Measurement of the $B_s^0 \to D_s^{(*)}+D_s^{(*)}\!\!\!$ branching fractions

The LHCb collaboration†

Abstract

The branching fraction of the decay $B_s^0 \to D_s^{(*)}+D_s^{(*)}\!\!\!$ is measured using $pp$ collision data corresponding to an integrated luminosity of $1.0\,\text{fb}^{-1}$, collected using the LHCb detector at a centre-of-mass energy of $7\,\text{TeV}$. It is found to be

$$B(B_s^0 \to D_s^{(*)}+D_s^{(*)}\!\!\! ) = (3.05 \pm 0.10 \pm 0.20 \pm 0.34)\%,$$

where the uncertainties are statistical, systematic, and due to the normalisation channel, respectively. The branching fractions of the individual decays corresponding to the presence of one or two $D_s^{(*)}$ are also measured. The individual branching fractions are found to be

$$B(B_s^0 \to D_s^{(*)}+D_s^{(*)}\!\!\! ) = (1.35 \pm 0.06 \pm 0.09 \pm 0.15)\%,$$

$$B(B_s^0 \to D_s^{(*)}+D_s^{(*)}\!\!\! ) = (1.27 \pm 0.08 \pm 0.10 \pm 0.14)\%.$$

All three results are the most precise determinations to date.

Published in Phys. Rev. D 92, 092008 on 20th May 2016

© CERN on behalf of the LHCb collaboration, licence [CC-BY-4.0]

†Authors are listed at the end of this paper.
1 Introduction

Because of $B_s^0 - \overline{B_s^0}$ oscillations, the mass and flavour eigenstates of the $B_s^0$ system do not coincide. The $B_s^0$ meson mass eigenstates have a relative decay width difference $\Delta \Gamma_s / \Gamma_s$, where $\Delta \Gamma_s$ ($\Gamma_s$) is the difference (average) of the decay widths between the heavy and light states. The relative decay width difference is one of the key parameters of the $B_s^0$ system, and its precise determination allows stringent tests of the flavour sector of the Standard Model.

Under certain theoretical assumptions, $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decays were thought to saturate the CP-even contribution to $\Delta \Gamma_s$, and therefore the branching fraction of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ was used as a means of approximating $\Delta \Gamma_s / \Gamma_s$ [1]. This approximation is now considered to be a poor one [2], as the decay modes containing at least one $c$ quark, implied and summed over. The branching fraction ratio is determined as

\[
\frac{B(B_s^0 \to D_s^{(*)+} D_s^{(*)-})}{B(B_s^0 \to D_s^0 D_s^0)} = \frac{f_d}{f_s} \cdot \frac{\epsilon_{B^0}}{\epsilon_{B^0}} \cdot \frac{B(D^- \to K^+ \pi^- \pi^-)}{B(D_s^+ \to K^- K^+ \pi^+)} \cdot \frac{N_{B_s^0}}{N_{B_s^0}},
\]

(1)

where $f_d$ ($f_s$) is the fraction of $B^0$ ($\overline{B_s^0}$) mesons produced in the fragmentation of a $b$ quark, $\epsilon_{B^0}/\epsilon_{B^0}$ is the relative efficiency of the $B^0$ to the $B_s^0$ selections, $B(D^- \to K^+ \pi^- \pi^-)$ and $B(D_s^+ \to K^- K^+ \pi^+)$ are the branching fractions of the charm daughter decays, and $N_{B_s^0}/N_{B_s^0}$ is the relative yield of $B_s^0$ and $B^0$ candidates.

The branching fraction of the exclusive $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decay is determined in the same way, along with the branching fraction of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$. The branching fraction of the
$B_s^0 \rightarrow D_s^+ D_s^-$ decay has been previously measured by LHCb using the same data as this analysis [8], and is therefore not determined in this study. However, the selection efficiency and yield in the $B_s^0 \rightarrow D_s^+ D_s^-$ channel are determined in this analysis, as both are needed for the calculation of the total $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ selection efficiency.

2 Detector and simulation

The LHCb detector [9,10] 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 of the magnet. The tracking system provides a measurement of momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/$c$. The minimum distance of a track to a primary vertex, the impact parameter, is measured with a resolution of $(15 + 29/p_T) \mu$m, where $p_T$ is the component of the momentum transverse to the beam, in GeV/$c$. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons 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 event selection is performed in two stages, with an initial online selection followed by a tighter offline selection. The online event selection is performed by a trigger [11], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which performs a full event reconstruction.

In the simulation, $pp$ collisions are generated using PYTHIA 6 [12] with a specific LHCb configuration [13]. Decays of hadronic particles are described by EVTGEN [14], in which final-state radiation is generated using PHOTOS [15]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [16] as described in Ref. [17].

