0_s \rightarrow D^0\pi^+ K^-$ background in the $B^\pm \rightarrow D K^\pm$ channels is fixed relative to the corresponding $B^\pm \rightarrow D_s \pi^\pm$ yield. The systematic uncertainty due to the uncertainty on the relative rate is estimated via pseudoexperiments, where data sets are generated with the rate varied by $\pm 1\sigma$ and fitted using the default value. The maximal mean bias for each parameter is taken as the uncertainty. The resulting uncertainty is $0.1 \times 10^{-2}$ for all CP parameters.

An uncertainty is assigned to each CP parameter to accompany the correction that is applied for the small bias observed in the fit procedure. These uncertainties are determined by performing sets of pseudoexperiments, each generated with different values of $x_\pm$ and $y_\pm$ throughout a range around the values predicted by the world averages. The spread in observed bias is combined in quadrature with the uncertainty in the precision of the pseudoexperiments. This is taken as the systematic uncertainty and is $0.1 \times 10^{-2}$ for all CP parameters.

The systematic uncertainty from the effect of candidates being assigned the wrong Dalitz bin number is considered. The resolution in $m_\pm^2$ and $m_\mp^2$ is approximately 0.006 GeV$^2$/c$^4$ for candidates with long $K^0_S$ decays and 0.007 GeV$^2$/c$^4$ for candidates with downstream $K^0_S$ decays. While this is small compared to the typical width of a bin, net migration can occur in regions where the presence of resonances cause the density to change rapidly. To first order this effect is accounted for by use of the control channel. However, differences in the distributions of the Dalitz plots due to efficiency differences or the nonzero value of $r_B$ in the signal decay may cause residual effects. The uncertainty from this is determined via pseudoexperiments, in which different input $F_i$ values are used to reflect the residual migration. The size of any possible bias is found to be $0.1 \times 10^{-2}$ for all CP parameters.

There is a systematic uncertainty related to CP violation in the neutral kaon system due to the fact that the $K^0_S$ state is not an exact CP eigenstate and, separately, due to different nuclear interaction cross-sections of the $K^0$ and $\bar{K}^0$ mesons. The measurement is insensitive to global asymmetries, but is affected by the different Dalitz distributions of $D \rightarrow K^0_S h^- h^+$ and $D \rightarrow K^0_S h^+ h^-$ decays, as well as any correlations between Dalitz coordinates and the net material interaction. The potential bias on $x_\pm$ and $y_\pm$ is assessed using a series of pseudoexperiments, where data are generated taking the effects into account and fitted using the default fit. The $D \rightarrow K^0_S h^- h^+$ Dalitz distribution is estimated by transforming an amplitude model of $D \rightarrow K^0_S h^- h^+$ [22], following arguments and assumptions laid out in ref. [18]. The effect of material interaction is treated using the formalism described in ref. [56]. The size of the potential bias is found to be $\leq 0.2 \times 10^{-2}$ for all CP parameters, corresponding to a bias on $\gamma$ of approximately $0.8^\circ$, which is within expected limits [57].
The nonuniform efficiency profile over the Dalitz plot means that the values of \((c_i, s_i)\) appropriate for this analysis can differ from those measured by the CLEO collaboration, which correspond to the constant-efficiency case. Amplitude models are used to calculate the values of \(c_i\) and \(s_i\) both with and without the efficiency profiles determined from simulation. The models are taken from ref. [23] for \(D \to K_s^0 \pi^+ \pi^-\) decays and from ref. [24] for \(D \to K_s^0 K^+ K^-\) decays. The difference is taken as an estimate of the size of this effect. Pseudoexperiments are generated in which the values have been shifted by this difference, and then fitted with the default \((c_i, s_i)\) values. The resulting bias on \(x_\pm\) and \(y_\pm\) is found to be negligible.

The effect that a detection asymmetry between hadrons of opposite charge can have on the symmetry of the efficiency across the Dalitz plot is found to be negligible. Changes in the mass model used to describe the semileptonic control sample are also found to have a negligible effect on the \(F_i\) values.

