Aaij, R. et al. (2013) First observations of $\bar{B}_s^0 \rightarrow D^0 D^-$, $D^0 sD^−$ and $D^0 D^0$ decays. Physical Review D, 87 (9). Art. 092007. ISSN 1550-7998

Copyright © 2013 CERN, for the benefit of the LHCb collaboration

http://eprints.gla.ac.uk/82349/

Deposited on: 9 January 2014
First observations of $\bar{B}_s^0 \to D^+ D^-, D_s^+ D^-$ and $D^0 \bar{D}^0$ decays

R. Aaij et al.*
(LHCb Collaboration)
(Received 23 February 2013; published 21 May 2013)

First observations and measurements of the branching fractions of the $\bar{B}_s^0 \to D^+ D^-$, $\bar{B}_s^0 \to D_s^+ D^-$ and $\bar{B}_s^0 \to D^0 \bar{D}^0$ decays are presented using 1.0 fb$^{-1}$ of data collected by the LHCb experiment. These branching fractions are normalized to those of $\bar{B}^0 \to D^+ D^-$, $B^0 \to D_s^- D_s^+$ and $B^- \to D^0 D_s^+$, respectively. An excess of events consistent with the decay $\bar{B}^0 \to D^0 \bar{D}^0$ is also seen, and its branching fraction is measured relative to that of $B^- \to D^0 D_s^-$. Improved measurements of the branching fractions $\mathcal{B}(\bar{B}_s^0 \to D^+_s D^-_s)$ and $\mathcal{B}(B^- \to D_s^- D^-_s)$ are reported, each relative to $\mathcal{B}(B^0 \to D^- D^+_s)$. The ratios of branching fractions are $\frac{\mathcal{B}(\bar{B}_s^0 \to D^+_s D^-_s)}{\mathcal{B}(B^- \to D^- D^+_s)} = 1.08 \pm 0.20 \pm 0.10$, $\frac{\mathcal{B}(\bar{B}_s^0 \to D^- D^-_s)}{\mathcal{B}(B^- \to D^- D^+_s)} = 0.050 \pm 0.008 \pm 0.004$, $\frac{\mathcal{B}(\bar{B}^0 \to D^0 \bar{D}^0)}{\mathcal{B}(B^- \to D^- D^+_s)} = 0.019 \pm 0.003 \pm 0.003$, $\frac{\mathcal{B}(\bar{B}^0 \to D^- D^-_s)}{\mathcal{B}(B^- \to D^- D^+_s)} < 0.0024$ at 90% CL, $\frac{\mathcal{B}(\bar{B}_s^0 \to D^+_s D^-_s)}{\mathcal{B}(\bar{B}_s^0 \to D^- D^-_s)} = 1.22 \pm 0.02 \pm 0.07$, where the uncertainties are statistical and systematic, respectively.

DOI: 10.1103/PhysRevD.87.092007
PACS numbers: 13.25.Hw

I. INTRODUCTION

Double-charm decays of $B$ mesons can be used to probe the Cabibbo-Kobayashi-Maskawa matrix [1,2] elements and provide a laboratory to study final state interactions. The time-dependent $CP$ asymmetry in the $B^0 \to D^+ D^-$ decay provides a way to measure the $B^0$ mixing phase [3,4], where information from other double-charm final states can be used to account for loop (penguin) contributions and other nonfactorizable effects [5–9]. Double-charm decays of $B$ mesons can also be used to measure the weak phase $\gamma$, assuming $U$-spin symmetry [10,11]. The purely $CP$-even $B_s^0 \to D_s^+ D_s^-$ decay is also of interest, as it can be used to measure the $B_s^0$ mixing phase. Moreover, a lifetime measurement using the $B_s^0 \to D_s^+ D_s^-$ decay provides complementary information on $\Delta \Gamma_s$ [11–14] to that obtained from direct measurements [15], or from lifetime measurements in other $CP$ eigenstates [16,17].

The study of $B \to D D'$ decays $^1$ can also provide a better theoretical understanding of the processes that contribute to $B$ meson decay. Feynman diagrams contributing to the decays considered in this paper are shown in Fig. 1. The $B_s^0 \to D^0 \bar{D}^0$, $\bar{B}_s^0 \to D^+ D^-$ and $\bar{B}_s^0 \to D^0 \bar{D}^0$ decays are mediated by the $W$-exchange amplitude, along with penguin-annihilation contributions and rescattering [18]. The only other observed $B$ meson decays of this type are $B^0 \to D_s^{(*)} K^{(*)-}$ and $B_s^0 \to \pi^+ \pi^-$, with branching fractions of the order of $10^{-5}$ [19] and $10^{-6}$ [20], respectively. Predictions of the $\bar{B}_s^0 \to D^+ D^-$ branching fraction using perturbative approaches yield $\sim 5.0 \times 10^{-4}$ [21], while the use of nonperturbative approaches has led to a larger value of $4.5 \times 10^{-3}$ [22]. More recent phenomenological studies, which assume a dominant contribution from rescattering, predict a significantly lower branching fraction of $\mathcal{B}(\bar{B}_s^0 \to D^+ D^-) = \mathcal{B}(\bar{B}_s^0 \to D^0 \bar{D}^0) = (7.8 \pm 4.7) \times 10^{-5}$ [18].

