A search for rare $B \to D\mu^+\mu^-$ decays

The LHCb collaboration

E-mail: neville.harnew@physics.ox.ac.uk

ABSTRACT: A search for rare $B \to D\mu^+\mu^-$ decays is performed using proton-proton collision data collected by the LHCb experiment, corresponding to an integrated luminosity of 9 fb$^{-1}$. No significant signals are observed in the non-resonant $\mu^+\mu^-$ modes, and upper limits of $\mathcal{B}(B^0 \to D^0\mu^+\mu^-) < 5.1 \times 10^{-8}$, $\mathcal{B}(B^+ \to D^+\mu^+\mu^-) < 3.2 \times 10^{-8}$, $\mathcal{B}(B^0_s \to \bar{D}^0\mu^+\mu^-) < 1.6 \times 10^{-7}$ and $f_c/f_u \cdot \mathcal{B}(B^+_c \to D^+_s\mu^+\mu^-) < 9.6 \times 10^{-8}$ are set at the 95% confidence level, where $f_c$ and $f_u$ are the fragmentation fractions of a $B$ meson with a $c$ and $u$ quark respectively in proton-proton collisions. Each result is either the first such measurement or an improvement by three orders of magnitude on an existing limit. Separate upper limits are calculated when the muon pair originates from a $J/\psi \to \mu^+\mu^-$ decay. The branching fraction of $B^+_c \to D^+_s J/\psi$ multiplied by the fragmentation-fraction ratio is measured to be

$$\frac{f_c}{f_u} \cdot \mathcal{B}(B^+_c \to D^+_s J/\psi) = (1.63 \pm 0.15 \pm 0.13) \times 10^{-5},$$

where the first uncertainty is statistical and the second systematic.

KEYWORDS: B Physics, Flavour Physics, Hadron-Hadron Scattering, Rare Decay

ArXiv ePrint: 2308.06162
1 Introduction

Heavy-flavour decays that are suppressed according to the Standard Model (SM) are probes for new particle currents with mass scales beyond the reach of direct searches. Recent work focuses on the measurement of partial rates in $b \to s\ell\ell$ transitions ($\ell\ell = \mu\mu$ or $ee$) [1–3]. Such flavour-changing neutral currents mediate $B \to K^{(*)}\mu^+\mu^-$ decays, predominantly via electroweak penguin amplitudes. The weak amplitude whereby the quark and antiquark of the initial meson undergo annihilation can also contribute, however this contribution can often be neglected due to suppression by a factor $O(\Lambda_{QCD}/m_b)$ [4]. In order to validate this assumption, it is valuable to gather information about decays that are dominated by different modes of interaction between the quarks of the initial-state meson, either through a direct $b \to q\ell\ell$ decay, or proceeding through a $J/\psi \to \ell\ell$ intermediate state.

The $B_{(s)}^0 \to D^0\mu^+\mu^-$ decays\footnote{Throughout this paper, the inclusion of charge-conjugate states is implied.} proceed via an internal scatter process, specifically a $W$-exchange as shown in figure 1(a). Similar $W$-exchange decays, $B_{(s)}^0 \to D^0 J/\psi$, are depicted in figure 1(d). The clean experimental signature $J/\psi \to \mu^+\mu^-$ can be used to search for this decay by isolating a region in the dimuon spectrum around the known $J/\psi$ mass. For this mode, the BaBar collaboration set a limit on the branching fraction of $\mathcal{B} \left( B^0 \to D^0 J/\psi \right) < 1.3 \times 10^{-5}$ at the 90% confidence level (CL) [5], which is 2–3 orders of magnitude greater than the theoretical expectation of $10^{-8}$–$10^{-7}$ [6]. An operator product expansion is used in ref. [7] to predict a branching fraction of $\sim 3 \times 10^{-9}$ for $B^0 \to D^0 \ell^+\ell^-$ decays, integrated over a range of four-momentum transfer $q^2$ between 1.0–5.0 GeV$^2$/c$^4$. Elsewhere, a calculation based on perturbative QCD suggests a higher value, $\mathcal{O}(10^{-5})$ [8], although the treatment of long-distance divergences in that calculation has been questioned [9]. No experimental limit has been set for the short-distance $B^0 \to D^0 \ell^+\ell^-$ mode, but limits exist for the related
radiative mode $B \left( B^0 \to D^{0 \gamma} \right) < 2.5 \times 10^{-5}$ at the 90% CL [10], still greater than the $\sim 10^{-6}$ prediction [11].