Finally, several checks are conducted to assess the stability of the results. These include repeating the fits separately for both \(K_s^0\) categories, for each data-taking year, and by splitting the candidates depending on whether the hardware trigger decision was due to particles in the signal-candidate decay chain or other particles produced in the pp collision. No anomalies are found and no additional systematic uncertainties are assigned.

In total the systematic uncertainties are less than half of the corresponding statistical uncertainties. The correlation matrix obtained for the combined effect of the sources of experimental and strong-phase related systematic uncertainties is given in table 4.

### Table 4. Correlation matrix of the experimental and strong-phase related systematic uncertainties.

<table>
<thead>
<tr>
<th></th>
<th>(x_\pm)</th>
<th>(y_\pm)</th>
<th>(x_+)</th>
<th>(y_+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_\pm)</td>
<td>1</td>
<td>-0.25</td>
<td>0.43</td>
<td>-0.09</td>
</tr>
<tr>
<td>(y_\pm)</td>
<td>1</td>
<td>1</td>
<td>0.14</td>
<td>1</td>
</tr>
</tbody>
</table>

8 Results and interpretation

The \(CP\) observables are measured to be

\[
\begin{align*}
x_\pm &= (9.0 \pm 1.7 \pm 0.7 \pm 0.4) \times 10^{-2}, \\
y_\pm &= (2.1 \pm 2.2 \pm 0.5 \pm 1.1) \times 10^{-2}, \\
x_+ &= (-7.7 \pm 1.9 \pm 0.7 \pm 0.4) \times 10^{-2}, \\
y_+ &= (-1.0 \pm 1.9 \pm 0.4 \pm 0.9) \times 10^{-2},
\end{align*}
\]

where the first uncertainty is statistical, the second is the total experimental systematic uncertainty and the third is that arising from the precision of the CLEO measurements.
The signature for \( CP \) violation is that \((x_+, y_+) \neq (x_-, y_-)\). The distance between \((x_+, y_+)\) and \((x_-, y_-)\) is calculated, taking all uncertainties and correlations into account, and found to be

\[
|(x_+, y_+)-(x_-, y_-)| = (17.0 \pm 2.7) \times 10^{-2},
\]

which is different from zero by 6.4 standard deviations. This constitutes the first observation of \( CP \) violation in \( B^\pm \to D K^\pm \) decays for the \( D \to K^0_S h^+ h^- \) final states.

These results are compared to the expected central values of \( x_\pm \) and \( y_\pm \) that can be computed from \( r_B, \delta_B \), and \( \gamma \) as determined in the LHCb combination in ref. [54], and the results are shown in figure 10 (the later LHCb combination in ref. [58] includes the results of this measurement and is therefore unsuitable for comparison). The two sets of \((x_+, y_+)\) are in agreement within 1.6 standard deviations when the uncertainties and correlations of both the LHCb combination and this measurement are taken into account. There is a 2.7 standard deviation tension between the measured values of \((x_-, y_-)\) and the values calculated from the LHCb combination. This tension will be investigated further when this measurement and the LHCb combination are updated using data taken in 2017 and 2018.

The results for \( x_\pm \) and \( y_\pm \) are interpreted in terms of the underlying physics parameters \( \gamma, r_B \) and \( \delta_B \). The interpretation is done via a maximum likelihood fit using a frequentist treatment as described in ref. [59]. The solution for the physics parameters has a two-fold ambiguity as the equations are invariant under the simultaneous substitutions \( \gamma \to \gamma + 180^\circ \) and \( \delta_B \to \delta_B + 180^\circ \). The solution that satisfies \( 0 < \gamma < 180^\circ \) is chosen. The central values and 68% (95%) confidence intervals, calculated with the PLUGIN [60] method, are

\[
\begin{align*}
\gamma &= 87^\circ \pm 11^\circ \left( ^{+22^\circ}_{-23^\circ} \right), \\
r_B &= 0.086 \pm 0.013 \left( ^{+0.025}_{-0.014} \left( ^{+0.025}_{-0.027} \right) \right), \\
\delta_B &= 101^\circ \pm 11^\circ \left( ^{+22^\circ}_{-23^\circ} \right).
\end{align*}
\]
The values for $\gamma$ and $r_B$ are consistent with those presented in ref. [54]. This is the most precise measurement of $\gamma$ from a single analysis. The value of $\delta_B$ shows some disagreement with ref. [54], where the angle is determined to be $(139.9^{+4.8}_{-3.2})^\circ$.

