First observation of the rare \( B^+ \to D^+ K^+ \pi^- \) decay

R. Aaij et al.*
(LHCb Collaboration)
(Received 9 December 2015; published 3 March 2016)

The \( B^+ \to D^+ K^+ \pi^- \) decay is observed in a data sample corresponding to 3.0 fb\(^{-1}\) of \( pp \) collision data recorded by the LHCb experiment during 2011 and 2012. The signal significance is 8σ and the branching fraction is measured to be \( \mathcal{B}(B^+ \to D^+ K^+ \pi^-) = (5.31 \pm 0.90 \pm 0.48 \pm 0.35) \times 10^{-6} \), where the uncertainties are statistical, systematic and due to the normalization mode \( B^+ \to D^- K^+ \pi^- \), respectively. The Dalitz plot appears to be dominated by broad structures. Angular distributions are exploited to search for quasi-two-body contributions from \( B^+ \to D_s^+(2460)^0 K^+ \) and \( B^+ \to D^+ K^+(892)^0 \) decays. No significant signals are observed and upper limits are set on their branching fractions.

DOI: 10.1103/PhysRevD.93.051101

A key goal of flavor physics is to determine precisely the angle \( \gamma \) of the unitarity triangle constructed from pairs of elements of the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix \([1,2]\). The value of \( \gamma \equiv \arg[-V_{ud}V_{ub}^*/(V_{cd}V_{cb}^*)] \) is currently known to a precision of only about 10° \([3–5]\), which limits the sensitivity of tests of the Standard Model through global fits to the CKM matrix parameters.

A powerful method to determine \( \gamma \), known as the GLW method \([6,7]\), is to use \( B^+ \to D K^- \) decays with the neutral \( D \) meson decaying to \( CP \) eigenstates. The \( b \to u \) and \( b \to \bar{u} \) amplitudes both contribute to the decay, and the sensitivity to \( \gamma \) comes from their interference. A challenge with this method is that the ratio of magnitudes of the suppressed and favored amplitudes, \( r_B \), is not known independently and must be determined simultaneously with \( \gamma \). This is usually addressed by using in addition other decays of the \( D \) meson that provide complementary information on \( r_B \) and \( \gamma \) \([8,9]\).

In the case of \( B^+ \to D^{*+} K^+ \) decays, where \( D^{*+} \) represents an excited \( D \) or \( \bar{D} \) meson such as the \( D_s^+(2460) \) state which can decay to both \( D^\pm \pi^\mp \) and \( D \pi^0 \), it is possible to obtain a clean determination of \( r_B \) \([10]\). The relative branching fractions of the \( b \to u \) mediated \( B^+ \to D^{*0}K^+ \to D^+\pi^-K^+ \) and the \( b \to \bar{u} \) mediated \( B^+ \to D^{*+}K^+ \to D^-\pi^+K^+ \) processes give the value of \( r_B^2 \), while the \( B^+ \to D^{*+}K^+ \to D\pi^0 K^+ \) final state, where the \( D \) meson is reconstructed using \( CP \) eigenstate decay modes, provides sensitivity to \( \gamma \). Decay diagrams for \( B^+ \to D_s^+(2460)^0 K^+ \) and \( B^+ \to D_s^+(2460)^0 K^+ \) decays are shown in Figs. 1(a) and 1(b).

Knowledge of the resonant structure of \( B^+ \to D^{*+} K^+ \pi^- \) and \( B^+ \to D^- K^+ \pi^+ \) decays is therefore needed. The latter channel has recently been studied with a Dalitz plot analysis \([11]\). Such a study would be difficult with the low yields expected for \( B^+ \to D^{*+} K^+ \pi^- \) decays in the available data samples, but an alternative approach exploiting the angular decay information to separate different spin states is viable in the region of the narrow \( D_s^+(2460)^0 \) resonance. The same method can also be used to search for \( B^+ \to D^+ K^+(892)^0 \) decays, which contribute to the \( D^+ K^+ \pi^- \) final state and are of interest since they are mediated by annihilation amplitudes, as shown in Fig. 1(c). A previous LHCb analysis of this mode set an upper limit \( \mathcal{B}(B^+ \to D^+ K^+(892)^0) < 1.8 \times 10^{-6} \) at the 90% credibility level \([12]\).