3 Signal selection

The $D_s^{(*)}$ meson decays to a $D_s^+$ meson and either a photon or a neutral pion (93.5 ± 0.7)% and (5.8 ± 0.7)% of the time, respectively, nearly saturating the total branching fraction. The remainder of the decays are ignored in this analysis. Neither of the neutral particles is reconstructed in the decay chain, and the individual $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ and $B_s \rightarrow D_s^{(*)+} D_s^{(*)-}$ decays are identified through the reconstructed invariant mass of the $D_s^+ D_s^-$ system. The individual peaks from $B_s^0 \rightarrow D_s^{(*)+} (\rightarrow D_s^{(*)+} \gamma) D_s^-$ and $B_s^0 \rightarrow D_s^{(*)+} (\rightarrow D_s^{(*)+} \pi^0) D_s^-$ are not resolved. Therefore the reconstructed $D_s^+ D_s^-$ mass distribution has three separate peaks, corresponding to decays containing zero, one, or two $D_s^{(*)}$ particles.
At the hardware trigger stage, events are required to have a muon with high \( p_T \) or a hadron, photon or electron with high transverse energy in the calorimeters. For hadrons, the transverse energy threshold is 3.5 GeV. Candidate \( B_s^0 \) and \( B^0 \) mesons are used in the analysis if at least one of the associated tracks is selected by the hardware trigger, or if the event is triggered independently of the particles in the signal decay. The software trigger considers all charged particles with \( p_T > 500 \text{ MeV}/c \) and constructs two-, three-, or four-track secondary vertices which require a significant displacement from the primary \( pp \) interaction vertices. At least one charged particle must have a transverse momentum \( p_T > 1.7 \text{ GeV}/c \) and be inconsistent with originating from a primary vertex. A multivariate algorithm \cite{18} is used for the identification of secondary vertices consistent with the decay of a \( b \) hadron. The selection to this point is hereafter referred to as the initial selection.

Signal \( B_s^0 \) and normalisation \( B^0 \rightarrow D_s^+ D^- \) candidates are required to satisfy a number of additional conditions in order to be included in the final samples. Kaons and pions are required to be identified by the particle identification (PID) system. All \( D_s^+ \) and \( D^- \) candidates must have an invariant mass within \( \pm 30 \text{ MeV}/c^2 \) of their known values \cite{19}. Signal \( B_s^0 \) candidates are required to have a reconstructed mass in the range \( 4750 - 5800 \text{ MeV}/c^2 \), whereas \( B^0 \) candidates must have a mass in the range \( 5050 - 5500 \text{ MeV}/c^2 \). After these requirements are applied there are still contributions from other \( b \)-hadron decays into final states with two charm particles. The decays \( A_s^0 \rightarrow A_s^+ (\rightarrow pK^-\pi^+)D_s^- \), where the \( p \) is misidentified as a \( K^+ \), and \( B^0 \rightarrow D_s^+ D^- (\rightarrow K^+\pi^-\pi^-) \), where a \( \pi^- \) is misidentified as a \( K^- \), result in background contamination in the signal channel, while the decay \( B_s^0 \rightarrow D_s^+ D_s^- \) contributes to the background in the normalisation channel if the \( K^+ \) in \( D_s^+ \rightarrow K^+ K^- \pi^+ \) is misidentified as a \( \pi^+ \). As these backgrounds accumulate in reconstructed mass close to the signal peaks, candidates consistent with any one of these background decay hypotheses are rejected in the selection by applying a veto based on the invariant mass of the candidate under the alternative particle type hypotheses. Candidate \( D_s^+ \) mesons are vetoed if they have a reconstructed mass in the range \( 2271 - 2301 \text{ MeV}/c^2 \) when the \( K^+ \) candidate is assumed to be a proton, or a mass in the range \( 1835 - 1905 \text{ MeV}/c^2 \) when the \( K^+ \) candidate is assigned the \( \pi^+ \) mass. Candidate \( D^- \) mesons are vetoed if they have a reconstructed mass in the range \( 1950 - 1990 \text{ MeV}/c^2 \) when a \( \pi^- \) candidate is assigned the kaon mass. In a simulated sample of \( B^0 \rightarrow D_s^+ D^- \) decays, 17.7% of the events meet all of the \( B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \) selection criteria before the \( D^\pm \) veto is applied. After the veto, only 0.05% of the simulated \( B^0 \rightarrow D_s^+ D^- \) sample still pass the full \( B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \) selection. The decay \( B^0 \rightarrow D_s^+ D_s^- \) and three-body \( B^\pm \rightarrow D_s^{(*)+} D_s^{(*)-} h^\pm \) decays, where \( h \) is either a kaon or pion, are examined as other potential background sources, but are all disregarded because of either a small selection efficiency or small branching fraction relative to the signal channels.