The present analysis is extended to include charged $B \to D \mu^+ \mu^-$ decays. Knowledge of the decays of doubly-heavy $B^+_c$ mesons is accumulating rapidly, though most observed decays can be attributed to favoured amplitudes with an external $W$-emission, e.g. $B^+_c \to J/\psi \ell^+ \nu$ and $B^+_c \to J/\psi \pi^+$ decays [12, 13]. The $B^+_c \to D^+_s J/\psi$ decay, first reported in ref. [14] and measured more recently in refs. [15, 16], is mediated by the external $W$-emission process but is also sensitive to colour-suppressed, annihilation and electroweak penguin amplitudes (figure 1(f), (e) and (c), respectively). To assess the relative magnitude of these sub-leading amplitudes, measurements across a range of suppressed decays must be made. The partial decay rate of the recently observed $B^+_c \to D^0 K^+$ decay together with the absence of $B^+_c \to D^0 \pi^+$ decays [17] implies the former is dominated by annihilation or electroweak penguin contributions. Charmless $B^+_c$ decays isolate the annihilation diagram, but no evidence for such decays has yet emerged [18, 19]. This paper presents the search for yet-unobserved $B^+_c \to D^+_s \mu^+ \mu^-$ decays, which are sensitive to electroweak penguin and radiative annihilation transitions only (figure 1(b) and (c), respectively). The SM branching fraction for this decay mode is predicted to be $O(10^{-8})$ but contributions from beyond-SM processes can raise this by an order of magnitude [20]. Similarly, since the $D^+_s$ meson in the $B^+ \to D^+_s \mu^+ \mu^-$ decay does not maintain any of the initial-state quark content, this transition must proceed via annihilation (figure 1(b) and (e)). No search results have yet been reported for such decays.

The $B \to D \mu^+ \mu^-$ decays depicted in figure 1 represent eight independent decay modes whose search is reported in this paper, namely:

- **$B^0$ decay modes:** $B^0 \to D^0 \mu^+ \mu^-$, $B^0 \to D^0 J/\psi$,
- **$B^+$ decay modes:** $B^+ \to D^+_s \mu^+ \mu^-$, $B^+ \to D^+_s J/\psi$,
- **$B^0_s$ decay modes:** $B^0_s \to D^{0}_s \mu^+ \mu^-$, $B^0_s \to D^{0}_s J/\psi$, 

![Figure 1. The leading diagrams for the decays: (a) $B^0_{(s)} \to D^0 \mu^+ \mu^-$, (b) $B^+_c \to D^+_s \mu^+ \mu^-$, (c) $B^+_c \to D^+_s J/\psi$, (d) $B^0_{(s)} \to D^0 J/\psi$, (e) $B^+_c \to D^+_s J/\psi$ and (f) $B^+_c \to D^+_s J/\psi$. Dashed lines represent the charged weak current, and a gluon exchange with unconnected quark lines is implied. Where two quark currents are indicated, the quark in parenthesis leads to a CKM-suppressed process.](image-url)
• $B_c^+$ decay modes: $B_c^+ \rightarrow D_s^+ \mu^+ \mu^-$, $B_c^+ \rightarrow D_s^+ J/\psi$.

The decay modes $B^+ \rightarrow D^+ \mu^+ \mu^-$ and $B_c^+ \rightarrow D^+ \mu^+ \mu^-$, either proceeding directly to $\mu^+ \mu^-$ or through a $J/\psi$ intermediate state, are not included here. There are large Cabbibo suppression factors associated with these modes with respect to the $B^+ \rightarrow D_s^+ \mu^+ \mu^-$ and $B_c^+ \rightarrow D_s^+ \mu^+ \mu^-$ modes, and as such are significantly below our current sensitivity.

The searches are performed using proton-proton collision data collected by the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV, corresponding to a total integrated luminosity of 9 fb$^{-1}$. The $D^0$ meson candidates are reconstructed in the $D^0 \rightarrow K^+ \pi^-$ channel, and the $D_s^+$ meson candidates in the $D_s^+ \rightarrow K^+ K^- \pi^+$ channel. The $J/\psi$ candidates are reconstructed in the $J/\psi \rightarrow \mu^+ \mu^-$ decay mode. For the non-$J/\psi$ modes, the square of the dimuon invariant mass, $q^2 (\mu^+ \mu^-)$, is selected within the range 0.044–8.0 GeV$^2$/c$^4$, which excludes the $J/\psi$ region. The selection, efficiency and background corrections are finalised before the signal measurements are performed, following a strategy of blinding to ensure unbiased results.