In this paper, the \( B^+ \to D^+ K^+ \pi^- \) channel is studied for the first time, and searches for \( B^+ \to D_s^+(2460)^0 K^+ \) and \( B^+ \to D^+ K^+(892)^0 \) decays are performed. The \( D^+ \) meson is reconstructed in the \( K^-\pi^+\pi^- \) final state. (The inclusion of charge conjugate processes is implied.) The \( B^+ \to D^- K^+ \pi^- \) decay \([11]\) is used for normalization. The analysis is based on 3.0 fb\(^{-1}\) of \( pp \) collision data collected with the LHCb detector during 2011 and 2012. The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range \( 2 < \eta < 5 \), described in detail in Refs. \([13,14]\). Simulated events are produced using the software described in Refs. \([15–20]\). To reduce the risk of biasing results, all analysis procedures were established before the data in the signal region were examined.

Candidates consistent with the decay chains \( B^+ \to D^{*+} K^+ \pi^- \) and \( B^+ \to D^- K^+ \pi^+ \) with \( D^\pm \to K^{\mp}\pi^\pm\pi^\pm \) are selected. The criteria for \( B^+ \to D^- K^+ \pi^+ \) and \( B^+ \to D^+ K^+ \pi^- \) candidates are identical, except for charge requirements, and are very similar to those described in Ref. \([11]\). A loose preselection is applied before two neural network classifiers (NNs) \([21]\) are used to separate signal decays from background events. The first NN separates true \( D^\pm \to K^{\mp}\pi^\pm\pi^\pm \) decays from random combinations, and the second (NN2) identifies signal \( B^+ \) decays. Both NNs

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
are trained with a sample of candidates from the topologically similar $B^+ \to D^+ \pi^- \pi^-$ decay. Additional selection requirements are imposed to reject contributions from specific decay modes. Only candidates with $DK$ mass, $m(DK)$, less than 5140 MeV/c^2 are kept, in order to remove backgrounds from $B^0 \to D^+ K^-$ decays combined with a random pion candidate. Similarly, potential $B^0 \to D^- \pi^+$ background is eliminated by requiring $m(D\pi) < 4790$ MeV/c^2.

Contributions from $B^+ \to D^+ \bar{D}^0$ decays, with $\bar{D}^0 \to K^+ \pi^-$, are removed by rejecting candidates within $\sim 3\sigma_{K_s}$ of the $\bar{D}^0$ mass, where $\sigma_{K_s}$ is the $K^+ \pi^-$ mass resolution, corresponding to $1830 < m(K\pi) < 1890$ MeV/c^2. Although each of these backgrounds affects only one of the final states, the vetoes are applied to both to avoid biasing the relative efficiency.

Signal candidates with invariant mass in the range 5100–5800 MeV/c^2 are retained for further analysis. Following all selection requirements, fewer than 1% of events contain more than one candidate; all are retained.

Extended maximum likelihood fits to the distributions of candidates in NN2 output and in $B$ candidate mass are used to determine the yields of $B^+ \to D^+ K^+ \pi^-$ and $B^+ \to D^- K^- \pi^+$ decays. Similar fitting techniques have been used successfully in several previous LHCb analyses [22–25]. A loose requirement is placed on the output of NN2 and the remaining data in each sample are divided into six bins of the NN2 output variable, each containing a similar number of signal decays. This binning scheme enhances the sensitivity while giving stable fit performance.