In order to further improve the purity of the signal and normalisation samples, a boosted decision tree (BDT) classifier is used to distinguish real \( B_s^{(*)} \) decays from combinatorial background \cite{20}. The BDT is trained using the AdaBoost algorithm \cite{21} to distinguish simulated \( B_s^{(*)} \) signal decays from background candidates obtained from mass sidebands in the data. Background candidates must contain a \( B_s^0 \) candidate with a mass greater than \( 5600 \text{ MeV}/c^2 \) and two \( D_s^\pm \) candidates with masses less than \( 1930 \text{ MeV}/c^2 \) or greater than
Table 1: Efficiencies of the various selection criteria for the three individual channels of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$, and for $B^0 \to D_s^+ D^-$. Each efficiency is presented relative to the previous cut and measured using simulated events, except for the PID efficiency which is obtained from data. The $D_s^+$ veto was only applied to the normalisation mode, $B^0 \to D_s^+ D^-$. 

| Selection | $B_s^0 \to D_s^+ D_s^-$ | $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ | $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ | $B^0 \to D_s^+ D^-$ |
|-----------|------------------------|---------------------------------|---------------------------------|---------------------|
| Reconstruction | 0.1184 ± 0.0003 | 0.1127 ± 0.0005 | 0.1061 ± 0.0005 | 0.1071 ± 0.0002 |
| Initial selection | 1.362 ± 0.008 | 1.250 ± 0.010 | 1.100 ± 0.010 | 1.416 ± 0.009 |
| Mass requirements | 89.4 ± 0.6 | 87.8 ± 1.0 | 88.3 ± 1.0 | 88.5 ± 0.6 |
| BDT | 97.9 ± 0.7 | 96.6 ± 1.1 | 96.7 ± 1.1 | 97.6 ± 0.7 |
| $D^+$ veto | 48.7 ± 0.5 | 50.3 ± 0.8 | 49.8 ± 0.8 | 68.7 ± 0.6 |
| $D_s^+$ veto | | | | |
| $A_s^+$ veto | 96.3 ± 1.0 | 96.3 ± 1.6 | 95.9 ± 1.6 | 98.2 ± 0.8 |
| Trig. requirement | 96.6 ± 0.7 | 96.7 ± 1.1 | 96.6 ± 1.1 | 96.8 ± 0.7 |
| PID requirements | 82.4 ± 0.2 | 82.4 ± 0.2 | 82.4 ± 0.2 | 84.2 ± 0.1 |
| Total | 0.0527 ± 0.0067 | 0.0460 ± 0.0095 | 0.0372 ± 0.0081 | 0.0467 ± 0.0060 |

The efficiencies of the selection criteria in both the signal and normalisation channels are listed in Table 1. The efficiencies of the background vetoes, trigger, reconstruction, and BDT selection are determined using simulated signal samples. The efficiencies of identifying $K^+$ and $\pi^+$ mesons are determined using a calibration data sample of $D^{+} \to D^0(\to K^−\pi^+)\pi^+$ decays, with kinematic quantities reweighted to match those of the signal candidates. The efficiency of the PID selection is found to be $(82.4 \pm 0.2)\%$ for signal $B_s^0$ decays and $(84.2 \pm 0.1)\%$ for $B^0$ decays. The efficiency of the full $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decay is determined by calculating a weighted average of the individual signal channels, with weights given by the relative yields in data. The relative efficiencies of the $B^0$ decay to the three individual channels and the full decay are given in Table 2.
Table 2: Efficiency of the normalisation channel \( B^0 \to D_s^+ D^- \) relative to the signal decays.

| Channel                  | \( \epsilon_{B^0} / \epsilon_{B^0_s} \) |
|--------------------------|-----------------------------------------|
| \( B^0 \to D_s^+ D^- \)  | 0.89 ± 0.02                             |
| \( B^0 \to D_s^{*\pm} D_s^{\mp} \) | 1.02 ± 0.03                         |
| \( B_s^0 \to D_s^{*+} D_s^{*-} \)  | 1.26 ± 0.03                             |
| \( B^0 \to D_s^{(*)+} D_s^{(*)-} \)  | 1.06 ± 0.02                             |

4 Signal and background shapes

The \( B_s^0 \) and \( B^0 \) yields in the signal channels and the normalisation mode are extracted by performing a three-dimensional extended unbinned maximum likelihood fit to the mass distributions of the \( B_s^0 \) meson and the two charm daughters.

In order to determine the yields for the individual signal peaks, the \( B_s^0 \) candidate mass distribution in each channel is modelled using simulated signal events. The \( B^0_s \to D_s^+ D^- \) peak is parameterised as the sum of a Crystal Ball function \(^{[22]}\) and a Gaussian function. The tail parameters of the Crystal Ball function, the ratio of the width of its Gaussian core to the width of the Gaussian function, and the relative weight of each function in the full distribution, are taken from simulation. The mean and width of the Gaussian core are allowed to float. The two \( D_s^{\pm} \) distributions are also parameterised using this model, with all shape parameters fixed to the values found in simulation.

Because of the kinematic differences between the \( D_s^{*+} \to D_s^+ \gamma \) and \( D_s^{*+} \to D_s^+ \pi^0 \) decays, the peak of the \( B^0_s \to D_s^{*\pm} D_s^{\mp} \) mass distribution is parameterised by a superposition of two Gaussian functions. The individual mean values, the ratio of the widths, and the fraction of each Gaussian function in the full distribution are fixed to values taken from simulation. The peak corresponding to \( B^0_s \to D_s^{*+} D_s^{*-} \) decays is modelled using a single Gaussian function, with the mean fixed to the value found from simulated events.