The values of $x_{\pm}, y_{\pm}$ measured in this analysis can be combined with those from the corresponding analysis of Run 1 data [9]. This procedure is done via a maximum likelihood fit, as implemented in the `gammacombo` package [59]. The previous measurements are identified by the index I, and the results within this paper are identified by the index II. When combining the two results, the fit determines the $(\hat{x}_{\pm}, \hat{y}_{\pm})$ parameters that maximize the multivariate Gaussian likelihood function

$$L(\hat{z}|\hat{\Sigma}) = ((2\pi)^8|\Sigma|)^{-1/2} \exp \left[-\frac{1}{2}(z - \hat{z})^T \Sigma^{-1}(z - \hat{z}) \right],$$

(8.1)

where $z = (x^I_{\pm}, y^I_{\pm}, x^{II}_{\pm}, y^{II}_{\pm})^T$ and $\hat{z} = (\hat{x}_{\pm}, \hat{y}_{\pm}, \hat{x}_{\pm}, \hat{y}_{\pm})^T$ are $8 \times 1$ vectors and $\Sigma$ is the $8 \times 8$ covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma^I & \Sigma^{I-II} \\ \Sigma^{II-I} & \Sigma^{II} \end{pmatrix}.$$  

(8.2)

The covariance matrix is expressed in terms of the covariance matrices obtained for the individual measurements, $\Sigma^I$ and $\Sigma^{II}$, and the cross-covariance matrix $\Sigma^{I-II}$ describing correlations between the measurements. The covariance matrix for this measurement, $\Sigma^{II}$, is calculated using the total statistical and systematic uncertainties, and the correlation matrices in tables 1 and 4. The covariance matrix for the Run 1 measurement, $\Sigma^I$, is taken from ref. [9], where it was calculated taking strong-phase-related correlations into account, but treating the experimental systematic uncertainties as uncorrelated. The impact of using the correlation matrix in table 4 for these instead is found to be negligible.

The dominant uncertainty in both measurements is the statistical uncertainty. As the measurements use independent data sets, the statistical uncertainties are uncorrelated. The cross-correlations of the systematic errors between measurements due to the strong phase inputs are obtained from the results of a series of fits to the two data sets in which the strong phases are varied identically. This mirrors the procedure used to evaluate the uncertainties within a single data set. The obtained cross-correlations between the fit results are given in table 5. The elements on the diagonal do not have unit value because the obtained correlations depend on the specific data sets for the two measurements.

The combination is performed assuming full correlation between the non-strong-phase related experimental systematic uncertainties in Run 1 and this measurement. The correlation matrix for the experimental uncertainties of this analysis is used as the cross-run correlation of the experimental systematic uncertainties. The complete correlation matrix for the experimental and strong-phase-related systematic uncertainties is given in table 6. The impact on the combination due to different assumptions on the cross-correlations of the systematic uncertainties is found to be negligible. This is unsurprising as both measurements remain limited in precision by their statistical uncertainties. The central values,
Table 5. Correlation matrix between Run 1 results (I) and the results presented in this paper (II), when fitting data while varying the inputs from the CLEO collaboration in a correlated way.

<table>
<thead>
<tr>
<th></th>
<th>$x_-^I$</th>
<th>$y_-^I$</th>
<th>$x_-^{II}$</th>
<th>$y_-^{II}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_-^I$</td>
<td>0.02</td>
<td>0.35</td>
<td>-0.32</td>
<td>-0.21</td>
</tr>
<tr>
<td>$y_-^I$</td>
<td>0.45</td>
<td>-0.23</td>
<td>0.03</td>
<td>-0.28</td>
</tr>
<tr>
<td>$x_-^{II}$</td>
<td>-0.19</td>
<td>0.01</td>
<td>0.55</td>
<td>-0.22</td>
</tr>
<tr>
<td>$y_-^{II}$</td>
<td>-0.30</td>
<td>-0.28</td>
<td>0.13</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 6. Total correlation matrix for the systematic uncertainties of the Run 1 results (I) and the results presented in this paper (II), including experimental and strong phase related systematic uncertainties.