The $B$ candidate mass shapes in the fit to $B^+ \to D^- K^+ \pi^+$ candidates are modeled in the same way as described in Ref. [11]. The signal is described by the sum of two Crystal Ball (CB) [26] functions, with a common mean. The tails on both sides of the peak are described by parameters that are fixed to values found in fits to simulated samples. Components are included for combinatorial background, modeled with an exponential shape, and for partially reconstructed backgrounds from $B^+ \to D^+ \bar{D}^0$ decays and misidentified $B^+ \to D^{(*)-} \pi^- \pi^+$ and $B^+ \to D_s^+ K^+ \pi^-$ decays, for which nonparametric descriptions are determined from simulation. Data-driven estimates of the misidentification probabilities, the phase-space distributions of the $B^+ \to D^{(*)-} \pi^- \pi^+$ decays [27,28] and the relative branching fractions of the $B^+ \to D_{s}^+ \pi^-$ and $B^+ \to D^+ \pi^- \pi^+$ modes [27,29] are used to obtain these shapes. For signal and partially reconstructed and combinatorial background components, the relative yields in each NN2 bin are free parameters of the fit, while those of misidentified $B^+ \to D^{(*)-} \pi^- \pi^+$ and $B^+ \to D_s^+ K^+ \pi^-$ decays are taken to be the same as for signal decays, since their NN2 responses are expected to be very similar.

A total of 25 parameters are determined from the fit to the $D^- K^+ \pi^+$ sample. These include yields of $B^+ \to D^- K^+ \pi^+$ decays, the combinatorial background, the partially reconstructed background, and the $B^+ \to D^{(*)-} \pi^- \pi^+$ and $B^+ \to D_s^+ K^+ \pi^-$ misidentified backgrounds. For the signal category, and for combinatorial and partially reconstructed backgrounds, the fractional yields $f_i$ of each component in NN2 bins 1–5 are free parameters, with the fraction in bin 6 determined as $f_6 = 1 - \sum_{i=1}^{5} f_i$. In addition, the exponential slope parameter of the combinatorial background and parameters of the signal invariant mass shape (the peak position, the width of the core CB function, the relative normalization and ratio of the CB widths) are allowed to vary. Figure 2 shows the combined $B^+$ candidate mass distribution of all NN2 bins weighted by $S/(S + B)$, where $S$ and $B$ are the fitted signal and background yields within $\pm 2.5\sigma_{CB}$ of the signal peak position and $\sigma_{CB}$ is the width of the core CB function. The fit results are summarized in Table I.

The model for the fit to $B^+ \to D^+ K^+ \pi^-$ candidates is similar to that for the $B^+ \to D^- K^+ \pi^+$ case. The functional forms for the mass shapes for signal and combinatorial background are identical. The signal peak position, the width of the core CB function, and the fractional yields in each NN2 bin are fixed to the values obtained from the $B^+ \to D^+ K^+ \pi^+$ fit. A component is included for partially reconstructed background, which is likely to be dominated by $B^0 \to D^+ K^- \pi^- \pi^-$ decays; although this channel is unobserved, it is expected to be a sizable source of background based on studies of similar decay modes [30–32]. As the resonant structure of this mode is unknown, its mass shape is modeled using a combination of simulated samples generated with various $D^+ \pi^-$, $D^+ K^+ \pi^-$ and $K^+ \pi^-$ resonances and nonresonant amplitudes. The unknown structure of this background could cause some disagreement between data and the fit result at low $m(D^- K^+ \pi^-)$. The fractional yields in each NN2 bin are fixed to be the same as those for partially reconstructed backgrounds in the $B^+ \to D^+ K^+ \pi^+$ fit. Potential partially reconstructed background from $B^0$ and $B^+$ decays with a missing pion hardly enter the fit region; any residual contributions are absorbed in the $B^0 \to D^+ K^+ \pi^- \pi^-$ mass shape.