There is also a component in the fit to describe the presence of background decays of the form \( B^0_s \to D_s^{(*)+} \bar{D}_s^{(*)-} \), where the \( D_s^{(*)+} \) can be either a \( D_{s1}(2460)^+ \) or a \( D_{s0}(2317)^+ \) meson that decays to a \( D_s^+ \) along with some combination of photons and neutral or charged pions. As some decay products are missed, this background is present only at the low mass region of the signal distribution. The shape of the distribution is determined by fitting to \( B_s^0 \to D_s(2460)^+ D_s^- \) simulated events, as the contribution from \( D_{s1}(2460)^+ \) is currently the best understood among the \( D_{sJ}^+ \) decays. It is found to be well modelled by an Argus function \(^{[23]}\), all shape parameters for which are fixed to the values found in simulation.

The combinatorial background shape in the \( B_s^0 \) candidate mass distribution is parameterised by a second-order polynomial, and the model is validated with candidates passing a wrong-sign version of the selection. The wrong-sign selection is identical to the signal selection but instead looks for events containing two \( D_s^+ \) mesons. The parameters of the combinatorial background distribution are allowed to float in the full fit to data, and are found to be compatible with the values obtained from the fit to the wrong-sign sample.
Table 3: The yields extracted from the fit to the $B^0_s \to D_s^{(*)+} D_s^{(*)-}$ candidate sample.

| Decay Mode | Yield   |
|------------|---------|
| $B^0_s \to D_s^+ D_s^-$ | $412 \pm 23$ |
| $B^0_s \to D_s^{*+} D_s^{*-}$ | $1032 \pm 39$ |
| $B^0_s \to D_s^{*\pm} D_s^{*-}$ | $786 \pm 48$ |
| Combinatorial background | $1342 \pm 47$ |
| $B^0_s \to D_s(2460)^\pm D_s^\mp$ | $432 \pm 42$ |

The combinatorial background shape in the $D_s^{\pm}$ distribution is determined using events taken from the high-mass sideband region of the $B^0_s$ distribution, and is found to be consistent with a first-order polynomial. The impact of adding a small Gaussian contribution to account for the presence of real $D_s^{\pm}$ mesons in the combinatorial background was found to be minimal, with the observed deviations from the nominal signal yields being smaller than the statistical uncertainty in each case.

The $B^0$ distribution is modelled using the same parameterisation as for the full $B^0_s$ distribution, with one exception. The peak where either the $D_s^+$ or $D^-$ comes from the decay of an excited state is modelled by a superposition of three Gaussian functions, rather than the two-Gaussian model used in the $B^0_s$ case, to account for the difference in distributions from $D_s^{*+}$ and $D^{*-}$ decays, as the $D^{*-}$ decay contains a $\pi^0$ in the final state more frequently than $D_s^{*+}$ decays. There is also a small contribution from the decay $B^0_s \to D_s^- D_s^+$, which is modelled with the same distribution as for the signal $B^0_s$ candidates.

5 Fit results

The fit to the signal data samples is shown in Fig. 1 where the triple peaked structure of the full decay is clearly visible. The yields for the individual signal channels and the two backgrounds are given in Table 3. The total $B^0_s \to D_s^{(*)+} D_s^{(*)-}$ yield is the sum of the individual signal channel yields, with the uncertainty calculated using the correlation coefficients between the individual yields, and is found to be $2230 \pm 63$. The full fit to the data sample for the normalisation mode is shown in Fig. 2 and the yields are given in Table 4. Almost all $B^0 \to D_s^{(*)+} D_s^{(*)-}$ decays are reconstructed with a mass lower than the 5050 MeV/$c^2$ mass cut imposed on the $B^0$ candidates. There is thus a relatively small yield from this channel. Only the main $B^0 \to D_s^+ D^-$ peak is used for normalisation purposes.

6 Systematic uncertainties

A number of systematic uncertainties affect the measurements of the ratios of branching fractions; the sources and magnitudes of these uncertainties are summarised in Table 5. The dominant source of uncertainty for two of the three branching fractions comes from the
Figure 1: Invariant mass distribution of the $B^0 \rightarrow D_s^{(*)}+D_s^{(*)}$ candidates. Also shown is the fit function and the individual components of the fit model.

