Observation of $B^0 \to D^0 K^+ K^-$ and Evidence for $B^{0}_s \to D^{0} K^{+} K^-$

R. Aaij et al.*

(LHCb Collaboration)

(Received 26 July 2012; published 26 September 2012)

The first observation of the decay $B^0 \to D^0 K^+ K^-$ is reported from an analysis of 0.62 fb$^{-1}$ of $pp$ collision data collected with the LHCb detector. Its branching fraction is measured relative to that of the topologically similar decay $B^0 \to D^0 \pi^+ \pi^-$. The as-yet unobserved decay $B^0 \to D^0 K^+ K^-$, where $D$ denotes a neutral charmed meson that is an admixture of $D^0$ and $\bar{D}^0$, gives a constraint on $\gamma$ with an uncertainty of around 20° [3]. Recent results from LHCb on $B^+ \to D K^+$ [4] have helped to reduce this uncertainty, but the use of additional channels to improve further the precision is of great interest. The first step in the program towards the measurement of $\gamma$ using the $B^0 \to D^{0} K^+ K^-$ decay is the observation of the channel. In this Letter the results of a search for neutral $B$ meson decays to $D^0 K^+ K^-$ are presented. The quantities measured include small contributions from decays to $D^0 K^+ K^-$ and $D^0 K^+ K^-$. The inclusion of charge conjugate modes is implied throughout.

The analysis uses 0.62 fb$^{-1}$ of LHC collision data at a center-of-mass energy of 7 TeV collected with the LHCb detector during 2011. In high energy $pp$ collisions all $b$ hadron species are produced, so both $B^0$ and $B^0_s$ decays are searched for simultaneously. The decay $B^0 \to D^0 K^+ K^-$ can be mediated by the decay diagrams shown in Fig. 1. These are a $W$-exchange diagram similar to that for the decay $B^0 \to D^0 K^+ K^-$ [9,10] [in this case an excited state that decays to $D^0 K^-$, such as $D^{*+} (2573)$, would be produced], and a color-suppressed tree diagram producing $D^0 h^0$, where $h^0$ is a light unflavored meson such as $a_0(980)$ that subsequently decays to $K^+ K^-$. Related $B$ decays with $s \bar{s}$ production, $B^+ \to D^{0} K^+ K^{(*)}$ [11] and $B^+ \to D^{(*)} K^+ K^+ K^+$ [12,13], have been measured to have branching fractions of $O(10^{-4})$.

The LHCb detector [14] 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 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream. The combined tracking system has momentum resolution $\Delta p/p$ that varies from 0.4% at 5 GeV/c to 0.6% at 100 GeV/c, and impact parameter (IP) resolution of 20 $\mu$m for tracks with high transverse momentum ($p_T$). Charged hadrons are identified using two ring-imaging Cherenkov (RICH) detectors. Photon, electron, and hadron candidates 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 trigger 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. In this analysis, signal candidates are accepted if one of the final state particles created a cluster in the calorimeter with sufficient transverse energy to fire the hardware trigger. Events that are triggered at hardware level by the decay products of the other $b$ hadron in the $pp \to b \bar{b} X$ event are also retained.

The software trigger requires characteristic signatures of $b$ hadron decays: at least one track, with high $p_T$ and a large IP with respect to any primary interaction vertex (PV) [15], that subsequently forms part of a two-, three- or

*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.
four-track secondary vertex with a high sum of the $p_T$ of the tracks and significant displacement from the PV [16]. In the offline analysis, the software trigger decision is required to be due to the candidate signal decay.