![Decay diagrams for (a) $B^+ \to D_s^{(2)}(2460)^0 K^+$, (b) $B^+ \to D_s^{(2)}(2460)^0 K^+$ and (c) $B^+ \to D^+ K^+(892)^0$.](image_url)

FIG. 1. Decay diagrams for (a) $B^+ \to D_s^{(2)}(2460)^0 K^+$, (b) $B^+ \to D_s^{(2)}(2460)^0 K^+$ and (c) $B^+ \to D^+ K^+(892)^0$ channels.
There remain 11 parameters that are varied in the fit to the $D^+K^+\pi^-$ sample: the yields for $B^+ \to D^+K^+\pi^-$ decays, combinatorial and partially reconstructed backgrounds; the fractional yields of the combinatorial background in each NN2 bin; the exponential slope parameter of the combinatorial background and the relative normalization and ratio of widths of the two CB functions. The results of this fit are summarized in Table I and shown in Fig. 2. The statistical significance of the $D^+K^+\pi^-$ peak, obtained from the square root of twice the change in negative log likelihood from the value obtained in a fit with zero signal yield, is $11\sigma$.

Systematic uncertainties on the ratio $B(B^+ \to D^+K^+\pi^-)/B(B^+ \to D^-K^+\pi^-)$ arise due to approximations made in the signal and background shapes used in the fit and uncertainties in the relative efficiencies. The largest uncertainties are associated with the particle identification and hardware trigger efficiencies (5.9%), the modeling of the combinatorial background in $B$ candidate mass and NN2 bins (4.4%) and the NN2 distributions of signal and partially reconstructed background (4.2%). Other sources, including the modeling of the $B^+_s \to D^+K^+\pi^+\pi^-$ background and potential biases that are either intrinsic to the fit procedure or related to the treatment of multiple candidates, contribute systematic uncertainties of 2.0% or less.

With all sources combined in quadrature, the total systematic uncertainty on the ratio of branching fractions is found to be 9.0%. The likelihood function is convolved with a Gaussian of width corresponding to the size of the systematic uncertainties that affect the signal yield, and the total significance of the signal is found to be $8\sigma$.

The relative branching fraction of $B^+ \to D^+K^+\pi^-$ and $B^+ \to D^-K^+\pi^-$ decays is determined from

$$\frac{B(B^+ \to D^+K^+\pi^-)}{B(B^+ \to D^-K^+\pi^-)} = \frac{N_{\text{corr}}(B^+ \to D^+K^+\pi^-)}{N_{\text{corr}}(B^+ \to D^-K^+\pi^-)}$$  \hspace{1cm} (1)$$

where the efficiency-corrected yield is $N_{\text{corr}} = \sum_i W_i/\epsilon_i$. Here the index $i$ runs over all candidates in the fit range, $W_i$ is the signal weight for candidate $i$, determined using the sPlot procedure [33], from the fits shown in Fig. 2, and $\epsilon_i$ is the efficiency for candidate $i$ as a function of its Dalitz plot position.

The average efficiencies are defined as $\bar{\epsilon} = N/N_{\text{corr}} = \sum_i W_i/N_{\text{corr}}$ and are found to be $\bar{\epsilon}(B^+ \to D^+K^+\pi^-) = (0.057 \pm 0.014\%)$ and $\bar{\epsilon}(B^+ \to D^-K^+\pi^-) = (0.079 \pm 0.003\%)$. These values include contributions from the LHCb detector acceptance, selection and trigger. The trigger efficiency and most selection efficiencies are calculated from simulated samples with data-driven corrections applied, while the particle identification efficiency is measured from a data control sample [34]. The difference between the efficiencies is mainly caused by the different Dalitz plot distributions of the data in each channel.