<table>
<thead>
<tr>
<th></th>
<th>$x_-^I$</th>
<th>$y_-^I$</th>
<th>$x_-^{II}$</th>
<th>$y_-^{II}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_-^I$</td>
<td>0.76</td>
<td>0.04</td>
<td>0.55</td>
<td>0.02</td>
</tr>
<tr>
<td>$y_-^I$</td>
<td>0.14</td>
<td>-0.06</td>
<td>-0.13</td>
<td>-0.25</td>
</tr>
<tr>
<td>$x_-^{II}$</td>
<td>0.58</td>
<td>-0.19</td>
<td>0.91</td>
<td>0.05</td>
</tr>
<tr>
<td>$y_-^{II}$</td>
<td>-0.05</td>
<td>-0.24</td>
<td>0.17</td>
<td>0.55</td>
</tr>
</tbody>
</table>

along with the combined statistical and systematic uncertainties for this combination are

\[
\begin{align*}
\langle x_- \rangle &= (7.0 \pm 1.7) \times 10^{-2}, \\
\langle y_- \rangle &= (4.1 \pm 2.0) \times 10^{-2}, \\
\langle x_+ \rangle &= (-7.8 \pm 1.7) \times 10^{-2}, \\
\langle y_+ \rangle &= (-1.4 \pm 1.7) \times 10^{-2}.
\end{align*}
\]

The interpretation in terms of the underlying physics parameters is performed on the combined values of $x_\pm$ and $y_\pm$ and the central values and their 68% (95%) confidence intervals are

\[
\begin{align*}
\gamma &= 80^\circ +10^\circ -9^\circ \left( +19^\circ -18^\circ \right), \\
r_B &= 0.080 +0.011 -0.011 \left( +0.022 -0.023 \right), \\
\delta_B &= 110^\circ +10^\circ -10^\circ \left( +19^\circ -20^\circ \right).
\end{align*}
\]

The results of the interpretation for both the combined and individual data sets are shown in figure 11, where the projections of the three-dimensional surfaces bounding the one and two standard deviation volumes on the $(\gamma, r_B)$ and $(\gamma, \delta_B)$ planes are shown. The uncertainty on $\gamma$ is inversely proportional to $r_B$. Therefore the lower central value of $r_B$ in the combined results lead to a larger than naively expected uncertainty on $\gamma$ when both data sets are used. The contribution of each source of uncertainty are estimated by
performing the combination while taking only subsets of the uncertainties into account. It is found that the statistical uncertainty on $\gamma$ is $8.5^\circ$, the uncertainty due to strong-phase inputs is $4^\circ$, and the uncertainty due to experimental systematic effects is $2^\circ$.

9 Conclusions

Approximately 4100 (560) $B^\pm \rightarrow D K^\pm$ decays with the $D$ meson decaying to $K_S^0 \pi^+ \pi^-$ ($K_S^0 K^+ K^-$) are selected from data corresponding to an integrated luminosity of 2.0 fb$^{-1}$ collected with the LHCb detector in 2015 and 2016. These samples are analysed to determine the $CP$-violating parameters $x_\pm = r_B \cos(\delta_B \pm \gamma)$ and $y_\pm = r_B \sin(\delta_B \pm \gamma)$, where $r_B$ is the ratio of the absolute values of the $B^+ \rightarrow D^{*0} K^-$ and $B^+ \rightarrow D^{*0} K^-$ amplitudes, $\delta_B$ is their strong-phase differences, and $\gamma$ is an angle of the Unitarity Triangle. The analysis is performed in bins of the $D$-decay Dalitz plot and existing measurements performed by the CLEO collaboration [18] are used to provide input on the $D$-decay strong-phase parameters ($c_i, s_i$). Such an approach allows the analysis to be free from model-dependent assumptions on the strong-phase variation across the Dalitz plot. This paper also gives the combination with the results obtained with an earlier data set, thereby allowing further improvements in the precision on $\gamma$. Considering only the data collected in 2015 and 2016 and choosing the solution that satisfies $0 < \gamma < 180^\circ$ yields $r_B = 0.086^{+0.013}_{-0.014}$, $\delta_B = (101 \pm 11)^\circ$, and $\gamma = (87^{+11}_{-12})^\circ$. The values of $r_B$ and $\gamma$ are consistent with world averages, while there is some tension in the determined value of $\delta_B$. This could be resolved by future analyses of the $B \rightarrow D K$ mode in a variety of $D$ decays, including those analysed here, utilising the data set that is being collected with LHCb in 2017 and 2018. The measurement reported in this paper represents the most precise determination of $\gamma$ from a single analysis.

Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies:
CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.A.). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (U.S.A.). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), ANR, Labex P2IO and OCEVU, and Région Auvergne-Rhône-Alpes (France), Key Research Program of Frontier Sciences of CAS, CAS PIFI, and the Thousand Talents Program (China), RFBR, RSF and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, the Royal Society, the English-Speaking Union and the Leverhulme Trust (United Kingdom).

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


[10] LHCb collaboration, A study of CP violation in B± → Dh± (h = K, π) with the modes D → K±π±π0, D → π±π−π0 and D → K±K−π0, Phys. Rev. D 91 (2015) 112014 [arXiv:1504.05402] [inSPIRE].
[11] LHCb collaboration, Measurement of CP observables in $B^{\pm} \to D K^{\pm}$ and $B^{\pm} \to D \pi^{\pm}$ with two- and four-body $D$ decays, Phys. Lett. B 760 (2016) 117 [arXiv:1603.08993] [SPIRE].


[13] LHCb collaboration, Measurement of CP asymmetry in $B_s^0 \to D^{\mp} K^\pm$ decays, JHEP 11 (2014) 060 [arXiv:1407.6127] [SPIRE].

[14] LHCb collaboration, Study of $B^- \to D K^- \pi^+ \pi^-$ and $B^- \to D \pi^- \pi^+ \pi^-$ decays and determination of the CKM angle $\gamma$, Phys. Rev. D 92 (2015) 112005 [arXiv:1505.07044] [SPIRE].

[15] LHCb collaboration, Constraints on the unitarity triangle angle $\gamma$ from Dalitz plot analysis of $B^0 \to D K^+ \pi^-$ decays, Phys. Rev. D 93 (2016) 112018 [arXiv:1602.03455] [SPIRE].


[18] CLEO collaboration, J. Libby et al., Model-independent determination of the strong-phase difference between $D^0$ and $\bar{D}^0 \to K_S^0 h^+ h^- (h = \pi, K)$ and its impact on the measurement of the CKM angle $\gamma/\phi_3$, Phys. Rev. D 82 (2010) 112006 [arXiv:1010.2817] [SPIRE].

[19] Belle collaboration, H. Aihara et al., First Measurement of $\phi_3$ with a model-independent Dalitz plot analysis of $B \to D K$, $D \to K_S^0 \pi^+ \pi^-$ decay, Phys. Rev. D 85 (2012) 112014 [arXiv:1204.6661] [SPIRE].

[20] Belle collaboration, K. Negishi et al., First model-independent Dalitz analysis of $B^0 \to D K^{*0}$, $D \to K_S^0 \pi^+ \pi^-$ decay, PTEP 2016 (2016) 043C01 [arXiv:1509.01098] [SPIRE].

[21] LHCb collaboration, Model-independent measurement of the CKM angle $\gamma$ using $B^0 \to D K^{*0}$ decays with $D \to K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$, JHEP 06 (2016) 131 [arXiv:1604.01525] [SPIRE].

[22] BABAR collaboration, B. Aubert et al., Measurement of $\gamma$ in $B^{\mp} \to D^{(*)} K^{(*)\mp}$ decays with a Dalitz analysis of $D \to K_S^0 \pi^- \pi^+$, Phys. Rev. Lett. 95 (2005) 121802 [hep-ex/0504039] [SPIRE].

