I. INTRODUCTION

The \( B^+ \) meson, the ground state of the \( \bar{b}c \) system, is unique, being the only weakly decaying heavy quarkonium system. Its lifetime [1,2] is almost 3 times smaller than that of other beauty mesons, pointing to the important role of the charm quark in weak \( B^+ \) decays. The \( B^+_c \) meson was first observed through its semileptonic decay \( B^+_c \rightarrow J/\psi \ell^+ \nu_X \) [3]. Only three hadronic modes have been observed so far: \( B^+_c \rightarrow J/\psi \pi^+ \) [4], \( B^+_c \rightarrow J/\psi \pi^+ \pi^+ \pi^- \) [5] and \( B^+_c \rightarrow \psi(2S)\pi^+ \) [6].

The first observations of the decays \( B^+_c \rightarrow J/\psi D^+_s \) and \( B^+_c \rightarrow J/\psi D^{*+}_s \) are reported in this paper. The leading Feynman diagrams of these decays are shown in Fig. 1. The decay \( B^+_c \rightarrow J/\psi D^+_s \) is expected to proceed mainly through spectator and color-suppressed spectator diagrams. In contrast to decays of other beauty hadrons, the weak annihilation topology is not suppressed and can contribute significantly to the decay amplitude.

Assuming that the spectator decay dominates and that factorization holds, the following approximations can be established:

\[
\mathcal{R}_{D^+_s/\pi^+} = \frac{\Gamma(B^+_c \rightarrow J/\psi D^+_s)}{\Gamma(B^+_c \rightarrow J/\psi \pi^+)} = \frac{\Gamma(B \rightarrow D^* D^+_s)}{\Gamma(B \rightarrow D^* \pi^+)} ,
\]

(1a)

\[
\mathcal{R}_{D^{*+}_s/D^+_s} = \frac{\Gamma(B^+_c \rightarrow J/\psi D^{*+}_s)}{\Gamma(B^+_c \rightarrow J/\psi D^+_s)} = \frac{\Gamma(B \rightarrow D^{*0} D^{*+}_s)}{\Gamma(B \rightarrow D^{*0} D^+_s)} ,
\]

(1b)

where \( B \) stands for \( B^+ \) or \( B^0 \) and \( D^* \) denotes \( D^{*0} \) or \( D^{*+} \). Phase space corrections amount to \( O(0.5\%) \) for Eq. (1a) and can be as large as 28\% for Eq. (1b), depending on the relative orbital momentum. The relative branching ratios estimated in this way, together with more detailed theoretical calculations, are listed in Table I, where the branching fractions for the \( B \rightarrow D^* D^+_s \) and \( B \rightarrow D^* \pi^+ \) decays are taken from Ref. [1].

The analysis presented here is based on a data sample, corresponding to an integrated luminosity of 3 fb\(^{-1}\), collected with the LHCb detector during 2011 and 2012 in pp collisions at center-of-mass energies of 7 and 8 TeV, respectively. The decay \( B^+_c \rightarrow J/\psi \pi^+ \) is used as a normalization channel for the measurement of the branching fraction \( \mathcal{B}(B^+_c \rightarrow J/\psi D^+_s) \). In addition, the low energy release (Q value) in the \( B^+_c \rightarrow J/\psi D^{*+}_s \) mode allows a determination of the \( B^+_c \) mass with small systematic uncertainty.

II. LHCB DETECTOR

The LHCb detector [12] 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 resolution of 20 \( \mu \)m for tracks with high transverse momentum. Charged hadrons are identified using two ring-imaging Cherenkov 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 [13] 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.

This analysis uses events collected by triggers that select the decay products of the dimuon decay of the \( J/\psi \) meson...
with high efficiency. At the hardware stage either one or two identified muon candidates are required. In the case of single muon triggers the transverse momentum $p_T$ of the candidate is required to be larger than 1.5 GeV/c. For dimuon candidates a requirement on the product of the $p_T$ of the muon candidates is applied, $\sqrt{p_T^1 p_T^2} > 1.3$ GeV/c. At the subsequent software trigger stage, two muons with invariant mass in the interval $2.97 < m_{\mu^+\mu^-} < 3.21$ GeV/$c^2$ and consistent with originating from a common vertex are required.