From Eq. (1), the ratio of branching fractions is determined to be

$$\frac{B(B^+ \to D^+K^+\pi^-)}{B(B^+ \to D^-K^+\pi^-)} = 0.073 \pm 0.012(\text{stat}) \pm 0.007(\text{syst}).$$

Taking $B(B^+ \to D^-K^+\pi^-) = (7.31 \pm 0.19 \pm 0.22 \pm 0.39) \times 10^{-5}$ [11] gives

$$B(B^+ \to D^+K^+\pi^-) = (5.31 \pm 0.90 \pm 0.48 \pm 0.35) \times 10^{-6},$$

where the third uncertainty is from $B(B^+ \to D^-K^+\pi^-)$, which arises mainly from the precision with which $B(B^+ \to D^-\pi^+\pi^+)$ [29] is known.
the limit that only $D^+\pi^-$ resonances contribute, weighting by $P_4$ will isolate the $D_s^0(2460)^0$ component, as shown in Ref. [11], where a more detailed description of contributions to each moment can be found. Similarly, at low $m(K^+\pi^-)$, weighting by $P_3$ can be used to determine the contribution from the $K^*(892)^0$ resonance. Higher moments may be present, due to tails of higher spin resonances or reflections from resonances in the other two-particle combination; these will also cause an excess of events in regions away from the resonance peak and therefore can be accounted for by sideband subtraction.

Candidates are selected within regions corresponding to approximately $\pm 2\Gamma$, where $\Gamma$ is the natural width [29], around the peaks of the $D_s^0(2460)^0$ resonance in $m(D^+\pi^+)$ and of the $K^*(892)^0$ resonance in $m(K^+\pi^-)$. The data are efficiency corrected and weighted according to the corresponding Legendre polynomial functions. Yields, denoted $N$, are then obtained from binned minimum $\chi^2$ fits to the $B^+$ candidate mass distribution. A variable-width binning scheme is used with bin widths chosen to avoid empty bins. The same procedure is applied for candidates in low and high sideband regions, between about 3$\Gamma$ and 5$\Gamma$ from the peak. For the normalization of the search for $D^+K^*(892)^0$ decays, the full efficiency-corrected $D^+K^+\pi^-$ sample is used without weighting by angular moment. The results are used to measure the ratios of branching fractions

$$\frac{B(B^+ \to D_s^0(2460)^0K^+)}{B(B^+ \to D_s^0(2460)^0K^+)} = \left(\frac{r_B(D_s^0(2460)^0K^+)}{N_{\text{corr}}(B^+ \to D_s^0(2460)^0K^+)}\right)^2$$

$$= \frac{N_{\text{corr}}(B^+ \to D_s^0(2460)^0K^+)}{N_{\text{corr}}(B^+ \to D_s^0(2460)^0K^+)}.$$ (2)

$$\frac{B(B^+ \to D^+K^+(892)^0 \to D^+K^+\pi^-)}{B(B^+ \to D^+K^+\pi^+)} = \frac{N_{\text{corr}}(B^+ \to D^+K^+(892)^0 \cdot \frac{1}{2})}{N_{\text{corr}}(B^+ \to D^-K^+\pi^-) \epsilon(K^+(892)^0)}.$$ (3)

where $N_{\text{corr}}$ are the yields obtained from the fit after accounting for subtraction of higher moments as estimated from the sideband regions. In Eq. (3) the correction of $\frac{1}{2}$ arises from the normalization of the Legendre polynomial functions and the factor of $\epsilon(K^+(892)^0) = 0.857 \pm 0.006$ is due to the efficiency of the $K^+(892)^0$ signal region [801.0 < $m(K^+\pi^-) < 990.6$ MeV/$c^2$] requirement. All efficiency, $D_s^0$ branching fraction and normalization effects cancel in Eq. (2).

The fit models used are based on those described above, but with some important simplifications. The angular weighting by $P_3$ or $P_4$ significantly reduces the combinatorial background, and therefore candidates in all NN2 bins are combined; moreover a linear shape is used instead of an

![Graph](image-url)
exponential function in order to allow for the possibility that the weighted background can fluctuate to negative values. The $B^+ \rightarrow D^+ K^+\pi^-$ and $\bar{B}^0 \rightarrow D^0 K^+\pi^-$ background shapes are given by nonparametric functions obtained from simulated samples with angular moment
weighting applied. No component is included for mis-
identified $B^+ \rightarrow D^- K^+\pi^−$ decays, as it is found to be removed by the weighting procedure.