This paper reports the first observations of the $\bar{B}_s^0 \to D^+ D^-$, $\bar{B}_s^0 \to D_s^+ D^-$ and $B^0 \to D^0 \bar{D}^0$ decays, and measurements of their branching fractions normalized relative to those of $B^0 \to D^+ D^-$, $B^0 \to D_s^- D_s^+$ and $B^- \to D^0 D_s^+$, respectively. An excess of events consistent with $\bar{B}_s^0 \to D^0 \bar{D}^0$ is also seen, and its branching fraction is reported. Improved measurements of the ratios of branching fractions $\mathcal{B}(\bar{B}_s^0 \to D_s^+ D_s^-) / \mathcal{B}(B^0 \to D^- D_s^+)$ and $\mathcal{B}(B^- \to D^0 D_s^-) / \mathcal{B}(B^0 \to D^- D_s^+)$ are also presented. All results are based upon a data sample corresponding to an integrated luminosity of 1.0 fb$^{-1}$ of $pp$ collision data at

---

*Full author list given at 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.

$^1$Throughout this paper, the notation $D$ (or $D'$) is used to refer to a $D^+$, $D^0$ or $D_s^+$ meson, and $B$ represents either a $B^0$, $B^-$ or $B_s^0$ meson.

FIG. 1 (color online). Feynman diagrams contributing to the double-charm final states discussed in this paper. They include (a) tree, (b) $W$ exchange and (c) penguin diagrams.
$\sqrt{s} = 7$ TeV recorded by the LHCb experiment in 2011. Inclusion of charge conjugate final states is implied throughout.

II. DETECTOR, TRIGGER AND DATA SAMPLES

The LHCb detector [23] 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 a momentum resolution ($\Delta p/p$) that varies from 0.4% at 5 GeV/c to 0.6% at 100 GeV/c, and an impact parameter (IP) resolution of 20 $\mu$m for tracks with high transverse momentum ($p_T$). The impact parameter is defined as the distance of closest approach of a given particle to the primary $pp$ interaction vertex (PV). Charged particles are identified by two ring-imaging Cherenkov detectors [24]. Discrimination of photons, electrons and charged hadrons is provided 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 [25] consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage that performs a partial event reconstruction (only tracks with $p_T > 0.5$ GeV/c are reconstructed and used). The software trigger requires a two-, three- or four-track secondary vertex with a large track $p_T$ sum and a significant displacement from any of the reconstructed PVs. At least one track must have $p_T > 1.7$ GeV/c and IP $\chi^2$ greater than 16 with respect to all PVs. The IP $\chi^2$ is defined as the difference between the $\chi^2$ of the PV reconstructed with and without the considered particle. A multivariate algorithm [26] is used to identify secondary vertices that originate from the decays of $b$ hadrons.

Signal efficiencies and specific backgrounds are studied using simulated events. Proton-proton collisions are generated using PYTHIA 6.4 [27] with a specific LHCb configuration [28]. Decays of hadronic particles are described by EVTGEN [29] in which final state radiation is generated using PHOTOS [30]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [31] as described in Ref. [32]. Efficiencies for identifying $K^+$ and $\pi^+$ mesons are determined using $D^{*+}$ calibration data, with kinematic quantities reweighted to match those of the signal particles [24].

Signal $B$ candidates are formed by combining pairs of $D$ meson candidates reconstructed in the following decay modes: $D^0 \rightarrow K^- \pi^+$ or $K^- \pi^+ \pi^- \pi^+$, $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D^{*+} \rightarrow K^- \pi^+ \pi^+ \pi^-$. The $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decay is only used for $\bar{B}^0 \rightarrow D^{*0} \bar{D}^0$ candidates, where a single $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decay in the final state is allowed, which approximately doubles the total signal efficiency. A refit of signal candidates with $D$ mass and vertex constraints is performed to improve the $B$ mass resolution.

Due to the similar kinematics of the $D^+ \rightarrow K^- \pi^+ \pi^+$, $D^+_s \rightarrow K^- K^+ \pi^+$ and $\Lambda^+_c \rightarrow p K^- \pi^+$ decays, there is cross feed between various $b$-hadron decays that have two charm particles in the final state. Cross feed between $D^+$ and $D^{*+}$ occurs when the $K^- \pi^+ h^+$ invariant mass is within 25 MeV/$c^2$ ($\sim$ 3 times the experimental resolution) of both the $D^+$ and $D^{*+}$ masses under the $h^+ = \pi^+$ and $h^+ = K^+$ hypotheses, respectively. In such cases, an arbitration is performed as follows: if either $|M(K^+ K^-) - m_\phi| < 10$ MeV/$c^2$ or $h^+$ satisfies a stringent kaon particle identification (PID) requirement, the $D$ candidate is assigned to be a $D^{*+}$ meson. Conversely, if $h^+$ passes a stringent pion PID requirement, the $D$ candidate is taken to be a $D^+$ meson. Candidates that do not pass either of these selections are rejected. A similar veto is applied to $D^0$ and $D^{*0}$ decays that are consistent with the $\Lambda^+_c \rightarrow p K^- \pi^+$ decay hypothesis if the proton is misidentified as a $\pi^+$ or $K^+$, respectively. The efficiencies of these $D$ selections are

<table>
<thead>
<tr>
<th>$B$ candidates</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{B}^0 \rightarrow D^{*+} D^-_s$</td>
<td>0.140</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*+} D^-_s$ (loose selection)</td>
<td>0.130</td>
</tr>
<tr>
<td>$\bar{B}^0 \rightarrow D^{*0} \bar{D}^0$, $K^- \pi^+ \pi^+$</td>
<td>0.447</td>
</tr>
<tr>
<td>$B^- \rightarrow D^{*0} D^-_s$</td>
<td>0.238</td>
</tr>
</tbody>
</table>

092007-2
determined using simulated signal decays to model the
kinematics of the decay and \( D^{*+} \rightarrow D^0 \pi^+ \) calibration
data for the PID efficiencies. Their values are given
in Table I.