Table 4: The yields extracted from the fit to the $B^0 \rightarrow D_s^{(*)\pm}D_s^{(*)\mp}$ candidate sample.

| Decay Mode                        | Yield   |
|-----------------------------------|---------|
| $B^0 \rightarrow D_s^+D^-$        | 3636 ± 64 |
| $B^0 \rightarrow D_s^{*+}D^-/D_s^{*-}D_s^+$ | 3579 ± 110 |
| $B^0 \rightarrow D_s^{*+}D^{*-}$  | 166 ± 86  |
| $B_s^0 \rightarrow D_s^{*+}D_s^{*+}$ | 85 ± 13   |
| Combinatorial background          | 1542 ± 56 |

$b$ fragmentation fraction ratio, $f_s/f_d = 0.259 \pm 0.015$ [24]. Part of the uncertainty on this ratio is due to the ratio of the charm branching fractions $B(D_s^+ \rightarrow K^+\pi^-\pi^-)/B(D^- \rightarrow K^+\pi^-\pi^-) = 0.594 \pm 0.020$ [24], the inverse of which is used in the measurements presented in this paper, as shown in Eq. [1]. With the two values from Ref. [24], the part of the uncertainty on $f_s/f_d$ due to the charm branching fractions cancels, leading to a total uncertainty for the product $f_d/f_s \times B(D^- \rightarrow K^+\pi^-\pi^-)/B(D_s^+ \rightarrow K^+K^-\pi^+) = 4.7\%$.

The fit model used for the yield extraction is validated using pseudoexperiments and is found to be unbiased. The uncertainty due to the imperfect knowledge of the shape of the full mass distribution is investigated by measuring the yields using alternative models for each of the peaks. The $D_s^+D_s^-$ peak is modelled with an Apollonios function [25] or a Cruijff function [26], the $D_s^{*+}D_s^{*-}$ peak is modelled using a single Gaussian function, and the $D_s^{*+}D_s^{*-}$ peak is modelled using a combination of two Gaussian functions. The
\( B^0 \rightarrow D_s^{(*)\pm} D^{(*)\mp} \) fit model uncertainty is assessed by modeling the \( B^0 \rightarrow D^- D^+ \) peak with both an Apollonios function and a Cruijff function. In all cases, the systematic uncertainty is taken to be the RMS deviation of the sets of yields with respect to the nominal yields found using the standard fits. The \( B^0 \rightarrow D^- D^+ \) uncertainty is added in quadrature to the signal channel uncertainties, leading to a systematic uncertainty of 3.4\% for the \( D_s^{(*)\pm} D_s^{(*)\mp} \) branching fraction ratio, 2.2\% for the \( D_s^{(*)+} D_s^{(*)-} \) branching fraction ratio, and 2.2\% for the total \( B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \) branching fraction ratio.

The uncertainty on the combinatorial background yield is determined by considering the differences when instead fitting this background with an exponential function, and is of the order of 1.5\% for all of the branching fraction ratios.

The dominant uncertainty for the \( B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \) decay channel results from the lack of knowledge of the \( B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \) background decays. The shape of this background overlaps mostly with the \( B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \) signal decays, and therefore the systematic uncertainty due to this background shape is much larger for this channel (5.0\%) than for the other two exclusive branching fractions (0.2\% – 0.4\%). The uncertainty is measured by repeating the fit with the cut-off point of the Argus function varied from 5050 MeV/c^2 to 5200 MeV/c^2, where the upper limit is chosen in order to account for the presence of decays containing \( D_{s0}(2317)^+ \) mesons. The changes to the yields from the values found in the nominal fit are calculated in each case. The systematic uncertainty in each channel is then assigned as the RMS of the full set of deviations. The uncertainty on the overall branching fraction ratio is also determined in this way, and is found to be 1.9\%.

The uncertainties on the overall efficiencies due to the limited size of the simulated
Table 5: Systematic uncertainties, in % of the relevant branching fraction ratio, for the $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ branching fraction ratios.

| Source                                      | $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ | $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ | $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ |
|---------------------------------------------|---------------------------------|---------------------------------|---------------------------------|
| $f_d/f_s \times \frac{\mathcal{B}(D^- \to K^+\pi^+\pi^-)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)}$ | 4.7                             | 4.7                             | 4.7                             |
| Fit model                                   | 3.4                             | 2.2                             | 2.2                             |
| Comb. background                            | 1.2                             | 1.9                             | 1.5                             |
| $D_{sJ}^+$ background                       | 0.4                             | 5.0                             | 1.9                             |
| Simulation statistics                       | 1.9                             | 2.1                             | 1.9                             |
| PID efficiency                              | 1.4                             | 1.8                             | 1.5                             |
| Trigger efficiency                          | 1.5                             | 1.5                             | 1.5                             |
| Total                                       | 6.6                             | 8.1                             | 6.4                             |

samples are calculated individually for each channel. For the total measurement, $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$, a weighted average of the individual uncertainties is used, with weights proportional to the final yield values obtained from data. These uncertainties on the efficiencies are then propagated to the branching fraction ratios.