The detector acceptance and response are estimated with simulated data. Proton-proton collisions are generated using PYTHIA 6.4 [14] with the configuration described in Ref. [15]. Particle decays are then simulated by EVTGEN [16] in which final state radiation is generated using PHOTOS [17]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [18] as described in Ref. [19].

### III. EVENT SELECTION

Track quality of charged particles is ensured by requiring that the $\chi^2$ per degree of freedom, $\chi^2_{dof}$/ndf, is less than 4. Further suppression of fake tracks created by the reconstruction is achieved by a neural network trained to discriminate between these and real particles based on information from track fit and hit pattern in the tracking detectors. A requirement on the output of this neural network, $P_{\text{fake}} < 0.5$, allows us to reject half of the fake tracks.

Duplicate particles created by the reconstruction are suppressed by requiring the symmetrized Kullback-Leibler divergence [20] $\Delta_{\text{KL}}^{\text{min}}$, calculated with respect to all particles in the event, to be in excess of 5000. In addition, the transverse momentum is required to be greater than $550(250)$ MeV/c for each muon (hadron) candidate.

Well-identified muons are selected by requiring that the difference in logarithms of the likelihood of the muon hypothesis, as provided by the muon system, with respect to the pion hypothesis, $\Delta^{\mu/\pi}$ in $L$ [21], is greater than zero. Good quality particle identification by the ring-imaging Cherenkov detectors is ensured by requiring the momentum of the hadron candidates, $p$, to be between 3.2 and 100 GeV/c, and the pseudorapidity to be in the range $2 < \eta < 5$. To select well-identified kaons (pions) the corresponding difference in logarithms of the likelihood of the kaon and pion hypotheses [22] is required to be $\Delta^{K/\pi}$ in $L > 2(<0)$. These criteria are chosen to be tight enough to reduce significantly the background due to misidentification while ensuring good agreement between data and simulation.

To ensure that the hadrons used in the analysis are inconsistent with being directly produced in a pp interaction vertex, the impact parameter $\chi^2$, defined as the difference between the $\chi^2$ of the reconstructed pp collision vertex formed with and without the considered track, is required to be $\chi^2_{IP} > 9$. When more than one vertex is reconstructed, that with the smallest value of $\chi^2_{IP}$ is chosen.

As in Refs. [23–25] the selection of $J/\psi \rightarrow \mu^+\mu^-$ candidates proceeds from pairs of oppositely charged muons forming a common vertex. The quality of the vertex is ensured by requiring the $\chi^2$ of the vertex fit, $\chi^2_{\text{vertex}}$, to be less than 30. The vertex is forced to be well separated from the reconstructed pp interaction vertex by requiring the decay length significance $S_{\text{flight}}$, defined as the ratio of the projected distance from pp interaction vertex to $\mu^+\mu^-$ vertex on direction of $\mu^+\mu^-$ pair momentum and its uncertainty, to be greater than 3. Finally, the mass of the dimuon combination is required to be within $\pm 45$ MeV/$c^2$ of the known $J/\psi$ mass [1], which corresponds to a $\pm 3.5\sigma$ window, where $\sigma$ is the measured $J/\psi$ mass resolution.

Candidate $D^+_s$ mesons are reconstructed in the $D^+_s \rightarrow (K^+K^-)\phi\pi^+$ mode using criteria similar to those in Ref. [26]. A good vertex quality is ensured by requiring $\chi^2_{\text{vertex}} < 25$. The mass of the kaon pair is required to be consistent with the decay $\phi \rightarrow K^+K^-$, $|m_{K^+K^-} - m_\phi| < 20$ MeV/$c^2$. Finally, the mass of the candidate is required to be within $\pm 20$ MeV/$c^2$ of the known $D^+_s$ mass [1].

### TABLE I. Predictions for the ratios of $B^+_c$ meson branching fractions. In the case of $R_{D^+_s}/D^+_s$, the second uncertainty is related to the unknown relative orbital momentum.

<table>
<thead>
<tr>
<th>$R_{D^+_s}/n^+$</th>
<th>$R_{D^+_s}/D^+_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.90 ± 0.42</td>
<td>2.20 ± 0.35 ± 0.62</td>
</tr>
<tr>
<td>1.58 ± 0.34</td>
<td>2.07 ± 0.52 ± 0.52</td>
</tr>
<tr>
<td>1.3</td>
<td>3.9</td>
</tr>
<tr>
<td>2.6</td>
<td>1.7</td>
</tr>
<tr>
<td>2.0</td>
<td>2.9</td>
</tr>
<tr>
<td>2.2</td>
<td>⋮</td>
</tr>
<tr>
<td>1.2</td>
<td>⋮</td>
</tr>
</tbody>
</table>

Eqs. (1) with $B^0$

Eqs. (1) with $B^+$

Ref. [7]

Ref. [8]

Ref. [9]

Ref. [10]

Ref. [11]
which corresponds to a ±3.5σ window, where σ is the measured \(D_s^+\) mass resolution, and its transverse momentum to be >1 GeV/c.

Candidate \(B_c^+\) mesons are formed from \(J/\psi D_s^+\) pairs with transverse momentum in excess of 1 GeV/c. The candidates should be consistent with being produced in a pp interaction vertex by requiring \(\chi^2_{\text{IP}} < 9\) with respect to reconstructed pp collision vertices. A kinematic fit is applied to the \(B_c^+\) candidates [27]. To improve the mass and lifetime resolution, in this fit, a constraint on the pointing of the candidate to the primary vertex is applied together with mass constraints on the intermediate \(J/\psi\) and \(D_s^+\) states. The value of the \(J/\psi\) mass is taken from Ref. [1].

For the \(D_s^+\) meson the value of \(m_{D_s^+} = 1968.31 \pm 0.20\) MeV/c\(^2\) is used, that is, the average of the values given in Refs. [1,28]. The \(\chi^2\) per degree of freedom of this fit, \(\chi^2_{\text{DF}}/\text{ndf}\), is required to be less than 5. The decay time of the \(D_s^+\) candidate, \(c\tau(D_s^+),\) determined by this fit, is required to satisfy \(c\tau > 75\) \(\mu\)m. The corresponding signed significance \(S_c\), defined as the ratio of the measured decay time and its uncertainty, is required to be in excess of 3. Finally, the decay time of the \(B_c^+\) candidate, \(c\tau(B_c^+),\) is required to be between 75 \(\mu\)m and 1 mm. The upper edge, in excess of seven lifetimes of \(B_c^+\) meson, is introduced to remove badly reconstructed candidates.

**IV. OBSERVATION OF \(B_c^+ \rightarrow J/\psi D_s^+\)**

The mass distribution of the selected \(B_c^+ \rightarrow J/\psi D_s^+\) candidates is shown in Fig. 2. The peak close to the known mass of the \(B_c^+\) meson [1,29] with a width compatible with the expected mass resolution is interpreted as being due to the \(B_c^+ \rightarrow J/\psi D_s^+\) decay. The wide structure between 5.9 and 6.2 GeV/c\(^2\) is attributed to the decay \(B_c^+ \rightarrow J/\psi D_s^+\), followed by \(D_s^+ \rightarrow D_s^0\) or \(D_s^+ \rightarrow D_s^0\) \(\pi^0\) decays, where the neutral particles are not detected. The process \(B_c^+ \rightarrow J/\psi D_s^+\) being the decay of a pseudoscalar particle into two vector particles is described by three helicity amplitudes: \(A_{\pm\pm}, A_{00}\), and \(A_{-\pm}\), where indices correspond to the helicities of the \(J/\psi\) and \(D_s^+\) mesons. Simulation studies show that the \(J/\psi D_s^+\) mass distributions are the same for the \(A_{\pm\pm}\) and \(A_{00}\) amplitudes. Thus, the \(J/\psi D_s^+\) mass spectrum is described by a model consisting of the following components: an exponential shape to describe the combinatorial background, a Gaussian shape to describe the \(B_c^+ \rightarrow J/\psi D_s^+\) signal and two helicity components to describe the \(B_c^+ \rightarrow J/\psi D_s^+\) contributions corresponding to the \(A_{\pm\pm}\) and \(A_{00}\) amplitudes. The shape of these components is determined using the simulation where the branching fractions for \(D_s^+ \rightarrow D_s^0\) \(\gamma\) and \(D_s^+ \rightarrow D_s^0\) \(\pi^0\) decays are taken from Ref. [1].

To estimate the signal yields, an extended unbinned maximum likelihood fit to the mass distribution is performed. The correctness of the fit procedure together with the reliability of the estimated uncertainties has been extensively checked using simulation. The fit has seven free parameters: the mass of the \(B_c^+\) meson, \(m_{B_c^+}\), the signal resolution \(\sigma_{B_c^+}\), the relative amount of the \(A_{\pm\pm}\) helicity amplitudes of total \(B_c^+ \rightarrow J/\psi D_s^+\) decay rate, \(f_{\pm\pm}\), the slope parameter of the exponential background and the yields of the two signal components, \(N_{B_c^+ \rightarrow J/\psi D_s^+}\) and \(N_{B_s^+ \rightarrow J/\psi D_s^+}\), and of the background. The values of the signal parameters obtained from the fit are summarized in Table II. The fit result is also shown in Fig. 2.

To check the result, the fit has been performed with different models for the signal: a double-sided Crystal Ball function [30,31], and a modified Novosibirsk function [32]. For these tests the tail and asymmetry parameters are fixed using the simulation values, while the parameters representing the peak position and resolution are left free to vary. As alternative models for the background, the product of an exponential function and a fourth-order polynomial function are used. The fit parameters obtained are stable with respect to the choice of the fit model and the fit range interval.

The statistical significance for the \(B_c^+ \rightarrow J/\psi D_s^+\) signal is estimated from the change in the likelihood function

**TABLE II.** Signal parameters of the unbinned extended maximum likelihood fit to the \(J/\psi D_s^+\) mass distribution.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m_{B_c^+}) [MeV/c(^2)]</td>
<td>6276.28 ± 1.44</td>
</tr>
<tr>
<td>(\sigma_{B_c^+}) [MeV/c(^2)]</td>
<td>7.0 ± 1.1</td>
</tr>
<tr>
<td>(N_{B_c^+ \rightarrow J/\psi D_s^+})</td>
<td>28.9 ± 5.6</td>
</tr>
<tr>
<td>(N_{B_s^+ \rightarrow J/\psi D_s^+})</td>
<td>2.37 ± 0.56</td>
</tr>
<tr>
<td>(f_{\pm\pm}) [%]</td>
<td>52 ± 20</td>
</tr>
</tbody>
</table>

FIG. 2 (color online). Mass distributions for selected \(J/\psi D_s^+\) pairs. The solid curve represents the result of a fit to the model described in the text. The contribution from the \(B_c^+ \rightarrow J/\psi D_s^+\) decay is shown with thin green dotted and thin yellow dash-dotted lines for the \(A_{\pm\pm}\) and \(A_{00}\) amplitudes, respectively. The inset shows a zoom of the \(B_c^+\) mass region.
\[ S_\text{c} = \sqrt{2 \ln \frac{L_B}{L_\text{B} + S}} \]

where \( L_B \) is the likelihood of a background-only hypothesis and \( L_{B+S} \) is the likelihood of a background-plus-signal hypothesis. The significance has been estimated separately for the \( B_c^+ \rightarrow J/\psi D_s^+ \) and \( B_c^+ \rightarrow J/\psi D_s^{*+} \) signals. To exclude the look-elsewhere effect \([33]\), the mass and resolution of the peak are fixed to the values obtained with the simulation. The minimal significance found varying the fit model as described above is taken as the signal significance. The statistical significance for both the \( B_c^+ \rightarrow J/\psi D_s^+ \) and \( B_c^+ \rightarrow J/\psi D_s^{*+} \) signals estimated in this way is in excess of 9 standard deviations.

The low \( Q \) value for the \( B_c^+ \rightarrow J/\psi D_s^+ \) decay mode allows the \( B_c^+ \) mass to be precisely measured. This makes use of the \( D_s^+ \) mass value, evaluated in Sec. III, taking correctly into account the correlations between the measurements. The calibration of the momentum scale for the dataset used here is detailed in Refs. \([28,34]\). It is based upon large calibration samples of \( B_c \) mesons. Adding these in quadrature gives

\[ m_{B_c^+} = 6276.28 \pm 1.44 \text{(stat)} \pm 0.36 \text{(syst)} \text{ MeV}/c^2. \]

The uncertainty on the \( D_s^+ \) meson mass and on the momentum scale largely cancels in the mass difference

\[ m_{B_c^+} - m_{D_s^+} = 4307.97 \pm 1.44 \text{(stat)} \pm 0.20 \text{(syst)} \text{ MeV}/c^2. \]

**V. Normalization to the \( B_c^+ \rightarrow J/\psi \pi^+ \) Decays Mode**

A large sample of \( B_c^+ \rightarrow J/\psi \pi^+ \) decays serves as a normalization channel to measure the ratio of branching fractions for the \( B_c^+ \rightarrow J/\psi D_s^+ \) and \( B_c^+ \rightarrow J/\psi \pi^+ \) modes. Selection of \( B_c^+ \rightarrow J/\psi \pi^+ \) events is performed in a manner similar to that described in Sec. III for the signal channel. To further reduce the combinatorial background, the transverse momentum of the pion for the \( B_c^+ \rightarrow J/\psi \pi^+ \) mode is required to be in excess of 1 GeV/c. The mass distribution of the selected \( B_c^+ \rightarrow J/\psi \pi^+ \) candidates is shown in Fig. 3.

To determine the yield, an extended unbinned maximum likelihood fit to the mass distribution is performed. The signal is modeled by a double-sided Crystal Ball function and the background with an exponential function. The fit gives a yield of 3009 ± 79 events. As cross-checks, a modified Novosibirsk function and a Gaussian function for the signal component and a product of exponential and polynomial functions for the background are used. The difference is treated as systematic uncertainty.

The ratio of the total efficiencies (including acceptance, reconstruction, selection, and trigger) for the \( B_c^+ \rightarrow J/\psi D_s^+ \) and \( B_c^+ \rightarrow J/\psi \pi^+ \) modes is determined with simulated data to be 0.148 ± 0.001, where the uncertainty is statistical only. As only events explicitly selected by the \( J/\psi \) triggers are used, the ratio of the trigger efficiencies for the \( B_c^+ \rightarrow J/\psi D_s^+ \) and \( B_c^+ \rightarrow J/\psi \pi^+ \) modes is close to unity.

**VI. Systematic Uncertainties**

Uncertainties on the ratio \( R_{D_s^+} \) related to differences between the data and simulation efficiency for the selection requirements are studied using the abundant \( B_c^+ \rightarrow J/\psi \pi^+ \) channel. As an example, Fig. 4 compares the distributions of \( \chi^2(B_c^+) \) and \( \chi^2(B_c^{*+}) \) for data and simulated \( B_c^+ \rightarrow J/\psi \pi^+ \) events. For background subtraction the sPlot technique \([36]\) has been used. It can be seen that the agreement between data and simulation is good. In addition, a large sample of selected \( B^+ \rightarrow J/\psi (K^+ K^-) \bar{K} K^+ \) events has been used to quantify differences between data and simulation. Based on the deviation, a systematic uncertainty of 1% is assigned.

The agreement of the absolute trigger efficiency between data and simulation has been validated to a precision of 4% using the technique described in Refs. \([13,31,37]\) with a large sample of \( B^+ \rightarrow J/\psi (K^+ K^-) \bar{K} K^+ \) events. A further cancellation of uncertainties occurs in the ratio of branching fractions resulting in a systematic uncertainty of 1.1%.

The systematic uncertainties related to the fit model, in particular to the signal shape, mass and resolution for the
The yields of $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$ and $B^+ \rightarrow J/\psi K^+$ decays for different data taking periods and dipole magnet polarities. This results in a further 2.5% uncertainty.

The largest systematic uncertainty is due to the knowledge of the branching fraction of the $D_s^+ \rightarrow (K^- K^+)_{\phi} \pi^+$ decay, with a kaon pair mass within $\pm 20$ MeV/$c^2$ of the known $\phi$ meson mass. The value of $(2.24 \pm 0.11 \pm 0.06)$% from Ref. [39] is used in the analysis. The systematic uncertainties on $\mathcal{R}_{D_s^+/D_s^0}$ are summarized in Table III.

The ratio $\mathcal{R}_{D_s^+/D_s^0}$ is estimated as

$$\mathcal{R}_{D_s^+/D_s^0} = \frac{N_{B_s^- \rightarrow J/\psi D_s^+}}{N_{B_s^- \rightarrow J/\psi D_s^0}},$$

where the ratio of yields is given in Table II. The uncertainty associated with the assumption that the efficiencies for the $B_s^- \rightarrow J/\psi D_s^+$ and $B_s^- \rightarrow J/\psi D_s^0$ modes are equal is evaluated by studying the dependence of the relative yields for these modes for loose (or no) requirements on the $\chi^2_{fit}(B_s^\pm)$, $\chi^2_{IP}(B_s^\pm)$ and $c\tau(B_s^\pm)$ variables. For this selection the measured ratio of $B_s^- \rightarrow J/\psi D_s^0$ to $B_s^- \rightarrow J/\psi D_s^+$ is summarized in Table III.

### Table III. Relative systematic uncertainties for the ratio of branching fractions of $B_s^- \rightarrow J/\psi D_s^+$ and $B_s^- \rightarrow J/\psi D_s^0$.

<table>
<thead>
<tr>
<th>Source</th>
<th>Uncertainty [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated efficiencies</td>
<td>1.0</td>
</tr>
<tr>
<td>Trigger</td>
<td>1.1</td>
</tr>
<tr>
<td>Fit model</td>
<td>1.8</td>
</tr>
<tr>
<td>Track reconstruction</td>
<td>2.0</td>
</tr>
<tr>
<td>Hadron interactions</td>
<td>2.0</td>
</tr>
<tr>
<td>Track quality selection</td>
<td>1.4</td>
</tr>
<tr>
<td>Kaon identification</td>
<td>3.0</td>
</tr>
<tr>
<td>$B_s^+$ lifetime</td>
<td>1.0</td>
</tr>
<tr>
<td>Stability for various data taking conditions</td>
<td>2.5</td>
</tr>
<tr>
<td>$\mathcal{B}(D_s^+ \rightarrow (K^- K^+)_{\phi} \pi^+)$</td>
<td>5.6</td>
</tr>
<tr>
<td>Total</td>
<td>8.4</td>
</tr>
</tbody>
</table>
$B_c^+ \rightarrow J/\psi D_s^+$ events changes to 2.27 ± 0.59. An uncertainty of 4% is assigned to the $R_{D_s^+}/D^+$ ratio.

The uncertainty on the fraction of the $A_{zz}$ amplitude, $f_{zz}$, has been studied with different fit models for the parameterization of the combinatorial background, as well as different mass resolution models. This is negligible in comparison to the statistical uncertainty.

VII. RESULTS AND SUMMARY

The decays $B_c^+ \rightarrow J/\psi D_s^+$ and $B_c^+ \rightarrow J/\psi D_s^+$ have been observed for the first time with statistical significances in excess of 9 standard deviations. The ratio of branching fractions for $B_c^+ \rightarrow J/\psi D_s^+$ and $B_c^+ \rightarrow J/\psi \pi^+$ is calculated as

$$\frac{\mathcal{B}(B_c^+ \rightarrow J/\psi D_s^+)}{\mathcal{B}(B_c^+ \rightarrow J/\psi \pi^+)} = \frac{1}{\mathcal{B}_{D_s^+}} \times \frac{e_{B_c^+ \rightarrow J/\psi \pi^+}^{\text{tot}}}{e_{B_c^+ \rightarrow J/\psi D_s^+}^{\text{tot}}} \times \frac{N(B_c^+ \rightarrow J/\psi D_s^+)}{N(B_c^+ \rightarrow J/\psi \pi^+)}, \tag{3}$$

where the value of $\mathcal{B}_{D_s^+} = \mathcal{B}(D_s^+ \rightarrow (K^- K^+)\phi \pi^+)$ [39] with the mass of the kaon pair within ±20 MeV/c^2 of the known value of the $\phi$ mass is used, together with the ratio of efficiencies, and the signal yields given in Secs. IV and V. This results in

$$\frac{\mathcal{B}(B_c^+ \rightarrow J/\psi D_s^+)}{\mathcal{B}(B_c^+ \rightarrow J/\psi \pi^+)} = 2.90 \pm 0.57(\text{stat}) \pm 0.24(\text{syst}).$$

The value obtained is in agreement with the naive expectations given in Eq. (1a) from $B^0$ decays, and the values from Refs. [8–10] but larger than predictions from Refs. [7,11] and factorization expectations from $B^+$ decays.

The ratio of branching fractions for the $B_c^+ \rightarrow J/\psi D_s^+$ and $B_c^+ \rightarrow J/\psi D_s^+$ decays is measured to be

$$\frac{\mathcal{B}(B_c^+ \rightarrow J/\psi D_s^+)}{\mathcal{B}(B_c^+ \rightarrow J/\psi D_s^+)} = 2.37 \pm 0.56(\text{stat}) \pm 0.10(\text{syst}).$$

This result is in agreement with the naive factorization hypothesis [Eq. (1b)] and with the predictions of Refs. [8,9].

ACKNOWLEDGMENTS

We thank A. Luchinsky and A. K. Likhoded for advice on aspects of $B_c^+$ physics. 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 (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 (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.

Observation of $B^+_s \rightarrow J/\psi D_{s}^{+}$...
38Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
39Physik-Institut, Universität Zürich, Zürich, Switzerland
40Nikhef National Institute for Subatomic Physics, Amsterdam, Netherlands
41Nikhef National Institute for Subatomic Physics (NSC KIPT), Kharkiv, Ukraine
42Department of Physics, University of Warwick, Coventry, United Kingdom
43Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
44University of Birmingham, Birmingham, United Kingdom
45H. H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
46STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
47School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
48School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
49School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
50Department of Physics, University of Oxford, Oxford, United Kingdom
51Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
52Imperial College London, London, United Kingdom
53Department of Physics, University of Oxford, Oxford, United Kingdom
54Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
55University of Cincinnati, Cincinnati, Ohio, USA
56University of Cagliari, Cagliari, Italy
57Syracuse University, Syracuse, New York, USA
58Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
[associated with Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil]
59Institut für Physik, Universität Rostock, Rostock, Germany [associated with Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany]

aAlso at P. N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
bAlso atUniversità di Bari, Bari, Italy.
cAlso atUniversità di Bologna, Bologna, Italy.
dAlso atUniversità di Cagliari, Cagliari, Italy.
eAlso atUniversità di Ferrara, Ferrara, Italy.
fAlso atUniversità di Firenze, Firenze, Italy.
gAlso atUniversità di Urbino, Urbino, Italy.
hAlso atUniversità di Modena e Reggio Emilia, Modena, Italy.
iAlso atUniversità di Genova, Genova, Italy.
jAlso atUniversità di Milano Bicocca, Milano, Italy.
kAlso atUniversità di Roma Tor Vergata, Roma, Italy.
lAlso atUniversità di Roma La Sapienza, Roma, Italy.
mAlso atUniversità della Basilicata, Potenza, Italy.
nAlso atLIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
oAlso atIFIC, Universitat de Valencia-CSIC, Valencia, Spain.
pAlso atHanoi University of Science, Hanoi, Vietnam.
qAlso atUniversità di Padova, Padova, Italy.
rAlso atUniversità di Pisa, Pisa, Italy.
sAlso atScuola Normale Superiore, Pisa, Italy.