To suppress contributions from non-\( D \bar{D}' \) final states, the
reconstructed \( D \) decay vertex is required to be downstream
of the reconstructed \( B \) decay vertex, and the \( B \) and \( D \) decay
vertices are required to have a vertex separation (VS) \( \chi^2 \)
larger than two. Here, the VS \( \chi^2 \) is the difference in \( \chi^2 \)
between the nominal \( B \) vertex fit and a vertex fit where the
\( D \) is assumed to have zero lifetime. The efficiencies of this
set of requirements are obtained from simulation and are
included in Table I.

To further improve the purity of the \( B \rightarrow D \bar{D}' \) samples, a
boosted decision tree (BDT) discriminant is used to
distinguish signal \( D \) mesons from backgrounds [33, 34].

The BDT uses five variables for the \( D \) meson and 23 for
each of its children. The variables include kinematic quanti-
ties, track quality, and vertex and PID information. The
signal and background distributions used to train the BDT
are obtained from \( B^0 \rightarrow D^+ \pi^- \), \( B^- \rightarrow D^0 \pi^- \) and \( B^0_s \rightarrow D_s^+ \pi^- \) decays from data. The signal distributions are
background subtracted using weights [35] obtained from
a fit to the \( B \) candidate invariant mass distribution. The
background distributions are taken from the high \( B \) mass
sidebands in the same data sample.

It is found that making a requirement on the product
of the two \( D \) meson BDT responses provides better
discrimination than applying one to each BDT response
individually. The optimal BDT requirement in each decay is
chosen by maximizing \( N_S/\sqrt{N_S + N_B} \). The number
of signal events, \( N_S \), is computed using the known
(or estimated, if unknown) branching fractions, selection
efficiencies from simulated events, and the BDT efficien-
cies from the \( B^0 \rightarrow D^+ \pi^- \), \( B^- \rightarrow D^0 \pi^- \) and \( B^0_s \rightarrow D_s^+ \pi^- \) calibration samples, reweighted to account for
small differences in kinematics between the calibration
and signal samples. The number, \( N_B \), is the expected
background yield for a given BDT requirement. To obtain
the BDT efficiency in a given signal mode, the kinematical
properties and correlations between the two \( D \) mesons are
taken from simulation, while the actual BDT response
distributions are obtained from \( B \rightarrow D \pi^- \) data. The result-
ing optimal BDT efficiencies are listed in Table I.

For the purpose of measuring \( B(B^0 \rightarrow D_s^+ D_s^-)/B(B^0 \rightarrow D^- D_s^+) \), the BDT optimization leads to loose
BDT requirements since the expected yields are relatively
large. On the other hand, for \( B(B^0 \rightarrow D_s^+ D^-)/B(B^0 \rightarrow D^- D_s^+) \), the expected signal yield of \( B^0 \rightarrow D_s^+ \pi^- \) decays is small; in this case both the signal
and normalization modes are required to pass the same tighter
BDT requirement. The different BDT selections applied to
the \( B^0 \rightarrow D^- D_s^+ \) decay are referred to as the “loose
selection” and the “tight selection.” Since the final state
is identical for the tight selection, the BDT efficiency
cancels in the ratio of branching fractions, and is not
included in Table I.

For \( B^0_{(s)} \rightarrow D^0 \bar{D}' \) candidates, a peaking background
from \( B \rightarrow D^{*+} \pi^- \rightarrow (D^0 \pi^+) \pi^- \) decays, where the \( \pi^+ \)
is misidentified as a \( K^+ \), is observed. This contribution is
removed by requiring the mass difference, \( M(K^- \pi^+ \pi^-) -
M(K^- \pi^+) > 150 \text{ MeV}/c^2 \), where the \( K^+ \) in the recon-
structed decay is taken to be a \( \pi^+ \). After the final selection,
around 2\% of events in the \( B^0_{(s)} \rightarrow D_s^+ D_s^- \) decay mode
contain multiple candidates; for all other modes the multi-
tiple candidate rate is below 1\%. All candidates are kept for
the final analysis.

For the ratios of branching fractions between modes
with identical final states, no requirements are made on
the hardware trigger decision. When the final states differ,
a trigger selection is applied to facilitate the determination
of the relative trigger efficiency. The selection requires that
either (i) at least one of the tracks from the reconstructed
signal decay is associated with energy depositions in the
calorimeters that passed the hardware trigger requirements,
or (ii) the event triggered independently of the signal
decay particles, e.g., on the decay products of the other \( b \)
hadron in the event. A small fraction (\( ~5\% \)) of events are
triggered by a combination of both the signal \( b \)-hadron
daughters and one or more other particles in the event.
These events are discarded.