There is a systematic uncertainty arising from the calculation of the efficiencies of the PID cuts. The calibration of the data samples is performed in bins of momentum and pseudorapidity, which results in an uncertainty on the calculated efficiencies owing to the finite size of the $D^{(*)+} \to D^0\pi^+$ calibration samples and the binning scheme used. The uncertainty resulting from the calibration sample size and binning scheme is determined by redoing the calibration using different binning schemes. Another systematic uncertainty is due to the presence of a small combinatorial background component in the samples that are used to determine the PID efficiencies. The systematic uncertainty due to this contamination is estimated by comparing the efficiencies found in data to those found when calibrating simulated signal events. The total uncertainties due to the PID efficiency calculation for the three branching fraction ratios presented in this paper are shown in Table 5. The value for $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ is again the weighted average of the contributing channels, with the uncertainty for the $D_s^0 D_s^-$ contribution being 1.1%.

The uncertainty of 1.5% from the trigger response is assessed by considering variations in the response between data and simulation. The individual uncertainties are combined in quadrature to give the total relative systematic uncertainties for each measurement given in Table 5.
7 Summary and discussion

Inserting the measured yields and relative efficiencies into Eq. [1], along with the $f_s/f_d$ and $B(D^- \to K^+\pi^-\pi^-)/B(D_s^+ \to K^+K^-\pi^+)$ values taken from [24], gives

$$\frac{B(B_s^0 \to D_s^{(*)+}D_s^{(*)-})}{B(B^0 \to D_s^+D^-)} = 4.24 \pm 0.14 \text{ (stat)} \pm 0.27 \text{ (syst)},$$

$$\frac{B(B_s^0 \to D_s^{(*)+}D_s^{(*)-})}{B(B^0 \to D_s^+D^-)} = 1.88 \pm 0.08 \text{ (stat)} \pm 0.12 \text{ (syst)},$$

$$\frac{B(B_s^0 \to D_s^{(*)+}D_s^{(*)-})}{B(B^0 \to D_s^+D^-)} = 1.76 \pm 0.11 \text{ (stat)} \pm 0.14 \text{ (syst)}.$$ 

Using the current world average measurement of the $B^0 \to D_s^+D^-$ branching fraction of $(7.2 \pm 0.8) \times 10^{-3}$ [19], gives

$$B(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) = (3.05 \pm 0.10 \pm 0.20 \pm 0.34)\%,$$

$$B(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) = (1.35 \pm 0.06 \pm 0.09 \pm 0.15)\%,$$

$$B(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) = (1.27 \pm 0.08 \pm 0.10 \pm 0.14)\%,$$

where the uncertainties are statistical, systematic, and due to the branching fraction of the normalisation channel, respectively.

Figure 3 shows the LHCb measurement of the total $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ branching fraction, along with the previous measurements by Belle [5], CDF [6], and D0 [7], the average of these previous measurements as calculated by HFAG [27], and the theoretical value from Ref. [3]. The theoretical prediction is for a decay time $t = 0$, while the measurements integrate over the $B_s^0$ meson lifetime; the correction factor for mixing is known [28], but has not been applied. The LHCb result is consistent with all previous measurements and calculations, and is the most precise determination to date. In addition, the LHCb measurements of the individual $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ and $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ branching fractions are consistent with, and more precise than, all previous measurements.

Using this measurement of the branching fraction of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays to calculate $\Delta \Gamma_s/\Gamma_s$ as detailed in Ref. [1] gives a value approximately half as large as the most recent HFAG determination [27], suggesting that indeed $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays do not saturate the $CP$-even modes [2]. The measurements presented in this analysis will be useful in improving the understanding of hadronisation effects in $B_s^0$ decays via the $b \to c\bar{c}s$ quark transition, and in determining a precise value of the inclusive branching fraction of these $b \to c\bar{c}s$ decays.
Figure 3: The $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ branching fraction measurements by Belle, CDF, and D0, the current world average, the theoretical prediction from Ref. [3], and the new LHCb result.

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to the communities behind the multiple open source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).
References

[1] R. Aleksan et al., Estimation of $\Delta \Gamma_s$ for the $B^0_s - \bar{B}^0_s$ system. Exclusive decays and the parton model, Phys. Lett. B 316 (1993) 567

[2] A. Lenz, Theoretical update of $B$-mixing and lifetimes, arXiv:1205.1444.

[3] C.-K. Chua, W.-S. Hou, and C.-H. Shen, Long-distance contribution to $\Delta \Gamma_s/\Gamma_s$ of the $B^0_s - \bar{B}^0_s$ system, Phys. Rev. D84 (2011) 074037, arXiv:1107.4325.

[4] F. Krinner, A. Lenz, and T. Rauh, The inclusive decay $b \rightarrow c\bar{c}s$ revisited, Nucl. Phys. B 876 (2013) 31

[5] Belle collaboration, S. Esen et al., Precise measurement of the branching fractions for $B_s \rightarrow D_s^{(*)}D_s^{(*)}$ and first measurement of the $D_s^{+}D_s^{-}$ polarization using $e^+e^-$ collisions, Phys. Rev. D87 (2013) 031101, arXiv:1208.0323.

[6] CDF collaboration, T. Aaltonen et al., Measurement of $B^0_s \rightarrow D_s^{(*)}D_s^{(*)}$ branching ratios, Phys. Rev. Lett. 108 (2012) 201801, arXiv:1204.0536.