For normalisation, kinematically similar decays with higher branching fractions are also reconstructed. The $B^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^{*0} (\rightarrow K^+ \pi^-)$ decay is used to normalise the neutral $B^0$ and $B^0_s$ signal modes, while the $B^+ \rightarrow J/\psi K^+$ decay is used to normalise the two $B^+$ signal modes and the $B_c^+ \rightarrow D_s^+ J/\psi$ measurement. The $B_c^+ \rightarrow D_s^+ \mu^+ \mu^-$ decay is normalised with respect to the observed $B_c^+ \rightarrow D_s^+ J/\psi$ decay mode.

2 Detector, triggering and simulation

The LHCb detector [21, 22] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, and 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 (VELO) [23] surrounding the proton-proton interaction region, a large-area silicon-strip detector [24] 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 [25] placed downstream of the magnet. The tracking system provides a measurement of the 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 (IP), 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, expressed in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov (RICH) detectors [26]. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers [27].

The online event selection is performed by a trigger [28, 29], consisting of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage that performs a full event reconstruction. The hardware trigger selects $B \rightarrow D \mu^+ \mu^-$ candidates containing at least one muon with a large transverse momentum, a pair of muons with a large product of their transverse momenta, or containing large energy deposits in
the calorimeters from particles related to the $D$-meson decay. The transverse momentum threshold of the muon triggers varied in the range between 1 and 3 GeV/c, depending on the data-taking conditions, and similarly the hadronic energy threshold varied between 3 and 4 GeV/c. At the software stage, the trigger required a two-, three- or four-track secondary vertex with a significant displacement from any proton-proton interaction vertex [30]. At least one charged particle must have a $p_T$ greater than 1.5 GeV/c.

Simulation is used to measure the efficiency of both the detector acceptance and the applied selection requirements. In the simulation, proton-proton collisions are generated using Pythia [31, 32] with a specific LHCb configuration [33]. Unstable particles are described by EvtGen [34], with the rare- $B$ modes decayed uniformly in phase space and final-state radiation generated using Photos [35]. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit [36, 37] as described in ref. [38].

3 Candidate selection

When reconstructing $D$ meson candidates, all hadron tracks are required to be compatible with their respective particle hypotheses according to a likelihood-based variable that uses mainly the RICH input from the particle identification (PID) system. The candidates must have a good-quality vertex and have travelled a significant distance from the $B$-candidate decay vertex. The $\bar{D}^0 \rightarrow K^+\pi^-$ ($D_s^+ \rightarrow K^+K^-\pi^+$) combinations are required to be within $\pm25$ MeV/$c^2$ ($\pm20$ MeV/$c^2$) of the known $\bar{D}^0$ ($D_s^+$) mass [39], corresponding to roughly $\pm3$ standard-deviation windows around the respective mass peaks.

A PID cut is applied to both muon candidates, using an equivalent likelihood-based variable type as for the hadrons. For the signal modes involving a $J/\psi \rightarrow \mu^+\mu^-$ reconstruction, the dimuon invariant mass is required to be within $\pm36$ MeV/$c^2$ of the known $J/\psi$ mass [39], i.e. a $\pm3$ standard-deviation window. For the non-$J/\psi$ signal decays, $q^2(\mu^+\mu^-)$ is selected within the range 0.044–8.0 GeV$^2$/c$^4$.

With the $D$ meson and muon pair identified, $B$ meson candidates are then reconstructed with the $D$, and where appropriate the $J/\psi$, masses constrained to their known values [40]. For modes with a primary $B^0$, $B^0_s$ or $B^+$ meson, a decay time greater than 0.2 ps is required, whereas for modes with a primary $B_c^+$ meson, the minimum decay time is relaxed to 0.05 ps due to the shorter lifetime of this particle. These values can be compared to the measured lifetime resolution at LHCb which is approximately 50 fs [23]. The decay-time requirements correspond to typical flight paths of around 7 mm for the $B^0$, $B_s^0$ and $B^+$ and 2.5 mm for the $B_c^+$, to be compared to the flight-path resolution of approximately 230 $\mu$m.