III. SIGNAL AND BACKGROUND SHAPES

To determine the signal yields, the mass distributions are
parameterized as the sum of two Crystal Ball (CB) func-
tions [36], which account for non-Gaussian tails on both
sides of the signal peak. The asymmetric shapes account
for both non-Gaussian mass resolution effects (on both
sides) and energy loss due to final state radiation. The
two CB shapes are constrained to have equal area and a
common mean. Separate sets of shape parameters are
determined for \( B^0 \rightarrow D^+ D_s^- \), \( B^0_s \rightarrow D_s^+ D_s^- \) and \( B^0_s \rightarrow D^- D_s^+ \) using simulated signal decays, although their shapes
are very similar. In the fits to data, the signal shape pa-
rameters are fixed to the simulated values, except for a
smearing factor that is added in quadrature to the widths
from simulation. This number is allowed to vary indepen-
dently in each fit, but is consistent with about 4.6 \( \text{MeV}/c^2 \)
across all modes, resulting in a mass resolution of about
9 \( \text{MeV}/c^2 \). For the more rare \( B^0_{(s)} \rightarrow D^0 \bar{D}' \) and \( B^0_{(s)} \rightarrow D^{*+} D^- \) decay modes, the \( B^0_{(s)} \rightarrow D_s^+ D_s^- \) signal shape parameters are used. In determining the signal significan-
ces, the signal shape is fixed to that for \( B^0_s \rightarrow D_s^+ D_s^- \),
including an additional smearing of 4.6 \( \text{MeV}/c^2 \). The impact of using the \( B^0 \rightarrow D^- D_s^+ \) or \( B^- \rightarrow D^0 D_s^- \) signal shapes on the signal significances is negligible.

Several specific backgrounds contribute to the \( D \bar{D}' \)
mass spectra. In particular, decays such as \( B \rightarrow D^{(s)} \bar{D}' \),
where the \( D^* \) mesons decay through pion or photon
emission, produce distinct structures in all decays under

092007-3
consideration (due to angular momentum conservation). The shapes of these backgrounds are derived from simulation, which are corrected for known resolution differences between data and simulated events, and then fixed in fits to the data. The relative yield of the two peaks in the characteristic structure from the decay $D^+ \rightarrow D_0^0 \pi^+$ is allowed to vary freely, to enable better modeling of the background in the low mass region. Since these events are below the signal peak by at least the pion mass, the impact on the signal yield determinations is negligible.

A source of peaking background that contributes to $B \rightarrow D_0^+ D_s^+$ modes are the $B \rightarrow D K^{*0} K^+ \rightarrow D K^- \pi^+ K^+$ decays, where the $K^{*0} K^+$ is not produced in a $D_s^+$ decay. Although the branching fractions for these decays [37] are about twice as large as that of the $B \rightarrow D D_s^+$ decay channel, the 25 MeV/c$^2$ mass window around the known $D_s^+$ mass and the VS $\chi^2 > 2$ requirement reduces this contribution to about 1% of the signal yield. This expectation is corroborated by studying the $D_s^+$ candidate mass sidebands. The shape of this background is obtained from simulation, and is described by a single Gaussian function which has a width about 2.5 times larger than that of the signal decay and peaks at the nominal $B$ meson mass. The larger width in this decay than in the signal mode is a result of the $D$ mass and vertex-constrained fit applied in the $B$ reconstruction.

After the charm cross-feed vetoes (see Sec. II), the cross-feed rate from $B^0 \rightarrow D^+ D_s^+$ decays into the $B_s \rightarrow D_s^+ D_s^-$ sample is $(0.7 \pm 0.2)%$. The shape of this misidentification background is obtained from simulation. A similar cross-feed background contribution from $\Lambda_b \rightarrow \Lambda_s^+ D_s^-$ decays is also expected due to events passing the $\Lambda_b$ veto. Taking into account the observed yields of these decays in data, we fix the $B^0 \rightarrow D^- D_s^+$ and $\Lambda_b \rightarrow \Lambda_s^+ D_s^-$ cross feed yields to 35 and 15 events, respectively. Investigation of the $D$ mass sidebands reveals no additional contributions from non-$D\bar{D}$ backgrounds.

The combinatorial background shape is described by an exponential function whose slope is determined from wrong-sign candidates. Wrong-sign candidates include the $D_s^+ D_s^-$, $D^0 D_s^0$, or $D_s^0(K^+ \pi^-)D_s^+$ final states, in which no signal excesses should be present [neglecting the small contribution from the doubly Cabibbo suppressed $B^- \rightarrow D^0(K^+ \pi^-)D_s^+$ decay]. For the $B_{(s)}^0 \rightarrow D^+ D^-$ decay, the exponential shape parameter is allowed to vary in the fit due to an insufficient number of wrong-sign $D^+ D^-$ candidates.

IV. FIT RESULTS

Figure 2 shows the invariant mass spectra for $B_s^0 \rightarrow D_s^+ D_s^-$ and $B^0 \rightarrow D^- D_s^+$ candidates. The results of unbinned extended maximum likelihood fits to the distributions are overlaid, with the signal and background components indicated in the legends. Signal yields of $451 \pm 23 B_s^0 \rightarrow D_s^+ D_s^-$ and $5157 \pm 64 B^0 \rightarrow D^- D_s^+$ decays are observed.

Figure 3 shows the invariant mass spectrum for $B^0 \rightarrow D^+ D^- D_s^+$ and $B_{(s)}^0 \rightarrow D_s^+ D^- D_s^-$ candidates, where the tight BDT selection requirements have been applied as discussed previously. We observe $36 \pm 6 B_s^0 \rightarrow D_s^+ D^- D_s^-$ signal decays, with $2832 \pm 53$ events in the $B^0 \rightarrow D^- D_s^+$ normalization mode. The statistical significance of the $B_s^0 \rightarrow D_s^+ D^- D_s^-$ signal corresponds to 10σ by computing $\sqrt{-2 \ln \left( L_0 / L_{\max} \right)}$, where $L_{\max}$ and $L_0$ are the fit likelihoods with the signal yields allowed to vary and fixed to zero, respectively. Variations in the signal and background model have only a marginal impact on the signal significance. The $B_s^0 \rightarrow D^+ D^- D_s^+$ decay is thus observed for the first time.