[7] D0, V. M. Abazov et al., Evidence for decay $B^0_s \rightarrow D_s^{(*)}D_s^{(*)}$ and a measurement of $\Delta \Gamma_s^{CP}/\Gamma_s$, Phys. Rev. Lett. 102 (2009) 091801, arXiv:0811.2173.

[8] LHCb collaboration, R. Aaij et al., First observations of $\bar{B}_s^0 \rightarrow D^+D^-, D_s^+D_s^-$ and $D^0\bar{D}^0$ decays, Phys. Rev. D87 (2013) 092007, arXiv:1302.5854.

[9] LHCb collaboration, A. A. Alves Jr. et al., The LHCb detector at the LHC, JINST 3 (2008) S08005.

[10] LHCb collaboration, R. Aaij et al., LHCb detector performance, Int. J. Mod. Phys. A30 (2015) 1530022, arXiv:1412.6352.

[11] R. Aaij et al., The LHCb trigger and its performance in 2011, JINST 8 (2013) P04022, arXiv:1211.3055.

[12] T. Sjöstrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, JHEP 05 (2006) 026, arXiv:hep-ph/0603175.

[13] I. Belyaev et al., Handling of the generation of primary events in Gauss, the LHCb simulation framework, J. Phys. Conf. Ser. 331 (2011) 032047.

[14] D. J. Lange, The EvtGen particle decay simulation package, Nucl. Instrum. Meth. A462 (2001) 152.

[15] P. Golonka and Z. Was, PHOTOS Monte Carlo: A precision tool for QED corrections in $Z$ and $W$ decays, Eur. Phys. J. C45 (2006) 97, arXiv:hep-ph/0506026.
[16] Geant4 collaboration, J. Allison et al., Geant4 developments and applications, IEEE Trans. Nucl. Sci. 53 (2006) 270; Geant4 collaboration, S. Agostinelli et al., Geant4: A simulation toolkit, Nucl. Instrum. Meth. A506 (2003) 250.

[17] M. Clemencic et al., The LHCb simulation application, Gauss: Design, evolution and experience, J. Phys. Conf. Ser. 331 (2011) 032023.

[18] V. V. Gligorov and M. Williams, Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree, JINST 8 (2013) P02013, arXiv:1210.6861.

[19] Particle Data Group, K. A. Olive et al., Review of particle physics, Chin. Phys. C38 (2014) 090001 and 2015 update.

[20] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and regression trees, Wadsworth international group, Belmont, California, USA, 1984.

[21] R. E. Schapire and Y. Freund, A decision-theoretic generalization of on-line learning and an application to boosting, Jour. Comp. and Syst. Sc. 55 (1997) 119.

[22] T. Skwarnicki, A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[23] ARGUS collaboration, H. Albrecht et al., Search for hadronic b → u decays, Phys Lett. B 241 (1990) 278.

[24] LHCb collaboration, Updated average $f_s/f_d$ b-hadron production fraction ratio for 7 TeV pp collisions, [LHCb-CONF-2013-011].

[25] D. Martinez Santos and F. Dupertuis, Mass distributions marginalized over per-event errors, Nucl. Instrum. Meth. A764 (2014) 150, arXiv:1312.5000.

[26] BaBar collaboration, P. del Amo Sanchez et al., Study of $B → X\gamma$ decays and determination of $|V_{td}/V_{ts}|$, Phys. Rev. D82 (2010) 051101, arXiv:1005.4087.

[27] Heavy Flavor Averaging Group, Y. Amhis et al., Averages of b-hadron, c-hadron, and $\tau$-lepton properties as of summer 2014, arXiv:1412.7515, updated results and plots available at http://www.slac.stanford.edu/xorg/hfag/.

[28] K. De Bruyn et al., Branching ratio measurements of $B_s$ decays, Phys. Rev. D86 (2012) 014027, arXiv:1204.1735.
LHCb collaboration