A boosted decision tree (BDT) algorithm [41, 42], implemented using the TMVA toolkit [43, 44], is used to separate signal from combinatorial background. The BDT classifier is trained and applied independently for three families of modes: neutral $B$ mesons including both $B^0$ and $B_s^0$ (since the final states for the searches are identical), charged $B$ mesons, and $B_c^+$ candidates. Variables describing the properties of the $B$ and $D$ meson candidates are provided to train the BDT classifiers: a total of 15 for the neutral decays and an additional two for $B^+$ and $B_s^+$ modes which include a $D_s^+$ decay. The variables used comprise kinematic and topological quantities such as $p_T$ and IP, flight distance, direction of the weakly
decaying heavy mesons and the angle between their reconstructed trajectory and momentum. Simulated samples containing signal decays are used as target samples in the BDT training, while background samples are taken from data. For the neutral B meson BDT algorithm, decays with a B^0 invariant mass greater than 5800 MeV/c^2 are used to form the background sample. For the charged B meson BDT classifiers, a pure combinatorial background sample of wrong-charge B \rightarrow D \mu^+ \mu^- combinations is used; this is to avoid feed-down effects from B^+_c decays if training on the B^+ upper mass sideband. The selection that is applied to the BDT output is optimised using a Punzi figure of merit [45] to choose a balance between signal efficiency and background rejection for each of the eight search categories.

The normalisation samples containing prolific B^0 \rightarrow J/\psi K^{*0} and B^+ \rightarrow J/\psi K^+ decays are selected without a multivariate technique. The J/\psi and K^{*0} candidates are required to have a good-quality vertex and be within \pm 50 MeV/c^2 of their known masses. The particle types of their decay-product tracks are required to be correctly identified using PID selection criteria. Additional selection requirements are applied on the transverse momenta of the final-state tracks. The yields of B^0 \rightarrow J/\psi K^{*0} and B^+ \rightarrow J/\psi K^+ decays are determined using binned maximum-likelihood fits to data. An S-wave contribution of 4% under the K^{*0} resonance peak is numerically subtracted from the fitted B^0 \rightarrow J/\psi K^{*0} yield as for ref. [46].

4 Signal yield estimation

Six invariant-mass distributions of B candidate decays are fitted, corresponding to B^{0(s)} \rightarrow \bar{D}^0 \mu^+ \mu^-, B^+ \rightarrow D_s^+ \mu^+ \mu^- and B^+_c \rightarrow D_s^+ \mu^+ \mu^-, selected with and without a J/\psi intermediate state. In all samples, the distributions are described by probability density functions (PDFs) with shape parameters that are free to vary. The yields of signal and background contributions are measured using extended unbinned maximum-likelihood fits to the invariant-mass distributions, with the single exception of the B^{0(s)} \rightarrow \bar{D}^0 J/\psi decay where the background contribution is fixed as described below. Signal B \rightarrow D \mu^+ \mu^- decays are modelled by the sum of two Crystal-Ball functions [47], each consisting of a Gaussian core and a power-law tail, where the widths and tail parameters are fixed based on simulation. In the case of the neutral B invariant-mass distributions, two such functions are used to describe each of the B^0 and B^0_s candidate decays. All means are fixed from values measured in the appropriate normalisation-mode fit or, for the B^{0(s)}_s modes where normalisation-mode yields are lower, to their known values [39] with a small offset derived from simulation. In all samples, the combinatorial backgrounds are modelled by exponential distributions with parameters that are free to vary. Specific background contributions are also included, which are described below. The fits to the invariant-mass distributions for the B \rightarrow D \mu^+ \mu^- and B \rightarrow DJ/\psi decays are shown in figures 2, 3 and 4 for neutral B^{0(s)} mesons, charged B^+ and B^+_c mesons, respectively. The fit ranges are defined by the limits of the plots.

It is necessary to include additional background components to fully describe the data for specific decay modes. In the B^0 \rightarrow \bar{D}^0 \mu^+ \mu^- invariant-mass distribution, B^0 \rightarrow \bar{D}^0 \pi^- \mu^+ \nu_\mu and B^0 \rightarrow \bar{D}^{*0} \pi^- \mu^+ \nu_\mu decays with \bar{D}^{*0} \rightarrow D^0 (\pi^0/\gamma) constitute a background if the \pi^- is misidentified as a \mu^- and the neutral particles are not reconstructed. This contribution is modelled using a kernel density estimation technique [48] on events from a simplified
as for the background in such a way that it is determined using simulation and the yield is fixed in the same manner. For the mass fit of the sidebands, the non-resonant contribution within the mass range is estimated to be 5.2 ± 2.9 events, with the yield then fixed to this central value. For the mass fit of the $B^+ \rightarrow D^+_s J/\psi$ decay, $B^+ \rightarrow J/\psi K^+ \pi^- \pi^+$ decays may enter if the $\pi^-$ is misidentified as a $K^-$. The shape of such a background is determined using simulation and the yield is fixed in the same manner as for the background in the $B^0$ samples.