The invariant mass spectrum for $B_{(s)}^0 \rightarrow D^+ D^-$ candidates is shown in Fig. 4 (left). Peaks are seen at both the $B^0$ and $B_s^0$ meson masses, with yields of $165 \pm 13$ and $43 \pm 7$ signal events, respectively. In the lower mass region, two prominent peaks from $B_s^0 \rightarrow D_s^+ D^- D_s^-$ decays are also evident. The significance of the $B_s^0 \rightarrow D^+ D^-$ signal yield is computed as described above, and corresponds to 11σ, establishing the first observation of this decay mode.

![Diagram](image.png)

**FIG. 2** (color online). Invariant mass distributions for (left) $B_{(s)}^0 \rightarrow D_s^+ D_s^-$ and (right) $B^0 \rightarrow D^- D_s^+$ candidates in the data with the loose BDT selection applied to the latter. The signal and background components are indicated in the legend. The $\Lambda_b^0 \rightarrow \Lambda_s^+ D_s^-$, $B_s^0 \rightarrow D_s^+ K^- K^+ \pi^-$ and $B^0 \rightarrow D^- K^+ K^- \pi^+$ background components are too small to be seen, and are excluded from the legends.
FIG. 3 (color online). Invariant mass distribution for $B^0 \to D^- D^*_s$ and $\bar{B}^0 \to D^+_s D^-$ candidates in the data, with the tight BDT selection applied. The distribution is plotted on a (left) linear and (right) logarithmic scale to highlight the suppressed $\bar{B}^0 \to D^+_s D^-$ signal. Signal and background components are indicated in the legend.

FIG. 4 (color online). Invariant mass distributions for (left) $\bar{B}^0 \to D^+ D^-$ (right) $\bar{B}^0 \to D^0 \bar{D}^0$ candidates in the data. Signal and background components are indicated in the legend.

Figure 4 (right) shows the $D^0 \bar{D}^0$ invariant mass distribution and the results of the fit. Both $(K^- \pi^+, K^+ \pi^-)$ and $(K^- \pi^+, K^+ \pi^- \pi^+ \pi^-)$ combinations are included. A $\bar{B}^0 \to D^0 \bar{D}^0$ signal is seen with a significance of $11\sigma$, which establishes the first observation of this decay mode. The data also show an excess of events at the $B^0$ mass. The significance of that excess corresponds to $2.4\sigma$, including both the statistical and systematic uncertainty. The fitted yields in the $\bar{B}^0 \to D^0 \bar{D}^0$ and $B^0 \to D^0 \bar{D}^0$ decay modes are $45 \pm 8$ and $13 \pm 6$ events, respectively. If both the $\bar{B}^0 \to D^0 \bar{D}^0$ and $B^0 \to D^0 \bar{D}^0$ decays proceed through $W$-exchange diagrams, one would expect the signal yield in $B^0 \to D^0 \bar{D}^0$ to be $\sim (f_{\text{fs}}/f_{\text{sd}}) \times |V_{cd}/V_{cs}|^2 \approx 0.2$ of the yield in $\bar{B}^0 \to D^0 \bar{D}^0$, where we have used $|V_{cd}/V_{cs}|^2 = 0.054$ [19] and the $B$ fragmentation fraction ratio $f_{\text{fs}}/f_{\text{sd}} = 0.256 \pm 0.020$ [38]. The fitted yields are consistent with this expectation. The decay $B^- \to D^- D^*_s$ is used as the normalization channel for both the $\bar{B}^0 \to D^0 \bar{D}^0$ and $B^0 \to D^0 \bar{D}^0$ branching fraction measurements, where only the $D^0 \to K^- \pi^+$ decay mode is used. The fitted invariant mass distribution for $B^- \to D^0 D^- \pi^+$ candidates is shown in Fig. 5. The fitted signal yield is $5152 \pm 73$ events.

The measured yields, $N_{B \to D\bar{D}}$, relevant for the branching fraction measurements are summarized in Table II. The branching fractions are related to the measured yields by

FIG. 5 (color online). Invariant mass distribution for $B^- \to D^0 D^- \pi^+$ candidates in the data. Signal and background components are indicated in the legend. The $B^- \to D^0 K^- K^+ \pi^-$ background components are too small to be seen, and are excluded from the legend.
The measured ratios of branching fractions are computed to be

\[ \frac{\mathcal{B}(B_s^0 \to D_s^{-} D_s^0)}{\mathcal{B}(B_s^{-} \to D_s^{-} D_s^0)} = \frac{f_d}{f_s} \cdot \epsilon'_{rel} \cdot \kappa \cdot \frac{N_{B_s^0 \to D_s^{-} D_s^0}}{N_{B_s^{-} \to D_s^{-} D_s^0}}, \]

\[ \frac{\mathcal{B}(B_s^0 \to D^+ D^0)}{\mathcal{B}(B_s^{-} \to D^+ D_s^0)} = \frac{f_d}{f_s} \cdot \epsilon'_{rel} \cdot \frac{N_{B_s^0 \to D^+ D^0}}{N_{B_s^- \to D^+ D_s^0}}. \]