Candidates that are consistent with the decay chain $B_{(s)}^0 \rightarrow D^0 K^+ K^-$, $D^0 \rightarrow K^+ \pi^-$ are selected. In order to reduce systematic uncertainties in the measurement, the topologically similar decay $\bar{D}^0 \pi^+ \pi^-$, which has previously been well studied [17,18], is used as a normalization channel. The $D^0$ candidate invariant mass is required to satisfy $1844 < m_{K\pi} < 1884 \text{ MeV}/c^2$. Tracks are required to be consistent with either the kaon or pion hypothesis, as appropriate, based on particle identification (PID) information from the RICH detectors. All other selection criteria were tuned on the $\bar{D}^0 \pi^+ \pi^-$ channel. The large yields available in the normalization sample allow the selection to be based on data, though the efficiencies are determined using Monte Carlo (MC) simulated events. For the simulation, $pp$ collisions are generated using PYTHIA 6.4 [19] with a specific LHCb configuration [20]. Decays of hadronic particles are described by EVTGEN [21]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [22] as described in Ref. [23].

Selection requirements are applied to obtain a clear signal peak in the $D^0 \pi^+ \pi^-$ normalization channel. The selection includes criteria on the track quality of the tracks forming the signal candidate, their $p$, $p_T$ and inconsistency with originating from the PV ($\chi^2_{D^0}$). Requirements are also placed on the corresponding variables for candidate composite particles ($\bar{D}^0$, $B_{(s)}^0$) together with restrictions on the consistency of the decay fit ($\chi^2_{\text{vertex}}$), the flight distance significance ($\chi^2_{\text{flight}}$), and the angle between the momentum vector and the line joining the PV to the $B_{(s)}^0$ vertex ($\cos\theta_{\text{dir}}$) [24].

Further discrimination between signal and background categories is achieved by calculating weights for the remaining $D^0 \pi^+ \pi^-$ candidates [25]. The weights are used by the NEUROBAYES neural network package [26] to maximize the separation between categories. A total of 15 variables are used in the network. They include the $\chi^2_{D^0}$ of the four candidate tracks, the $\chi^2_{\text{IP}}$, $\chi^2_{\text{vertex}}$, $\chi^2_{\text{flight}}$, and $\cos\theta_{\text{dir}}$ of the $\bar{D}^0$ and $B_{(s)}^0$ candidates, and the $B_{(s)}^0$ candidate $p_T$. Variables describing the $p_T$ asymmetry and track multiplicity in a 1.5 rad cone [4] around the $B_{(s)}^0$ candidate flight direction are also used. The input quantities to the neural network only depend weakly on the kinematics of the $B_{(s)}^0$ decay. A requirement on the network output is imposed that reduces the combinatorial background by an order of magnitude while retaining about 80% of the signal. No bias is observed by using data driven selection requirements.

To improve the $B_{(s)}^0$ candidate invariant mass resolution, the four-momenta of the tracks from the $D^0$ candidate are adjusted so that their combined invariant mass matches the world average value [3]. An additional $B^0$ mass constraint is applied in the calculation of the Dalitz plot coordinates, which are used in the determination of event-by-event efficiencies. A small fraction ($\sim 5\%$ within the mass range described below) of candidates with invariant masses far from the $B_{(s)}^0$ peak fail the mass constrained fit, and are removed from the analysis.

To remove a large potential background from $B^0 \rightarrow D^+ (2010) \pi^+$, candidates in the $D^0 \pi^+ \pi^-$ sample are rejected if $m_{D\pi}-m_D$ (for either pion charge) lies within $\pm 2.5 \text{ MeV}/c^2$ of the nominal $D^+ - \bar{D}^0$ mass difference [3]. Candidates in the $\bar{D}^0 K^+ K^-$ sample are also rejected if the invariant mass difference calculated under the pion mass hypothesis satisfies the same criterion. This removes 3.3% of $D^0 K^+ K^-$ candidates. Less than 1% of $\bar{D}^0 K^+ K^-$ combinations are rejected by requiring that the pion from the $\bar{D}^0$ candidate together with the two kaons do not form an invariant mass in the range 1950–1975 MeV/$c^2$, which removes potential background from $B^0 \rightarrow D^+_s K^+ \pi^0$ decays.

After all selection requirements are applied, less than 1% of events with at least one candidate also contain a second candidate. Such multiple candidates are retained and treated the same as other candidates; the associated systematic uncertainty is negligible.