The analysis method is validated using the $B^+ \rightarrow D^- K^+\pi^−$ channel and simulated pseudo-experiments. The fit fraction for $B^+ \rightarrow D_2^{(*)}(2460)^0 K^+$ decays obtained from a full Dalitz plot analysis in Ref. [11] is reproduced within the expected range. The procedure is tested by searching for a fake $K'$ resonance in $m(K^+\pi^+)$, and the yield is found to be consistent with zero.

The results of the fits to $P_4$-weighted and efficiency-
corrected $B^+ \rightarrow D^+ K^+\pi^−$ and $B^+ \rightarrow D^- K^+\pi^−$ data samples in the $D_2^{(*)}(2460)^0$ resonance region are shown in Fig. 4. The procedure isolates the $B^+ \rightarrow D_2^{(*)}(2460)^0 K^+$ decay, as expected, but no evidence is seen for the suppressed $B^+ \rightarrow D_2^{(*)}(2460)^0 K^+$ channel. The corresponding fits for the $B^+ \rightarrow D^+ K^+(892)^0$ search are also shown in Fig. 4; there is no evidence for this decay. The yields are given in Table II.

Systematic uncertainties arise due to the fit models and background subtraction used to determine $\tilde{N}_{\text{corr}}$ in Eqs. (2) and (3). The uncertainties are evaluated from the effects on the yields of the following variations: the combinatorial background shape is changed from linear to flat; the

<table>
<thead>
<tr>
<th>Table II. Results of the binned minimum $\chi^2$ fits to efficiency-corrected $B^+$ candidate invariant mass distributions in each resonance region and with weighting according to angular distributions as described in the text.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Sideband</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>$\tilde{N}(B^+ \rightarrow D_2^{(*)}(2460)^0 K^+)$</td>
</tr>
<tr>
<td>$\tilde{N}(B^+ \rightarrow D_2^{(*)}(2460)^0 K^+)$</td>
</tr>
<tr>
<td>$\tilde{N}(B^+ \rightarrow D^- K^+\pi^−)$</td>
</tr>
<tr>
<td>$\tilde{N}(B^+ \rightarrow D^- K^+\pi^−)$</td>
</tr>
</tbody>
</table>
\( \bar{B}_s ^0 \rightarrow D^+ K^+ \pi^- \pi^- \) background component is removed; all other fit components are varied in the same way as described previously. The limited precision of the knowledge of the efficiencies as functions of Dalitz plot position also causes a small uncertainty. An uncertainty is assigned due to the effect of changing the sideband regions from the default of 3\( \Gamma \leftrightarrow 5\Gamma \) to 4\( \Gamma \leftrightarrow 6\Gamma \). The uncertainty in \( \epsilon(K^*(892)^0) \) of Eq. (3) is also accounted for. The total systematic uncertainty is obtained by combining all sources in quadrature.

The ratio of branching fractions is thus measured to be

\[
\frac{\mathcal{B}(B^+ \rightarrow D_s^+(2460)^0 K^+)}{\mathcal{B}(B^+ \rightarrow D_s^-(2460)^0 K^+)} = 0.002 \pm 0.015(\text{stat}) \pm 0.005(\text{syst}),
\]

which in turn gives

\[
r_B(D_s^0(2460)K^+) = 0.04 \pm 0.18(\text{stat}) \pm 0.06(\text{syst}).
\]

Assuming Gaussian uncertainties, upper limits at 90(95)% confidence level (CL) are obtained by integrating the likelihood in the region of positive branching fraction,

\[
(r_B(D_s^0(2460)K^+))^2 < 0.027(0.033) \quad \text{and} \quad r_B(D_s^0(2460)K^+) < 0.30(0.36).
\]