R. Aaij, C. Abellán Beteta, B. Adeva, M. Adinolfi, A. Affolder, Z. Ajaltouni, S. Akar, J. Albrecht, F. Alessio, M. Alexander, S. Ali, M. Alkhazov, P. Alvarez Cartelle, A.A. Alves Jr, S. Amato, S. Amerio, Y. Anghis, L. Ani, L. Anderlini, G. Andreassi, M. Andreotti, J.E. Andrews, R.B. Appleby, O. Aquines Gutierrez, F. Archilli, P. d’Argenti, A. Artamonov, M. Artuso, E. Aslanides, G. Auriemma, M. Balaouch, S. Bachmann, J.J. Back, A. Badalov, C. Baesso, W. Baldini, R.J. Barlow, C. Barschel, S. Barsuk, W. Barter, V. Batozskaya, V. Battista, A. Bay, L. Beaucourt, J. Beddow, F. Bedeschi, I. Bediaga, L.J. Bel, V. Bellee, N. Belloli, I. Belyaev, E. Ben-Haim, G. Bencivenni, S. Benson, J. Benton, A. Berezin, A. Bertolli, M.-J. Betti, M. van Beuzekom, S. Bifani, P. Billoir, T. Bird, A. Birnkraut, A. Bizzeti, I. El Rifai, A. Dosil Suárez, G. Bencivenni, N.H. Brook, E. Buchanan, C. Burr, A. Bursche, J. Buytaert, S. Caddei, R. Calabrese, M. Calvi, M. Calvo Gomez, D. Campora Perez, L. Capriotti, A. Carbone, R. Cardinale, A. Cardini, P. Carniti, K. Carvalho Akiba, G. Casse, L. Cassina, L. Castillo Garcia, M. Cattaneo, Ch. Cauet, G. Cavallero, R. Cenci, M. Charles, Ph. Charpentier, M. Chefestievile, S. Chen, F.-S. Cheung, N. Chiapolini, M. Chrzasteczko, C. Cid Vidal, G. Ciezarek, P.E.L. Clarke, M. Clemente, H.V. Cliff, J. Cloies, V. Coco, J. Cogan, E. Cogneras, V. Cogoni, L. Cojocariu, G. Collazuol, P. Collins, A. Comerma-Montells, A. Contu, A. Cook, M. Combes, S. Coquereau, G. Corti, M. Corvo, B. Couturier, G.A. Cowan, D.C. Craik, A. Crocombe, M. Cruz Torres, S. Cunliffe, R. Currie, C. D’Ambrosio, E. Dall’Occo, J. Dalseno, P.N.Y. David, A. Davis, O. De Aguiar Francisco, K. De Bruyn, S. De Capua, M. De Cian, J.M. De Miranda, L. De Paula, P. De Simone, C.-T. Dean, D. Decamp, M. Deckenhoff, L. Del Buono, N. Déllé, M. Demmer, D. Derkach, O. Deschamps, F. Dettori, B. Dey, A. Di Canto, F. Di Ruscio, H. Dijkstra, S. Donleavy, F. Dorder, M. Dorigo, A. Dosil Suárez, A. Dovbnya, K. Dreimanis, L. Dufour, G. Dujany, K. Dungs, P. Durante, R. Dzhelyadin, A. Dziurda, A. Dzyubenko, S. Easo, U. Egede, V. Egorychev, S. Eidelman, S. Eisenhardt, U. Eitschberger, R. Elkhalil, L. Eklund, I. El Rifi, Ch. Elsasser, L. El, S. Eles, S. Esen, H.M. Evans, T. Evans, A. Falabella, C. Färber, N. Farley, S. Farry, R. Fay, D. Ferguson, V. Fernandez Albor, F. Ferrari, F. Ferreira Rodrigues, M. Ferro-Luzzi, S. Filipov, M. Fiore, M. Fiorini, M. Firlej, C. Fitzpatrick, T. Fintowski, F. Fleuret, K. Fohl, P. Foll, M. Fontana, F. Fontanelli, J.C. Forsaw, R. Forty, M. Frank, C. Frei, M. Froshini, J. Fri, E. Furfaro, A. Gallas Torreira, D. Galli, S. Gallorini, S. Gambetta, M. Gandelman, P. Gandini, Y. Gao, J. García Pardiñas, J. Garra Tico, G. Garrido, D. Gascon, C. Gaspar, R. Gaul, L. Gavardi, G. Gazzoni, D. Gerick, E. Gersabeck, M. Gersabeck, T. Gershon, Ph. Ghez, S. Giani, V. Gibson, O.G. Girard, L. Giubega, V.V. Gligorov, C. Göbel, D. Golubkov, A. Golutvin, A. Gomes, A. Gotti, M. Grabalosa Gándara, R. Graciani Díaz, L.A. Granado Cardoso, E. Graugés, E. Graverini, G. Graziani, A. Greco, E. Greenwood, P. Griffith, L. Grillo, O. Grünberg, B. Guí, E. Gushchin, Y. Gu, 39
Yandex School of Data Analysis, Moscow, Russia, associated to 32
Instituto de Fisica Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain, associated to 37
Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to 42

a Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
b Laboratoire Leprince-Ringuet, Palaiseau, France
c P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
d Università di Bari, Bari, Italy
e Università di Bologna, Bologna, Italy
f Università di Cagliari, Cagliari, Italy
g Università di Ferrara, Ferrara, Italy
h Università di Urbino, Urbino, Italy
i Università di Modena e Reggio Emilia, Modena, Italy
j Università di Genova, Genova, Italy
k Università di Milano Bicocca, Milano, Italy
l Università di Roma Tor Vergata, Roma, Italy
m Università di Roma La Sapienza, Roma, Italy
n Università della Basilicata, Potenza, Italy
o AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
p LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
q Hanoi University of Science, Hanoi, Viet Nam
r Università di Padova, Padova, Italy
s Università di Pisa, Pisa, Italy
t Scuola Normale Superiore, Pisa, Italy
u Università degli Studi di Milano, Milano, Italy

† Deceased