In addition to combinatorial background, candidates may be formed from misidentified or partially reconstructed $B_{(s)}^0$ decays, or from $B_{(s)}^0$ decays to identical final states but without intermediate charmed mesons (referred to below as charmless peaking background). Contributions from partially reconstructed decays are reduced by requiring the invariant mass of the $B_{(s)}^0$ candidate to be above 5150 MeV/$c^2$. Sources of misidentified backgrounds are investigated using simulation. Most potential sources are found to have a broad invariant mass distribution, and are absorbed in the combinatorial background shape used in the fit described below. Backgrounds from $\Lambda^+_b \rightarrow \bar{D}^0 \bar{p} K^+$ and $\Lambda^0_b \rightarrow \bar{D}^0 \bar{p} \pi^+$ [27], $B^0 \rightarrow \bar{D}^0 K^+ \pi^-$ and

![FIG. 1. Sample decay diagrams that contribute to the $B^0 \rightarrow D^0 K^+ K^-$ final state via (left) $W$ exchange, (right) rescattering from a color-suppressed decay.](image-url)
The contributions from charmless peaking background are investigated using candidates, reconstructed without the $D^0$ mass constraint, in sideband regions around the $D^0$ mass. The distributions are fitted with double Gaussian signal and linear background probability density functions (PDFs). Extrapolating to the $D$ mass signal region, 773 ± 30 (126 ± 18) charmless background decays are expected in the $B^0 \rightarrow D^0 \pi^+ \pi^-(B^0 \rightarrow D^0 K^+ K^-)$ distributions. No peaking background is observed in the $B^0$ region.

The signal yields are obtained from unbinned maximum likelihood fits to the $D^0 \pi^+ \pi^-$ and $D^0 K^+ K^-$ invariant mass distributions in the range 5150–5600 MeV/$c^2$. There are 14214 $D^0 \pi^+ \pi^-$ and 2990 $D^0 K^+ K^-$ candidates. The $D^0 \pi^+ \pi^-$ fit includes a double Gaussian shape for signal, together with an exponential component for partially reconstructed background, and a PDF for $\Lambda_b^0 \rightarrow D^0 X$ decays modeled using a nonparametric function obtained from simulation. The $D^0 K^+ K^-$ fit includes a second double Gaussian component to account for the possible presence of both $B^0$ and $B^0_s$ decays, and peaking background PDFs for $\Lambda_b^0 \rightarrow D^0 p K^+$, $B^0 \rightarrow D^0 K^+ \pi^-$, and $B^0_s \rightarrow D^0 K^+$, all modeled using nonparametric functions. The shape of the combinatorial background is essentially linear, but is multiplied by a function that accounts for the fact that candidates with high invariant masses are more likely to fail the $B^0$ mass constrained fit.

The result of the fit to $D^0 \pi^+ \pi^-$ candidates is shown in Fig. 2. There are nine free parameters in this fit: the double Gaussian peak position, core width and fraction in the core, the linear slope of the combinatorial background and the exponential shape parameter of the partially reconstructed background, and the yields of the four categories. The relative width of the broader to the core Gaussian component is constrained within uncertainty to the value obtained in simulation. The fit yields 8060 ± 150 $B^0 \rightarrow D^0 \pi^+ \pi^-$ decays, including charmless peaking background.

Since the fit to $D^0 K^+ K^-$ candidates, shown in Fig. 2, has more components, additional constraints are imposed in order to improve the stability of the results. The parameters of the double Gaussian shapes are constrained to be identical for $B^0$ and $B^0_s$ signals, with an offset in their mean values fixed to the known $B^0$-$B^0_s$ mass difference [3]. The slope of the combinatorial component is constrained to the value obtained in the fit to $D^0$ mass sideband events. The exponential shape parameter is constrained to the value obtained in the $D^0 \pi^+ \pi^-$ fit. The fit yields 558 ± 49 $B^0 \rightarrow D^0 K^+ K^-$ decays, including charmless peaking background, and 104 ± 29 $B^0_s \rightarrow D^0 K^+ K^-$ decays. All background yields are consistent with their expectations within uncertainties.

The ratio of branching fractions is obtained after subtracting the charmless peaking background, and applying event-by-event efficiencies as a function of the Dalitz plot position.

\[
\frac{\mathcal{B}(B^0 \rightarrow D^0 K^+ K^-)}{\mathcal{B}(B^0 \rightarrow D^0 \pi^+ \pi^-)} = R(B^0, B^0) = \frac{N_{\text{corr}}(D^0 K^+ K^-)}{N(D^0 K^+ K^-)} \left(1 - \frac{N_{\text{peak}}(D^0 K^+ K^-)}{N(D^0 K^+ K^-)}\right)
\]

where $N$ is the yield obtained from the fit, $N_{\text{peak}}$ is the charmless peaking background contribution, and the efficiency corrected yield $N_{\text{corr}} = \sum W_i / e_{\text{tot}}^i$. Here the index $i$ runs over all candidates in the fit range, $W_i$ is the signal weight for candidate $i$ [25] from the fit shown in Fig. 2, and $e_{\text{tot}}^i$ is the efficiency for candidate $i$, which depends only on its Dalitz plot position. The statistical uncertainty on the branching fraction ratio incorporates the effects of the shape parameters that are allowed to vary in the fit, the dilution due to event weighting, and the charmless peaking.

FIG. 2 (color online). Fits to the $B^0$ candidate invariant mass distributions for the (left) $D^0 \pi^+ \pi^-$ and (right) $D^0 K^+ K^-$ samples. Data points are shown in black, the full fitted PDFs as solid blue lines and the components as detailed in the legends. Yields of the partially reconstructed and peaking backgrounds are all small for the $D^0 K^+ K^-$ sample.
The ratio of fragmentation fractions is instead determined as

\[ \frac{f_s}{f_d} = 0.267^{+0.021}_{-0.020} \]

The ratio of branching fractions is

\[ \frac{\mathcal{B}(B^0 \to \bar{D}^0 K^+ K^-)}{\mathcal{B}(B^0 \to D^0 K^+ K^-)} = R(B^0, B^0) = \left( \frac{f_s}{f_d} \right)^{-1} \frac{N(B^0 \to DKK)}{N(B^0 \to DKK) - N_{\text{peak}}(B^0 \to DKK)} . \]  

The Dalitz plot distributions obtained from the signal weights are shown in Fig. 3. The $B^0 \to \bar{D}^0 \pi^+ \pi^-$ distribution shows contributions from the $\rho^0(770)$ and $f_2(1270)$ resonances (upper diagonal edge of the Dalitz plot) and from the $D_s^+(2460)$ state (horizontal band), as expected from previous studies of this decay [17,18]. The $B^0 \to \bar{D}^0 K^+ K^-$ distribution shows a possible contribution from the $D_s^+(2573)$ resonance, together with an enhancement of events at low $K^+ K^-$ invariant mass (upper diagonal edge).

The branching fraction of the $B^0_s$ decay to $\bar{D}^0 K^+ K^-$ is measured relative to that of $B^0$ to the same final state. Because of the low yield in this decay, an event-by-event efficiency correction is not used. The ratio of branching fractions is instead determined as

\[ \frac{\mathcal{B}(B^0_s \to \bar{D}^0 K^+ K^-)}{\mathcal{B}(B^0 \to D^0 K^+ K^-)} = R(B^0_s, B^0) = \left( \frac{f_s}{f_d} \right)^{-1} \frac{N(B^0_s \to DKK)}{N(B^0 \to DKK) - N_{\text{peak}}(B^0 \to DKK)} . \]  

The Dalitz plot distributions for (left) $B^0 \to \bar{D}^0 \pi^+ \pi^-$ and (right) $B^0 \to \bar{D}^0 K^+ K^-$ obtained from the signal weights. Note that these distributions contain contributions from charmless peaking backgrounds.