In the fits to the invariant-mass distributions of the $B_c$ modes show in figure 4, no $B_c^+ \rightarrow D^+_s \mu^+\mu^-$ signal is observed whilst a $B_c^+ \rightarrow D^+_s J/\psi$ signal is clearly visible. For the $B_c^+ \rightarrow D^{*+}_s J/\psi$ candidates, $B_c^+ \rightarrow D^{*+}_s J/\psi$ decays are considered where $D^{*+}_s$ decays into $D^{*+}_s \rightarrow D^+_s \gamma$ or $D^{*+}_s \rightarrow D^+_s \pi^0$ and the neutral particles are not reconstructed. This results in
the broad structure seen between 5.9 and 6.2 GeV/$c^2$ in figure 4. As $B^+_s \rightarrow D^+_s \mu^+ \mu^-$ is a decay of a pseudoscalar into two vector particles, the decay is described by three helicity amplitudes: $A^{00}$ (longitudinal polarisation), $A^{++}$ and $A^{--}$ (transverse polarisations). The latter two are combined into a single $A^{\pm\pm}$ PDF because the $D^+_s J/\psi$ invariant-mass distribution is identical for the $A^{++}$ and $A^{--}$ amplitudes. The two helicity contributions are described by analytic PDFs in the fit to data. The $A^{\pm\pm}$ distribution is parameterised by an upward-open parabola whose range is defined by the kinematic endpoints of the decay and convolved with a Gaussian resolution function, resulting in a characteristic double-peaked shape. The $A^{00}$ distribution is described by a parabola exhibiting a maximum and convolved with a Gaussian resolution function, resulting in a broad single peak. These PDFs are further described in ref. [50]. All shape parameters are fixed from fits to simulation.

5 Branching fraction determination

The signal branching fractions are obtained from efficiency-corrected yield measurements of the signal and normalisation modes scaled by the known branching fractions. Specifically,

$$\mathcal{B}(B \rightarrow D \mu^+ \mu^-) = \frac{N^{B \rightarrow D \mu^+ \mu^-}}{N^{\text{norm}}} \cdot \frac{\varepsilon^{\text{norm}}}{\varepsilon^{B \rightarrow D \mu^+ \mu^-}} \cdot \frac{\mathcal{B}_{\text{sub-decays}}^{\text{norm}}}{\mathcal{B}_{\text{sub-decays}}^{B \rightarrow D \mu^+ \mu^-}},$$

where $\mathcal{B}(B \rightarrow D \mu^+ \mu^-)$ is the branching fraction of the relevant signal decay, $N^i$ refers to the observed yield for the signal decay or normalisation mode, $\varepsilon^i$ is the corresponding total selection efficiency, and $\mathcal{B}^{\text{norm}}$ is the branching fraction of the normalisation channel. The product of branching fractions of the relevant sub-decays, $\mathcal{B}_{\text{sub-decays}}^i$, are listed with all relevant branching fractions in table 1. These values are taken from ref. [39], with the exception of $\mathcal{B}(B^+_s \rightarrow D^+_s J/\psi)$ which uses the value measured in this paper when used as a normalisation mode for the $B^+_s \rightarrow D^+_s \mu^+ \mu^-$ decay. For the $B^0_s$ modes, the result takes into account the ratio of the fragmentation fractions of $B^0_s$ to $B^0$ meson production at LHCb, $f_s/f_d$. The value used in this paper is $0.249 \pm 0.020$, taken from ref. [51], which is a luminosity-weighted average taken over the values at the three centre-of-mass energies 7, 8
### Table 1.
A summary of the branching fractions of the decays (taken from ref. [39]) and the fragmentation ratio (taken from ref. [51]) which are used in the calculation of eq. (5.1). The value of $\mathcal{B}(K^*^0 \to K^+\pi^-)$ is taken from the ratio of isospin amplitudes.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{B}(B^0 \to J/\psi K^0)$</td>
<td>$(1.27 \pm 0.05) \times 10^{-3}$</td>
</tr>
<tr>
<td>$\mathcal{B}(B^+ \to J/\psi K^+)$</td>
<td>$(1.020 \pm 0.019) \times 10^{-3}$</td>
</tr>
<tr>
<td>$\mathcal{B}(\bar{D}^0 \to K^+\pi^-)$</td>
<td>$(3.947 \pm 0.030) \times 10^{-2}$</td>
</tr>
<tr>
<td>$\mathcal{B}(D^+_s \to K^+K^-\pi^+)$</td>
<td>$(5.38 \pm 0.10) \times 10^{-2}$</td>
</tr>
<tr>
<td>$\mathcal{B}(J/\psi \to \mu^+\mu^-)$</td>
<td>$(5.961 \pm 0.033) \times 10^{-2}$</td>
</tr>
<tr>
<td>$\mathcal{B}(K^{*0} \to K^+\pi^-)$</td>
<td>$2/3$</td>
</tr>
<tr>
<td>$f_s/f_d$</td>
<td>$0.249 \pm 0.020$</td>
</tr>
</tbody>
</table>