The result for \((r_B(D_s^0(2460)K^+))^2 \) and the product branching fraction \( \mathcal{B}(B^+ \rightarrow D_s^0(2460)^0 K^+) \times \mathcal{B}(D_s^0(2460)^0 \rightarrow D^+ \pi^-) = (23.2 \pm 1.1 \pm 0.6 \pm 1.0 \pm 1.6) \times 10^{-4} \) at 90(95)% C.L. [11] give

\[
\mathcal{B}(B^+ \rightarrow D_s^0(2460)^0 K^+) \times \mathcal{B}(D_s^0(2460)^0 \rightarrow D^+ \pi^-) = (0.4 \pm 3.5 \pm 1.1 \pm 0.1) \times 10^{-5},
\]

< 6.3(7.5) \times 10^{-5} \text{ at } 90(95)\% \text{ CL}.

These are the first experimental results on this decay mode. Similarly for \( B^+ \rightarrow D^+ K^+ (892)^0 \rightarrow D^+ K^+ \pi^- \),

\[
\frac{\mathcal{B}(B^+ \rightarrow D^+ K^+(892)^0 \rightarrow D^+ K^+ \pi^-)}{\mathcal{B}(B^+ \rightarrow D^- K^+ \pi^+)} = -0.0079 \pm 0.0039(\text{stat}) \pm 0.0028(\text{syst}),
\]

< 0.0044(0.0055) \text{ at } 90(95)\% \text{ C.L.}

The measured value \( \mathcal{B}(B^+ \rightarrow D^- K^+ \pi^+) = (7.31 \pm 0.19 \pm 0.22 \pm 0.39) \times 10^{-5} \) [11] and the isospin relation \( \mathcal{B}(K^+(892)^0 \rightarrow K^+ \pi^-) = \frac{3}{2} \) give

\[
\mathcal{B}(B^+ \rightarrow D^+ K^+(892)^0) = (-8.7 \pm 4.3 \pm 3.1 \pm 0.4) \times 10^{-7},
\]

< 4.9(6.1) \times 10^{-7} \text{ at } 90(95)\% \text{ C.L.,}

where the third uncertainty is due to the normalization channel branching fraction. This result supersedes the previous limit, which was obtained with a subset of the data [12].

In summary, the rare \( B^+ \rightarrow D^+ K^+ \pi^- \) decay has been observed for the first time with 8\( \sigma \) significance, based on a data sample of 3.0 fb\(^{-1} \) of \( pp \) collision data collected with the LHCb detector. The Dalitz plot appears to be dominated by broad structures. Searches for \( B^+ \rightarrow D_s^0(2460)^0 K^+ \) and \( B^+ \rightarrow D^+ K^+(892)^0 \) decays have been carried out by weighting the data according to the decay angle distributions, but no significant signals are seen. These results indicate that further studies, with larger data samples, of the Dalitz plot distribution of this mode will be of interest to understand the potential for a measurement of \( \gamma \) from \( B^+ \rightarrow D K^+ \pi^0 \) decays.

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); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA).

We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (U.S.). We are indebted to the communities behind the multiple open source software packages on which we depend. We are also thankful for the computing resources and the access to software R&D tools provided by Yandex LLC (Russia). Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Sklodowska-Curie Actions and ERC (European Union), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR (Russia), GVA, XuntaGal and GENCAT (Spain), The Royal Society and Royal Commission for the Exhibition of 1851 (United Kingdom).
FIRST OBSERVATION OF THE RARE...


[11] R. Aaij et al. (LHCb Collaboration), First observation and amplitude analysis of the $B^\pm \to D^\mp K^\mp \pi^\pm$ decay, Phys. Rev. D 91, 092002 (2015).

[12] R. Aaij et al. (LHCb Collaboration), First evidence for the annihilation decay mode $B^\pm \to D^\ast_\pm \phi$, J. High Energy Phys. 02 (2013) 043.