Most potential systematic effects cancel in the ratio. The PID efficiency is measured using a control sample of $D^- \to \bar{D}^0 \pi^-$, $\bar{D}^0 \to K^+ \pi^-$ decays to obtain background-subtracted efficiency tables for kaons and pions as functions of their $p$ and $p_T$ [28]. The kinematic properties of the tracks in signal decays are obtained from simulation, allowing the PID efficiency for each event to be obtained from the tables taking into account the correlation between the $p$ and $p_T$ values of the two tracks. The other contributions to the efficiency (detector acceptance, selection criteria, and trigger effects) are determined from simulation, and validated using data. All are found to be approximately constant across the Dalitz plane, apart from some modulations seen near the kinematic boundaries.

**FIG. 3.** Dalitz plot distributions for (left) $B^0 \to \bar{D}^0 \pi^+ \pi^-$ and (right) $B^0 \to \bar{D}^0 K^+ K^-$. Background subtraction. Most potential systematic effects cancel in the ratio.
A number of cross-checks are performed to test the stability of the result. The data sample is divided by dipole magnet polarity, data-taking period, and trigger category. Candidates were divided based upon the hardware trigger decision into three groups; events in which a particle from the signal decay created a large enough cluster in the calorimeter to fire the trigger, events that were triggered independently of the signal decay, and those events that were triggered by both the signal decay and the rest of the event. The neural network and PID requirements are tightened and loosened. The PID efficiency is evaluated using the kinematic properties from \( \bar{D}^0 \pi^+ \pi^- \) data instead of from simulation. The charmless peaking background contribution is determined from the upper and lower \( \bar{D}^0 \) mass sidebands separately. All give consistent results.

The significances of the signals are obtained from the changes in likelihood in fits to data with and without signal components, after accounting for systematic uncertainties and for charmless peaking background in \( B^0 \to \bar{D}^0 K^+ K^- \) only. They are found to be 5.8\( \sigma \) and 3.8\( \sigma \) for \( B^0 \to \bar{D}^0 K^+ K^- \) and \( B^0 \to \bar{D}^0 K^+ K^- \), respectively.

In summary, the decay \( B^0 \to \bar{D}^0 K^+ K^- \) has been observed for the first time, and its branching fraction relative to that of \( B^0 \to \bar{D}^0 \pi^+ \pi^- \) is measured to be

\[
\frac{\mathcal{B}(B^0 \to \bar{D}^0 K^+ K^-)}{\mathcal{B}(B^0 \to \bar{D}^0 \pi^+ \pi^-)} = 0.056 \pm 0.011 \pm 0.007,
\]

where the first uncertainty is statistical and the second is systematic. Using the known value of \( \mathcal{B}(B^0 \to \bar{D}^0 \pi^+ \pi^-) = (8.4 \pm 0.4 \pm 0.8) \times 10^{-4} \) [17], this gives

\[
\mathcal{B}(B^0 \to \bar{D}^0 K^+ K^-) = (4.7 \pm 0.9 \pm 0.6 \pm 0.5) \times 10^{-5},
\]

where the third uncertainty arises from \( \mathcal{B}(B^0 \to \bar{D}^0 \pi^+ \pi^-) \). Evidence for the \( B^0 \to \bar{D}^0 K^+ K^- \) decay has also been found, with relative branching fraction

\[
\frac{\mathcal{B}(B^0 \to \bar{D}^0 K^+ K^-)}{\mathcal{B}(B^0 \to \bar{D}^0 K^+ K^-)} = 0.90 \pm 0.27 \pm 0.20.
\]

A future study of the Dalitz plot distributions of these decays will provide insight into the dynamics of hadronic \( B \) decays. In addition, the \( B^0 \to \bar{D}^0 K^+ K^- \) decay may be used to measure the \( CP \) violating phase \( \gamma \).

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 CERN and at the LHCb institutes, and acknowledge support from the National Agencies: CAPES, CNPq, FAPERJ, and FINEP (Brazil); CERN; NSFC (China); CNRS/IN2P3 (France); BMBF, DFG, HGF, and MPG (Germany); INFN (Italy); FOM and NWO (Netherlands); SCSR (Poland); ANCS (Romania); MinES of Russia and Rosatom (Russia); MICINN, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (U.K.); NSF (U.S.). We also acknowledge the support received from the ERC under FP7 and the Region Auvergne.