Here, it is assumed that \( B^- \) and \( B^0 \) mesons are produced in equal numbers. The relative efficiencies, \( \epsilon'_{rel} \), are given in Table II. They account for geometric acceptance, detection and trigger efficiencies, and the additional VS \( \chi^2 \), BDT, and charm cross-feed veto requirements. The first four of these relative efficiencies are obtained from simulation, and the last two are determined using data-driven methods, as discussed in Sec. II. The indicated uncertainties on the relative efficiencies are due only to the finite sizes of the simulated signal decays. The average selection efficiency for \( B_s^- \to D_s^0 D_s^0 \) relative to \( B_s^0 \to D_s^0 D_s^0 \) is

\[ \epsilon'_{rel} = \frac{\epsilon_{B_s^- \to D_s^0 D_s^0} \cdot \mathcal{B}(D_s^0 \to K^+ K^- \pi^+) \mathcal{B}(D_s^0 \to K^- \pi^+)}{\epsilon_{K \pi, K \pi} \left( \mathcal{B}(D_s^0 \to K^- \pi^+) \right)^2 + 2 \epsilon_{K \pi, K \pi, K \pi} \mathcal{B}(D_s^0 \to K^- \pi^+) \mathcal{B}(D_s^0 \to K^- \pi^+)}. \]
upper limit. The upper limit is obtained by convolving the fitted likelihood with a Gaussian function whose width is the total systematic error, and integrating over the physical region.

V. SYSTEMATIC UNCERTAINTIES

A number of systematic uncertainties contribute to the measurements of the ratios of branching fractions. The sources and their values are summarized in Table III. The dominant source of uncertainty on the branching fraction ratios comes from the $b$ fragmentation fraction ratio, $f_d/f_s$, which has a total uncertainty of 7.8% [38], of which 5.3% is from the ratio of branching fractions $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)/\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$. For clarity, we have removed that portion of the uncertainty from $f_d/f_s$ and included its contribution in the row labeled $\mathcal{B}(D)$ in Table III. For $\mathcal{B}(B_s^0 \to D^+_sD^-_s)/\mathcal{B}(B^0 \to D^-D^+_s)$, the above $D^+_s/D^+$ branching fraction ratio from $f_d/f_s$ cancels with the corresponding inverted ratio in Eq. (1). On the other hand, in the ratio $\mathcal{B}(B^0 \to D^0D^0)/\mathcal{B}(B^- \to D^0D^+_s)$, the $D^+_s \to K^+K^-\pi^+$ branching fraction enters as the square, after considering the $D$ branching fractions used in computing $f_d/f_s$ [see Eq. (4)]. As a result, the uncertainty from $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$ contributes 9.8% to the total uncertainty on $\mathcal{B}(B^0 \to D^0D^0)/\mathcal{B}(B^- \to D^0D^+_s)$; smaller contributions from the limited knowledge of $\mathcal{B}(D^0 \to K^-\pi^+\pi^+) [1.3\%]$, $\mathcal{B}(D^0 \to K^-\pi^+\pi^+\pi^+) [2.5\%]$ and $\mathcal{B}(D^+ \to K^-\pi^+\pi^+) [2.1\%]$ are also included in the $\mathcal{B}(D)$ uncertainties.

Another significant uncertainty results from the precision on $b$-hadron lifetimes and decays of $B^0$ and $B^0_s$ to $CP$ eigenstates. Using the measured value of the width difference, $\Delta \Gamma_s = 0.116 \pm 0.018 \pm 0.006$ ps$^{-1}$ [40] we conservatively assume the $CP$-even lifetime to be in the range from 0.85 to 0.95 times the flavor-specific decay lifetime. With this allowed range a 2.9% uncertainty on the efficiencies for $B^0_s$ decays to $CP$ eigenstates is found. The average $B^0_s$ lifetime is known only to a precision of 3%, which leads to a 1.5% uncertainty on the selection efficiencies for $B^0_s$ decays to flavor-specific final states. The $B^0$ and $B^-$ lifetimes are known with sufficient precision that the associated uncertainty is negligible.

Several of the efficiency factors are estimated from simulation. Most, but not all, of the associated systematic uncertainties cancel due to the similar or identical final states for the signal and normalization modes. For modes with an unequal number of tracks in the final state, a 1% uncertainty due to small differences in the IP resolution between data and simulation is assigned. The efficiency of the VS $\chi^2$ requirement is checked using the large $B^0 \to D^-D^+_s$ signal in data, and the agreement to within 1% with the efficiency from simulation is the assigned uncertainty. For $\mathcal{B}(B^- \to D^0D^+_s)/\mathcal{B}(B^0 \to D^-D^+_s)$, a 1% uncertainty is attributed to the efficiency of track reconstruction. For $\mathcal{B}(B^0 \to D^0D^0)/\mathcal{B}(B^- \to D^0D^-_s)$, the one fewer track in the $D^0(K\pi)D^0(K\pi)$ final state is offset by the one extra track in $D^0(K\pi)D^0(K\pi\pi)$, relative to $D^0(K\pi)D^-_s(K\pi\pi)$, leading to a negligible tracking uncertainty. The mass resolution in data is slightly larger than in simulation, resulting in slightly different efficiencies for the reconstructed $D^0$, $D^+$ and $D^+_s$ invariant masses to lie within 25 MeV/c$^2$ of their known masses. This introduces a maximum of 1% uncertainty on the relative branching fractions. To estimate the uncertainty on the trigger efficiencies determined from simulation, the hadron trigger efficiency ratios were also determined using data. These efficiencies were measured using trigger-unbiased samples of kaons and pions identified in $D^0 \to D^0\pi^+$ decays. Using this alternative procedure, we find that the simulated trigger efficiency ratios have an uncertainty of 2%. The combined systematic uncertainties in the efficiencies obtained from simulation are given in Table III.