[23] BABAR collaboration, B. Aubert et al., Improved measurement of the CKM angle $\gamma$ in $B^{\mp} \to D^{(*)} K^{(*)\mp}$ decays with a Dalitz plot analysis of $D$ decays to $K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$, Phys. Rev. D 78 (2008) 034023 [arXiv:0804.2089] [SPIRE].


[27] Belle collaboration, A. Poluektov et al., Evidence for direct CP-violation in the decay $B \to D^{(*)}K^{*}$, $D \to K^{0}_{S}\pi^{+}\pi^{-}$ and measurement of the CKM phase $\phi_3$, Phys. Rev. D 81 (2010) 112002 [arXiv:1003.3360] [inSPIRE].

[28] LHCb collaboration, Measurement of CP violation and constraints on the CKM angle $\gamma$ in $B^{\pm} \to D K^{\pm} \to D^{(*)} K^{*} \to D^{(*)} K^{0}_{S}\pi^{+}\pi^{-}$ decays, Nucl. Phys. B 888 (2014) 169 [arXiv:1407.6211] [inSPIRE].


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


[50] LHCb collaboration, *Dalitz plot analysis of $B_s^0\to D^0 K^-\pi^+$ decays*, *Phys. Rev. D* 90 (2014) 072003 [arXiv:1407.7712] [inspire].


[56] LHCb collaboration, *Measurement of $CP$ asymmetry in $D^0\to K^-K^+$ and $D^0\to \pi^-\pi^+$ decays*, *JHEP* 07 (2014) 041 [arXiv:1405.2797] [inspire].

[57] Y. Grossman and M. Savastio, *Effects of $K^0-\Bar{K}^0$ mixing on determining $\gamma$ from $B^{\pm}\to DK^{\pm}$*, *JHEP* 03 (2014) 008 [arXiv:1311.3575] [inspire].


The LHCb collaboration

9. I. Physikalisches Institut, RWTH Aachen University, Aachen, Germany
10. Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
11. Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
12. Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
13. School of Physics, University College Dublin, Dublin, Ireland
14. INFN Sezione di Bari, Bari, Italy
15. INFN Sezione di Bologna, Bologna, Italy
16. INFN Sezione di Ferrara, Ferrara, Italy
17. INFN Sezione di Firenze, Firenze, Italy
18. INFN Laboratori Nazionali di Frascati, Frascati, Italy
19. INFN Sezione di Genova, Genova, Italy
20. INFN Sezione di Milano-Bicocca, Milano, Italy
21. INFN Sezione di Milano, Milano, Italy
22. INFN Sezione di Cagliari, Monserrato, Italy
23. INFN Sezione di Padova, Padova, Italy
24. INFN Sezione di Pisa, Pisa, Italy
25. INFN Sezione di Roma Tor Vergata, Roma, Italy
26. INFN Sezione di Roma La Sapienza, Roma, Italy
27. Nikhef National Institute for Subatomic Physics, Amsterdam, Netherlands
28. Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
29. AGH — University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
30. National Center for Nuclear Research (NCBJ), Warsaw, Poland
31. Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
32. Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
33. Institute of Theoretical and Experimental Physics (ITPP), Moscow, Russia
34. Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
35. Institute for Nuclear Research of the Russian Academy of Sciences (INR RAS), Moscow, Russia
36. Yandex School of Data Analysis, Moscow, Russia
37. Budker Institute of Nuclear Physics (SB RAS), Novosibirsk, Russia
38. Institute for High Energy Physics (IHEP), Protvino, Russia
39. ICCUB, Universitat de Barcelona, Barcelona, Spain
40. Instituto Galego de Física de Altas Enerxías (IGFAE), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
41. European Organization for Nuclear Research (CERN), Geneva, Switzerland
42. Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
43. Physik-Institut, Universität Zürich, Zürich, Switzerland
44. NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
45. Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
46. University of Birmingham, Birmingham, United Kingdom
47. H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
48. Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
49. Department of Physics, University of Warwick, Coventry, United Kingdom
50. STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
51. School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
52. School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
53. Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
54. Imperial College London, London, United Kingdom
55. School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom