Original citation:

Permanent WRAP url:
http://wrap.warwick.ac.uk/56742

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 3.0 (CC BY 3.0) license and may be reused according to the conditions of the license. For more details see: http://creativecommons.org/licenses/by/3.0/

A note on versions:
The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: publications@warwick.ac.uk
Observation of the Decay $B_c^+ \rightarrow B_s^0 \pi^+$

R. Aaij et al.*
(LHCb Collaboration)

(Received 22 August 2013; published 1 November 2013)

The result of a search for the decay $B_c^+ \rightarrow B_s^0 \pi^+$ is presented, using the $B_c^+ \rightarrow D^+_s \pi^+$ and $B_c^+ \rightarrow J/\psi \phi$ channels. The analysis is based on a data sample of $pp$ collisions collected with the LHCb detector, corresponding to an integrated luminosity of 1 fb$^{-1}$ taken at a center-of-mass energy of 7 TeV, and 2 fb$^{-1}$ taken at 8 TeV. The decay $B_c^+ \rightarrow B_s^0 \pi^+$ is observed with significance in excess of 5 standard deviations independently in both decay channels. The measured product of the ratio of cross sections and branching fraction is $[\sigma(B_c^+)/\sigma(B_s^0)] \times \mathcal{B}(B_c^+ \rightarrow B_s^0 \pi^+) = [2.37 \pm 0.31 (\text{stat}) \pm 0.11 (\text{syst})^{+0.17}_{-0.13}(\tau_{B_c^+})] \times 10^{-3}$, in the pseudorapidity range $2 < \eta(B) < 5$, where the first uncertainty is statistical, the second is systematic, and the third is due to the uncertainty on the $B_c^+$ lifetime. This is the first observation of a $B$ meson decaying to another $B$ meson via the weak interaction.

The $B_c^+$ meson is the ground state of the $b\bar{c}$ system. As such it is unique as it is the only weakly decaying doubly heavy meson. All measurements of $B_c^+$ meson decays to date are decays where the constituent $b$ quark decays weakly to a $c$ quark [1–8]. The decay of the $B_c^+$ meson to another $B$ meson, with the bottom quark acting as a spectator (see Fig. 1), has not previously been observed. This will improve the understanding of theoretical predictions and provide valuable information for the source of $B_c^0$ mesons at the LHC.

A wide range of predictions for the branching fraction $\mathcal{B}(B_c^+ \rightarrow B_s^0 \pi^+)$ exists, between 16.4% and 2.5%, based on, e.g., QCD sum rules [9,10], or quark-potential models (see Refs. [11–16] and references therein). Experimental clarification is needed to shed light on the present theoretical status. Unlike most other $B$ decays, the higher order corrections in the expansion of Heavy Quark Effective Theory within the framework of quantum chromodynamics (QCD) are relatively large. The expansion is described in powers of $m_c/m_b$ rather than $\Lambda_{\text{QCD}}/m_b$, due to the presence of two heavy quark constituents, where $\Lambda_{\text{QCD}}$ is the QCD scale, and $m_c$ ($m_b$) the charm (bottom) quark mass. In addition, the energy release in the decay is relatively small, leading to larger nonfactorizable effects compared to decays with lighter daughter particles. Study of the decay $B_c^+ \rightarrow B_s^0 \pi^+$ allows these models to be tested. Knowledge of the production of $B_s^0$ mesons from $B_c^+$ decays is also useful for time-dependent analyses of $B_c^+$ decays, to understand any associated decay-time bias due to the incorrect estimate of the $B_c^+$ decay time if originating from a $B_c^+$ decay, or to take advantage of flavor tagging capabilities using the accompanying (“bachelor”) pion.

The data used in this analysis were collected with the LHCb detector [17] from $pp$ collisions at $\sqrt{s} = 7$ TeV and 8 TeV, corresponding to integrated luminosities of 1 fb$^{-1}$ and 2 fb$^{-1}$, respectively.

The decays $B_c^+ \rightarrow D^- \pi^+$ and $B_c^+ \rightarrow J/\psi \phi$ are used, with the subsequent decays $D^- \rightarrow K^+K^-\pi^-$, $J/\psi \rightarrow \mu^+\mu^-$ and $\phi \rightarrow K^+K^-$. The inclusion of charge conjugate modes is implied throughout. The event selection and fits to the $B_c^+$ invariant mass distributions follow previous LHCb analyses based on these $B_c^+$ decay modes [18,19]. The two channels are analyzed independently and the final results are combined. The strategy is to normalize the final number of $B_c^+ \rightarrow B_s^0 \pi^+$ decays to the number of $B_c^+$ decays, which gives a result for the $B_c^+ \rightarrow B_s^0 \pi^+$ branching fraction multiplied by the ratio of $B_s^0$ and $B_c^+$ production rates, $[\sigma(B_c^+)/\sigma(B_s^0)] \times \mathcal{B}(B_c^+ \rightarrow B_s^0 \pi^+)$. The $B_c^+$ signal

*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.

DOI: 10.1103/PhysRevLett.111.181801 PACS numbers: 13.25.Hw, 12.15.Ji

FIG. 1. Leading-order Feynman diagram of the decay $B_c^+ \rightarrow B_s^0 \pi^+$. 

0031-9007/13/111(18)/181801(9) 181801-1 © 2013 CERN, for the LHCb Collaboration
The $B_s^0$ candidates are selected using the multivariate analysis known as the boosted decision tree (BDT) [28,29], to optimally discriminate between signal and background. In the training, simulated $B_s^0$ decays are used as the signal, whereas candidates in the $B_s^0$ mass sideband in data are used as the background. To avoid potential biases, only one sixth of the data is used in the training. It is verified that the distribution of the BDT discriminant is the same for the events used in the training, compared to those that were not. All events are used for the final result. The BDT training for the selection of $B_s^0 \rightarrow D_s^- \pi^+$ candidates uses only the upper sideband [5466, 5800] MeV/$c^2$, as the lower sideband contains a large amount of irreducible partially reconstructed $B$ decays, while the training for $B_s^0 \rightarrow J/\psi \phi$ uses both lower sideband [5200, 5316] MeV/$c^2$ and upper mass sideband [5416, 5550] MeV/$c^2$. The $B_s^0$ vertex quality ($\chi^{2}_{\text{vis}}$), flight distance, momentum $p$, and $p_T$ are used to discriminate the signal from the background. For the $D_s^- \pi^+$ final state we use, in addition, the $\chi^2_{\text{IP}}$ of the $J/\psi$ candidate and the $p_T$ of the $D_s^-$ candidate and the $p_T$ and $\chi^2_{\text{IP}}$ of the bachelor pion from the $B_s^0$ decay to suppress combinatorial background. The quantity $\chi^2_{\text{IP}}$ is defined as the difference in $\chi^2$ of a given primary vertex (PV) reconstructed with and without the considered track. The training for $J/\psi \phi$ candidates uses $p$, $p_T$, $\chi^2_{\text{vis}}$, and $\chi^2_{\text{IP}}$ of the $J/\psi$ and $\phi$ candidates, and the $p_T$ of the final state kaons and muons. In the selection of $B_s^0$ candidates from $B_c^-$ decays, variables that require the candidate to point to a primary vertex, such as the impact parameter of the $B_s^0$ candidate, are explicitly not included. The minimum value of the BDT discriminant is chosen by optimizing the $B_s^0$ signal significance $S/\sqrt{S+B}$, where $S$ and $B$ are the expected numbers of signal and combinatorial background events, respectively.

The total number of $B_s^0$ decays is obtained from extended unbinned maximum likelihood fits to the invariant mass distributions, using mass constraints for the $J/\psi$ candidates [30], and are shown in Fig. 2. The signal shapes are taken as double Crystal Ball functions [31] with a common peak value and with tails to either side of the

region was not examined until the event selection was finalized. Since the ratio of production rates, $\sigma(B_s^0)/\sigma(B^0)$, may depend on the kinematics of the produced $B$ meson, the result is quoted for 8 mesons produced in the pseudorapidity range $2 < \eta(B) < 5$, corresponding to the LHCb detector acceptance.

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, described in detail in Ref. [17]. The combined tracking system provides momentum measurement with relative uncertainty 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, $p_T$. The impact parameter (IP) is defined as the distance of closest approach between the track and a primary interaction. Charged hadrons are identified using two ring-imaging Cherenkov detectors. The charged pions from $B^+_s$ decays are selected with an efficiency of 93% while keeping the misidentification rate of kaons below 7%. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers with a typical efficiency of 97% at 1%–3% pion to muon misidentification probability. The trigger [20] 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. The $B_s^0$ candidates with muons in the final state are required to pass the hardware trigger, which selects muons with a transverse momentum, $p_T > 1.48$ GeV/$c$, whereas the $B_s^0$ candidates with only hadrons in the final state are selected by requiring a hadron in the calorimeter with $E_T > 3.6$ GeV/$c$.

Monte Carlo simulations, used to develop the $B_s^0$ candidate selection, are performed using BCVEGPy [21], interfaced with PYTHIA 6.4 [22] using a specific LHCb configuration [23]. Decays of hadronic particles are described by EVTGEN [24], in which final state radiation is generated using PHOTOS [25]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [26] as described in Ref. [27].

![Graphs showing invariant mass distributions](181801-2)

FIG. 2 (color online). Invariant mass distributions of (a) $B_s^0 \rightarrow D_s^- \pi^+$ and (b) $B_s^0 \rightarrow J/\psi \phi$ candidates. The different components are defined in the legend.
peak, to account for final state radiation and detector resolution effects. The parameters that describe the tails are obtained from simulation and are fixed in the fits. The peak and width parameters of the signal are allowed to vary. The combinatorial backgrounds are modeled with exponential distributions. The $B^0_d \to D^- \pi^+$ final state is contaminated by partially reconstructed $B$ decays such as $B^0 \to D^- \pi^+$ and $B^0 \to D^+ \rho^+$ decays, where the soft photon or neutral pion is not reconstructed, and by decays where one of the final state particles is misidentified as a kaon, such as $B^0 \to D^- \pi^+$ or $\Lambda^0 \to \Lambda^- \pi^+$ decays. The shapes of these backgrounds are fixed from simulation, following Ref. [18]. In total $103\,760 \pm 380 \ B^0_d \to J/\psi \phi$ and $73\,700 \pm 500 \ B^0_s \to D^- \pi^+$ decays are found.

Selected $B^0_d$ candidates with masses consistent with the known $B^0_d$ mass are combined with tracks that satisfy loose pion identification requirements. Subsequently, $B^+_c$ candidates are selected with a second BDT algorithm. In the training of the second BDT, simulated candidates with masses consistent with the $B^+_c$ mass [32] are used as the signal, and candidates in the $B^+_c$ mass sideband region in data are used as the background. For this, only the upper mass sideband is used in the case of $B^0 \to D^- \pi^+$, while also the lower mass sideband is used in the case of $B^0 \to J/\psi \phi$, to further suppress the larger combinatorial background at smaller values of the mass. Only one sixth of the total data set is used in the training. The second BDT uses the following variables: the $B^+_c$ candidate $p_T$, decay time, $X_{\text{ip}}$, $X_{\text{ip}}^2$, and the $B^+_c$ pointing angle, i.e., the angle between the $B^+_c$ candidate momentum vector and the line joining the associated PV and the $B^+_c$ decay vertex. The $B^0_d$ polar angle (the angle between $B^0_d$ flight direction and the beam axis), decay time, decay length, and pointing angle are also used. The $p$ and $p_T$ of the bachelor pion from the $B^+_c$ decay are the most discriminating observables in the second BDT. Differences between the analyses of the $D^- \pi^+$ and $J/\psi \phi$ final states are the use of $X_{\text{ip}}$ of the $B^0_d$ candidate and bachelor pion (from the $B^+_c$ decay), and $B^0_d$ and $B^+_c$ momentum for the former, and the use of the $B^+_c$ and $B^0_d$ decay-length uncertainties for the latter. The optimal selections are defined by maximizing figures of merit for a target level of significance of 3 standard deviations, $\varepsilon/(3/2 + \sqrt{B})$ [33], where $\varepsilon$ is the signal efficiency for a given BDT criterion. The figure of merit displays a plateau, and the chosen value is at the lower end to allow us to better constrain the shape of the combinatorial background. The chosen selection is very close to the optimal point for a target level of $5\sigma$ and for the expected significance $S/\sqrt{S + B}$. The trigger for $B^0 \to D^- \pi^+$ decays preferentially selects candidates with high $p_T$ with respect to the trigger for $B^0 \to J/\psi \phi$ decays, which results in higher efficiency for the second BDT requirement for the $B^0 \to D^- \pi^+$ final state. The $B^+_c$ and $B^0_d$ candidates are required to be produced in the pseudorapidity range $2 < \eta(B) < 5$.

The invariant mass distributions for the $B^+_c \to B^0_d \pi^+$ candidates are shown in Fig. 3, together with the resulting fits. The decay $B^+_c \to B^0_d \pi^+$ has a Q value of 770 MeV/c$^2$ (with $Q = m_{B^+_c} - m_{B^0_d} - m_{\pi^+}$), which results in a resolution of about 6 MeV/c$^2$ when a $B^0_d$ mass constraint is applied. The signal shape is modeled as a double Crystal Ball function, with its parameters obtained from simulated events. The larger number of $B^+_c$ candidates in the $B^0 \to D^- \pi^+$ channel allows variation of the peak position and the width in the fit. The combinatorial background is primarily due to signal $B^0_d$ decays combined with a random pion from the primary vertex, and is modeled with an exponential function. Backgrounds due to $B^+_c \to B^+_c \pi^+$ and $B^+_c \to B^0_d \rho^+$ decays, where the photon or neutral pion are not reconstructed, are simulated, and their shapes are modeled with Gaussian distributions, with parameters fixed in the fit, and yields allowed to vary. Statistical signal significances of 7.7$\sigma$ for $B^+_c \to B^0_s (\to D^- \pi^+) \pi^+$ and 6.1$\sigma$ for $B^+_c \to B^0_s (\to J/\psi \phi) \pi^+$ decays are obtained from the likelihood ratio of fits with and without the probability density function for the signal shape, $\sqrt{2\ln(L_s/L_{s+B})}$, with 64$\pm$10 and 35$\pm$8 signal decays, respectively.
In Fig. 3(a), the structure around 6225 MeV/c² is consistent with originating from \( B_c^+ \to B_d^- \pi^+ \) decays. However, this contribution is not significant.

To obtain the value for the \( B_c^+ \to B_d^- \pi^+ \) branching fraction, multiplied by the ratio of \( B_c^+ \) and \( B_0^0 \) production rates, the relative detection efficiency of \( B_0^0 \) decays compared to \( B_c^+ \to B_d^- \pi^+ \) decays is determined from simulation. Requiring the bachelor pion to be inside the LHCb acceptance reduces the \( B_c^+ \to B_d^- \pi^+ \) yield by about 19% with respect to the \( B_0^0 \) yield. The most significant reduction in the number of selected \( B_c^+ \) candidates comes from suppressing \( B_0^0 \) combinations with a random pion from the primary interaction, by means of the second BDT selection. The total relative detection efficiency of \( B_c^+ \to B_d^- \pi^+ \) decays with respect to \( B_0^0 \) decays is estimated to be 15.2% for the \( B_0^0 \to J/\psi \phi \) decay and 33.9% for the \( B_c^+ \to D^- \pi^+ \) final state. This difference in \( B_c^+ \) selection efficiencies is a consequence of the difference in \( B_0^0 \) trigger and selection requirements.

The sources of systematic uncertainty for the efficiency-corrected ratio of \( B_c^+ \) and \( B_0^0 \) yields are listed in Table I. The uncertainty on the \( B_0^0 \) yield in the \( D^- \pi^+ \) analysis is determined by varying the parameters that describe the tails of the signal mass distribution, and by reducing the exponent of the combinatorial background by a factor of 2. The uncertainty on the \( B_c^+ \to J/\psi \phi \) yield is obtained by comparing the fitted yield in simulated pseudoexperiments to the yield that was used as input to those experiments.

The uncertainty on the \( B_c^+ \) yield is quantified by varying the peak position and width in the fit to \( B_c^+ \to B_0^0 \to J/\psi \phi \) candidates. The signal model is validated using simulated pseudoexperiments in the \( J/\psi \phi \) analysis, whereas the tail parameters are varied by ±10% in the \( D^- \pi^+ \) analysis. In addition, the combinatorial background shape is changed to a straight line, and the difference in the signal yield is taken as the associated systematic uncertainty. The effect of partially reconstructed \( B_c^+ \to B_0^0 \rho^+ \) decays is estimated by excluding candidates with mass less than 6150 MeV/c² from the fit. The significance of the \( B_c^+ \to B_0^0 \pi^+ \) signal is reduced to 7.5σ for \( B_c^+ \to B_0^0 \to J/\psi \phi \pi^+ \) and 5.5σ for \( B_c^+ \to B_0^0 \to J/\psi \phi \pi^+ \) when the systematic uncertainties on the fit to the \( B_c^+ \) mass distribution are taken into account.

The relative detection efficiency of \( B_c^+ \) and \( B_0^0 \) events is determined from simulated events. The correspondence between data and simulation is quantified by varying the criterion on the BDT value, and by comparing the observed \( B_0^0 \) yield to the expected yield based on the change in efficiency as determined from simulation. The largest contribution is due to the 10% uncertainty on the \( B_c^+ \) lifetime [32], which was recently improved by the CDF Collaboration [34]. The change in selection efficiency when varying the \( B_c^+ \) lifetime by ±10% is assigned as the systematic uncertainty. A longer (shorter) \( B_c^+ \) lifetime corresponds to a larger (smaller) efficiency and therefore a smaller (larger) ratio. As a cross-check, the effect of the choice of different sets of BDT input variables is investigated and the result is found to be stable.

The contribution from Cabibbo suppressed \( B_c^+ \to B_0^0 K^+ \) decays, the uncertainty on the efficiency of reconstructing the extra pion, and the uncertainty on the efficiency of the particle identification requirement on the bachelor pion all give small contributions (< 1.0%) to the total systematic uncertainty, and are not itemized in the summary in Table I.

The \( B_0^0 \) and \( B_c^+ \) yields are corrected for the relative detection efficiencies, to obtain the efficiency-corrected ratios of \( B_c^+ \to B_0^0 \pi^+ \) over \( B_0^0 \) yields, \([2.54 \pm 0.40(\text{stat})^{+0.23}_{-0.17}(\text{syst})] \times 10^{-3}\) and \([2.20 \pm 0.49(\text{stat}) \pm 0.23(\text{syst})] \times 10^{-3}\) for the \( D^- \pi^+ \) and \( J/\psi \phi \) final states, respectively. The small fraction of \( B_0^0 \) candidates originating from \( B_c^+ \) decays is neglected. The uncertainty due to the uncertainty on the \( B_c^+ \) lifetime is correlated between the two measurements, and is accounted for in the combined result of the ratio of production rates multiplied with the branching fraction

\[
\frac{\sigma(B_c^+)}{\sigma(B_0^0)} \times B(B_c^+ \to B_d^- \pi^+)
\]

\[
= [2.37 \pm 0.31(\text{stat}) \pm 0.11(\text{syst})^{+0.17}_{-0.13}(\tau_{B_c^+})] \times 10^{-3},
\]

where the first uncertainty is statistical, the second is systematic, and the third is due to the uncertainty on the \( B_c^+ \) lifetime. Since \( \sigma(B_c^+)/\sigma(B_0^0) \) may depend on the kinematics of the produced \( B \) meson, the data are divided according to center-of-mass energy leading to \([1.27 \pm 0.42(\text{stat}) \pm 0.05(\text{syst})^{+0.07}_{-0.06}(\tau_{B_c^+})] \times 10^{-3}\) and \([2.92 \pm 0.40(\text{stat}) \pm 0.12(\text{syst})^{+0.21}_{-0.16}(\tau_{B_c^+})] \times 10^{-3}\) for \( \sqrt{s} = 7 \) and 8 TeV \( pp \) collisions, respectively. The lower value for the

---

TABLE I. Contributions of the various sources of (relative) systematic uncertainty on the efficiency-corrected ratio of event yields. The total systematic uncertainty is the quadratic sum of the individual contributions. The number of \( B_c^+ \to B_0^0 \to J/\psi \phi \) candidates is large enough that the peak position and width are freely varied in the fit, and hence the corresponding uncertainty is contained in the statistical uncertainty of the signal yield.