The limited sizes of the $B \to D\pi^-$ calibration samples lead to uncertainties in the BDT efficiencies. The uncertainties on the ratios vary from 1.0% to 2.0%. The uncertainty on the efficiency of the $D_{0/2}$ and $\Lambda_0^+$ vetoes is dominated by the PID efficiencies, but they only apply to

<table>
<thead>
<tr>
<th>Source</th>
<th>$B^0 \to D^0D^0$</th>
<th>$B^- \to D^0D^0$</th>
<th>$B^- \to D^+D^-$</th>
<th>$B^0 \to D^+D^-$</th>
<th>$B^- \to D^-D^+_s$</th>
<th>$B^- \to D^0D^-_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_d/f_s$</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
</tr>
<tr>
<td>$\mathcal{B}(D)$</td>
<td>⋯</td>
<td>5.3</td>
<td>5.3</td>
<td>5.3</td>
<td>10.2</td>
<td>10.2</td>
</tr>
<tr>
<td>$B$ meson lifetimes</td>
<td>2.9</td>
<td>1.5</td>
<td>2.9</td>
<td>2.9</td>
<td>10.2</td>
<td>10.2</td>
</tr>
<tr>
<td>Eff. from simulation</td>
<td>2.4</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>BDT selection</td>
<td>1.4</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Cross-feed vetoes</td>
<td>0.6</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>$D$ mass resolution</td>
<td>1.0</td>
<td>⋯</td>
<td>⋯</td>
<td>⋯</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Fit model</td>
<td>2.1</td>
<td>0.5</td>
<td>0.5</td>
<td>1.7</td>
<td>1.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Simulated sample size</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Total</td>
<td>8.0</td>
<td>8.5</td>
<td>8.9</td>
<td>11.7 (13.0)</td>
<td>5.5</td>
<td>5.5</td>
</tr>
</tbody>
</table>
the subset of $D$ candidates that fall within the mass window of two charm hadrons, e.g., both the $D^+$ and $D_s^+$ mesons, which occurs about 20% of the time for $D_s^+$ decays. Taking this fraction and the uncertainty in the PID efficiency into account, the veto efficiencies are estimated to have uncertainties of 1.0% for the $D^+$ veto, 0.5% for the $D_s^+$ veto, and 0.3% for the $A_s^+$ veto.

The fit model is validated using simulated experiments, and is found to be unbiased. To assess the uncertainty due to the imperfect knowledge of the various parameters used in the fit model, a number of variations are investigated. The only non-negligible uncertainties are due to the $B \rightarrow DK^- K^+ \pi^-$ background contribution, which is varied from 0% to 2%, and the cross feed from $B_s^0 \rightarrow D_s^+ D_s^-$ decays into the $B_s^0 \rightarrow D_s^+ D_s^-$ sample. The uncertainty varies from 1.7% to 2.1%. For $B(B_s^0 \rightarrow D^+ D^-) / B(B_s^0 \rightarrow D^+ D^-)$ and $B(B_s^0 \rightarrow D_s^+ D_s^-) / B(B_s^0 \rightarrow D^+ D^-)$, we assign an uncertainty of 0.5%, which accounts for potentially small differences in the signal shape for $B_s^0$ and $B_s^0$ decays (due to the $B_s^0$-$B_s^0$ mass difference). Lastly, the finite size of the samples of simulated decays contributes 3% uncertainty to all the measurements. In total, the systematic uncertainties on the branching fraction ratios range from 5.5% to 13.0%, as indicated in Table III.

VI. DISCUSSION AND SUMMARY

First observations and measurements of the relative branching fractions for the decays $B_s^0 \rightarrow D^+ D^-$, $B_s^0 \rightarrow D_s^+ D_s^-$ and $B_s^0 \rightarrow D^+ D^-$ have been presented, along with measurements of $B(B_s^0 \rightarrow D_s^+ D_s^-)$ and $B(B^+ \rightarrow D^+ D^-)$. The measured value of $B(B_s^0 \rightarrow D_s^+ D_s^-)$ is $0.55 \pm 0.06$ significantly lower than the naive expectation of unity for the case that both decays are dominated by tree amplitudes [see Fig. 1(a)], assuming small nonfactorizable effects and comparable magnitudes of the $B_{(s)} \rightarrow D_{(s)}$ form factors [41]. Unlike $B^0 \rightarrow D^+ D^-$, the $B_s^0 \rightarrow D_s^+ D_s^-$ decay receives a contribution from the $W$-exchange process [see Fig. 1(b)], suggesting that this amplitude may not be negligible. Interestingly, when comparing the $B_s^0 \rightarrow D_s^+ D_s^-$ and $B_s^0 \rightarrow D^+ D^-$ decays, which have the same set of amplitudes, one finds $|V_{cd}/V_{cs}|^2 \cdot B(B_s^0 \rightarrow D_s^+ D_s^-) / B(B^0 \rightarrow D^+ D^-) \sim 1$.

Taking the world average values for $B(B^0 \rightarrow D^- D_s^+)$, the absolute branching fractions are

$$B(B^+ \rightarrow D^+ D^-) = (8.6 \pm 0.2 \text{stat} \pm 0.4 \text{syst} \pm 1.0 \text{norm}) \times 10^{-3},$$

$$B(B_s^0 \rightarrow D_s^+ D_s^-) = (4.0 \pm 0.2 \text{stat} \pm 0.3 \text{syst} \pm 0.4 \text{norm}) \times 10^{-3}.$$  

The third uncertainty reflects the precision of the branching fraction for the normalization mode. These measurements are consistent with, and more precise than, both the current world average measurements [19] as well as the more recent measurement of $B(B_s^0 \rightarrow D_s^+ D_s^-)$ [42].