PHYSICAL REVIEW D 93, 051101(R) (2016)


[22] R. Aaij et al. (LHCb Collaboration), Measurement of the $B^0 \to \mu^+\mu^−$ Branching Fraction and Search for $B^0 \to \mu^+\mu^−$ Decays at the LHCb Experiment, Phys. Rev. Lett. 111, 101805 (2013).

[23] R. Aaij et al. (LHCb Collaboration), Search for the lepton flavour violating decay $\tau^− \to \mu^+\mu^−\mu^−$, J. High Energy Phys. 02 (2015) 121.

[24] R. Aaij et al. (LHCb Collaboration), Search for the decay $B^0_s \to D^0 f_0(980)$, J. High Energy Phys. 08 (2015) 005.

[25] R. Aaij et al. (LHCb Collaboration), Search for the lepton-flavour violating decay $D^0 \to e^+\mu^−\mu^−$, Phys. Lett. B 754, 167 (2016).


[27] K. Abe et al. (Belle Collaboration), Study of $B^+ \to D^{\ast+} \pi^− (D^{\ast0} \to D^{\ast+} \pi^−)$ decays, Phys. Rev. D 69, 112002 (2004).

[28] B. Aubert et al. (BABAR Collaboration), Dalitz plot analysis of $B^+ \to D^+ \pi^−\pi^−$, Phys. Rev. D 79, 112004 (2009).


[30] R. Aaij et al. (LHCb Collaboration), First observation of the decays $B^+_s \to D^{\ast+}_s K^−\pi^+\pi^−$ and $B^0 \to D_{s1}^{\ast\pm}(2536)^\mp\pi^\pm$, Phys. Rev. D 86, 112005 (2012).


FIRST OBSERVATION OF THE RARE ...

PHYSICAL REVIEW D 93, 051101(R) (2016)


1 Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil
2 Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
3 Center for High Energy Physics, Tsinghua University, Beijing, China
4 LAPP, Université Savoie Mont-Blanc, CNRS/IN2P3, Annecy-Le-Vieux, France
5 Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, Clermont-Ferrand, France
6 CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
7 LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
8 LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris, France
FIRST OBSERVATION OF THE RARE …

PHYSICAL REVIEW D 93, 051101(R) (2016)

63Departamento de Fisica, Universidad Nacional de Colombia, Bogota, Colombia
(associated with LPNHE, Université Pierre et Marie Curie,
Université Paris Diderot, CNRS/IN2P3, Paris, France)
64Institut für Physik, Universität Rostock, Rostock, Germany
(associated with Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany)
65National Research Centre Kurchatov Institute, Moscow, Russia
(associated with Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia)
66Yandex School of Data Analysis, Moscow, Russia
67Instituto de Fisica Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain
(associated with Universitat de Barcelona, Barcelona, Spain)
68Van Swinderen Institute, University of Groningen, Groningen, The Netherlands
(associated with Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands)

†Deceased.
aAlso at Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil.
bAlso at Laboratoire Leprince-Ringuet, Palaiseau, France.
cAlso at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
dAlso at Università di Bari, Bari, Italy.
eAlso at Università di Bologna, Bologna, Italy.
fAlso at Università di Cagliari, Cagliari, Italy.
gAlso at Università di Ferrara, Ferrara, Italy.
hAlso at Università di Urbino, Urbino, Italy.
iAlso at Università di Modena e Reggio Emilia, Modena, Italy.
jAlso at Università di Genova, Genova, Italy.
kAlso at Università di Milano Bicocca, Milano, Italy.
lAlso at Università di Roma Tor Vergata, Roma, Italy.
mAlso at Università di Roma La Sapienza, Roma, Italy.
nAlso at Università della Basilicata, Potenza, Italy.
oAlso at AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland.
pAlso at LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
qAlso at Hanoi University of Science, Hanoi, Viet Nam.
rAlso at Università di Padova, Padova, Italy.
sAlso at Università di Pisa, Pisa, Italy.
tAlso at Scuola Normale Superiore, Pisa, Italy.
uAlso at Università degli Studi di Milano, Milano, Italy.