<table>
<thead>
<tr>
<th>Source</th>
<th>( D^- \pi^+ ) (%)</th>
<th>( J/\psi \phi ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_0^0 ) fit model</td>
<td>3.0</td>
<td>1.2</td>
</tr>
<tr>
<td>( B_c^+ ) mean mass</td>
<td>...</td>
<td>2.0</td>
</tr>
<tr>
<td>( B_c^+ ) mass resolution</td>
<td>...</td>
<td>5.2</td>
</tr>
<tr>
<td>( B_c^+ ) signal model</td>
<td>1.5</td>
<td>1.7</td>
</tr>
<tr>
<td>Combinatorial background model</td>
<td>1.8</td>
<td>0.3</td>
</tr>
<tr>
<td>Partially reconstructed background</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Data-simulation difference</td>
<td>3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>( B_c^+ ) lifetime</td>
<td>+6.8</td>
<td>-3.5</td>
</tr>
<tr>
<td></td>
<td>+8.9</td>
<td>7.4</td>
</tr>
<tr>
<td>Total</td>
<td>-6.7</td>
<td>10.4</td>
</tr>
</tbody>
</table>
result of the 7 TeV data is attributed to a downward statistical fluctuation of the $B_s^+ \rightarrow B_0^0(\rightarrow J/\psi \phi)\pi^+$ yield in the 2011 data set, with a p value of 1.5%.

Assuming a value for $\mathcal{B}(B_s^+ \rightarrow J/\psi \pi^+)$ around 0.15% [11], combined with the results $[\sigma(B_s^+)/\sigma(B^+)] \times \mathcal{B}(B_s^+ \rightarrow J/\psi \pi^+)/\mathcal{B}(B^+ \rightarrow J/\psi K^+) = (0.68 \pm 0.10 \pm 0.03 \pm 0.05)\%$ [4], and measurements of $f_+/f_0$ [18] and $\mathcal{B}(B^+ \rightarrow J/\psi K^+)$ [32], results in a ratio of production rates of $B_s^+$ mesons over $B^0$ mesons of about 0.02. This leads to a branching fraction for $B_s^+ \rightarrow B_0^0\pi^+$ of about 10%. Although precise quantification requires improved understanding of $\sigma(B_s^+)$ and $\mathcal{B}(B_s^+ \rightarrow J/\psi \pi^+)$, even taking the lower estimates for $\mathcal{B}(B_s^+ \rightarrow J/\psi \pi^+)$ that are found in the literature [11], leads to a value of $\mathcal{B}(B_s^+ \rightarrow B_0^0\pi^+)$ which is the largest exclusive branching fraction of any known weak B meson decay.

In summary, the first observation of a weak decay of a $B$ meson to another $B$ meson is reported. This measurement will help to better understand flavor tagging and the decay time resolution in time-dependent $B_0^0$ analyses, and in addition will constrain models that predict branching fractions of $B_s^+$ decays.

We wish to thank A. K. Likhoded for useful discussions. 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 (Brazil); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); SCRF (Poland); MEN/IFA (Romania); MinES, Rosatom, RFBR and NRC “Kurchatov Institute” (Russia); MinECo, XuntaGal and GENCIAT (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 centers 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.

R. Aaij,40 B. Adeva,36 M. Adinolfi,45 C. Adrover,6 A. Affolder,51 Z. Ajaltouni,5 J. Albrecht,9 F. Alessio,37 M. Alexander,50 S. Ali,40 G. Alkhazov,29 P. Alvarez Cartelle,36 A. A. Alves, Jr.,24,37 S. Amato,2 S. Amerio,21

(LHCb Collaboration)

1Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil
2Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
3Center for High Energy Physics, Tsinghua University, Beijing, China
4LAPP, Université de Savoie, CNRS/IN2P3, Annecy-Le-Vieux, France
5Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, Clermont-Ferrand, France
Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.

Also at Università di Bari, Bari, Italy.

Also at Università di Bologna, Bologna, Italy.

Also at Università di Cagliari, Cagliari, Italy.

Also at Università di Ferrara, Ferrara, Italy.

Also at Università di Firenze, Firenze, Italy.

Also at Università di Urbino, Urbino, Italy.

Also at Università di Modena e Reggio Emilia, Modena, Italy.

Also at Università di Genova, Genova, Italy.

Also at Università di Milano Bicocca, Milano, Italy.

Also at Università di Roma Tor Vergata, Roma, Italy.

Also at Università di Roma La Sapienza, Roma, Italy.

Also at Università della Basilicata, Potenza, Italy.

Also at LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.

Also at Hanoi University of Science, Hanoi, Vietnam.

Also at Institute of Physics and Technology, Moscow, Russia.

Also at Università di Padova, Padova, Italy.

Also at Università di Pisa, Pisa, Italy.

Also at Scuola Normale Superiore, Pisa, Italy.