Using $B(B^0 \rightarrow D^+ D^-) = (2.11 \pm 0.31) \times 10^{-4}$ and $B(B^0 \rightarrow D^+ D^-) = (10.0 \pm 1.7) \times 10^{-3}$ [19], the following values for the branching fractions are obtained:

$$B(B_s^0 \rightarrow D^+ D^-) = (2.2 \pm 0.4 \text{stat} \pm 0.2 \text{syst} \pm 0.3 \text{norm}) \times 10^{-4},$$

$$B(B_s^0 \rightarrow D^0 \bar{D}^0) = (1.9 \pm 0.3 \text{stat} \pm 0.3 \text{syst} \pm 0.3 \text{norm}) \times 10^{-4},$$

$$B(B_s^0 \rightarrow D^0 \bar{D}^0) = (1.4 \pm 0.6 \text{stat} \pm 0.2 \text{syst} \pm 0.2 \text{norm}) \times 10^{-5}.$$  

These results are lower than, but consistent with, the perturbative-based calculations presented in Ref. [21]. The nonperturbative calculations for $B(B_s^0 \rightarrow D^+ D^-)$ give a result that is ~20 times larger than the measured value. The measured branching fractions are on the upper end (~1.5 – 2σ) of the predictions obtained by assuming that these decay amplitudes are dominated by rescattering [18]. As discussed above for the $B(B_s^0 \rightarrow D_s^+ D_s^-)$ measurement, this may also suggest that the $W$-exchange amplitude contribution is not negligible in $B \rightarrow D \bar{D}$ decays. For precise quantitative comparisons of these $B_s^0$ branching fraction measurements to theoretical predictions, one should account for the different total widths of the $CP$-even and $CP$-odd final states [13].

The Cabibbo suppressed $B_s^0 \rightarrow D_s^+ D_s^-$ decay is also observed for the first time. Its absolute branching fraction is

$$B(B_s^0 \rightarrow D_s^+ D_s^-) = (3.6 \pm 0.6 \text{stat} \pm 0.3 \text{syst} \pm 0.4 \text{norm}) \times 10^{-4}.$$  

This value is consistent with the expected suppression of $|V_{cd}/V_{cs}|^2$.

The results reported here are based on an integrated luminosity of 1.0 fb$^{-1}$. A data sample with approximately 2.5 times larger yields in these modes has already been collected in 2012, and larger samples are anticipated in the next few years. These samples give good prospects for $CP$-violation measurements, lifetime studies, and obtaining a deeper understanding of the decay mechanisms that contribute to $b$-hadron decays.

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); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF, and MPG (Germany); SFI (Ireland); INFN (Italy); FOM
and NWO (The Netherlands); SCSR (Poland); ANCS/IFA (Romania); MinES, Rosatom, RFBR, and NRC “Kurchatov Institute” (Russia); MinECo, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); NSF (USA). We also acknowledge the support received from the ERC under FP7. The Tier1 computing centres are supported by IN2P3 (France), KIT, and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom). We are thankful for the computing resources put at our disposal by Yandex LLC (Russia), as well as to the communities behind the multiple open source software packages that we depend on.

FIRST OBSERVATIONS OF...
25 Henryk Niewodniczanski Institute of Nuclear Physics, Kraków, Poland
26 AGH University of Science and Technology, Kraków, Poland
27 National Center for Nuclear Research (NCBJ), Warsaw, Poland
28 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
29 Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
30 Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
31 Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
32 Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
33 Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia
34 Institute for High Energy Physics (IHEP), Protvino, Russia
35 Universitat de Barcelona, Barcelona, Spain
36 Universidad de Santiago de Compostela, Santiago de Compostela, Spain
37 European Organization for Nuclear Research (CERN), Geneva, Switzerland
38 Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
39 Physik-Institut, Universität Zürich, Zürich, Switzerland
40 Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
41 NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
42 Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
43 University of Birmingham, Birmingham, United Kingdom
44 University of Bristol, Bristol, United Kingdom
45 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
46 Department of Physics, University of Warwick, Coventry, United Kingdom
47 Rutherford Appleton Laboratory, Science and Technology Facilities Council, Didcot, United Kingdom
48 School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
49 School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
50 Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
51 Imperial College London, London, United Kingdom
52 School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
53 School of Physics and Astronomy, University of Oxford, Oxford, United Kingdom
54 Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
55 Syracuse University, Syracuse, New York, USA
56 Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
57 Institut für Physik, Universität Rostock, Rostock, Germany (associated to Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany)
58 University of Cincinnati, Cincinnati, Ohio, USA (associated to Syracuse University, Syracuse, New York, USA)

a Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
b Also at Università di Bari, Bari, Italy.
c Also at Università di Bologna, Bologna, Italy.
d Also at Università di Cagliari, Cagliari, Italy.
e Also at Università di Ferrara, Ferrara, Italy.
f Also at Università di Firenze, Firenze, Italy.
g Also at Università di Urbino, Urbino, Italy.
h Also at Università di Modena e Reggio Emilia, Modena, Italy.
i Also at Università di Genova, Genova, Italy.
j Also at Università di Milano Bicocca, Milano, Italy.
k Also at Università di Roma Tor Vergata, Roma, Italy.
l Also at Università di Roma La Sapienza, Roma, Italy.
m Also at Università della Basilicata, Potenza, Italy.

n Also at LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
o Also at IFIC, Universitat de Valencia-CSIC, Valencia, Spain.
p Also at Hanoi University of Science, Hanoi, Vietnam.
q Also at Università di Padova, Padova, Italy.
r Also at Università di Pisa, Pisa, Italy.
s Also at Scuola Normale Superiore, Pisa, Italy.