and 13 TeV. The quoted error is associated with the averaging procedure, and also includes an uncertainty due to the integrated $p_T$ dependence of the $f_s/f_d$ values. The analogous quantity for $B^+_c$ production, $f_c/f_u = (7.5 \pm 1.8) \times 10^{-3}$ [52], is also luminosity averaged, and is kept as a separate factor in the quoted results for the two $B^+_c$ modes due to the relatively poor determination of this quantity to date.

The acceptance, reconstruction and selection efficiencies are determined from simulation, with corrections applied for the particle identification response of the detector. The latter corrections are obtained as a function of muon and hadron track momentum and pseudorapidity, using a set of calibration channels taken from data [53]. These give clean samples of particle species, independent of LHCb RICH- and muon-PID identification, from low-background decays such as $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$. For the branching fraction measurements of $B^+$ modes and the $B^+_c \to D^+_s J/\psi$ decay, the efficiency of the normalisation mode is corrected for the reconstruction efficiency of the different number of tracks in the signal modes: three versus five, respectively. The correction and its systematic uncertainty are derived from tag-and-probe tracking efficiency measurements made as a function of track momentum and pseudorapidity [54]. The ratio of efficiencies is also corrected for differences in the trigger efficiency between simulation and data. These corrections are measured in data using a technique that compares the trigger efficiency when a signal candidate is used in the trigger decision with the case where the decision is taken independently of the signal candidate [55].

All data at the three centre-of-mass energies are fitted together, and the final efficiency used in measuring the branching fraction for each mode is calculated from a luminosity-scaled combination of the efficiencies measured for each dataset, weighted by the ratio of the $b$-quark production cross-sections, given in refs. [56, 57]. After all contributions are considered, table 2 shows the fitted yields and efficiencies for the decays $B \to D\mu^+\mu^-$, $B \to DJ/\psi$ and the normalisation modes (before S-wave subtraction). The total efficiency is around 0.1% for the signal decays. The efficiencies after selection of the normalisation modes are 0.6% and 2.0% for $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ decays, respectively.

Many of the systematic uncertainties cancel in the ratio of efficiencies by construction of the signal branching fractions expressed in eq. (5.1), nevertheless several systematic
<table>
<thead>
<tr>
<th>Decay mode</th>
<th>Fitted yields</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \to \bar{D}^0 \mu^+ \mu^-$</td>
<td>0.3 ± 2.4</td>
<td>0.121 ± 0.003</td>
</tr>
<tr>
<td>$B^+ \to D^+_s \mu^+ \mu^-$</td>
<td>0.0 ± 1.1</td>
<td>0.084 ± 0.002</td>
</tr>
<tr>
<td>$B^0_s \to \bar{D}^0 \mu^+ \mu^-$</td>
<td>0.0 ± 1.1</td>
<td>0.129 ± 0.004</td>
</tr>
<tr>
<td>$B^+_c \to D^+_s \mu^+ \mu^-$</td>
<td>0.0 ± 1.1</td>
<td>0.025 ± 0.001</td>
</tr>
<tr>
<td>$B^0 \to \bar{D}^0 J/\psi$</td>
<td>2.4 ± 2.9</td>
<td>0.115 ± 0.003</td>
</tr>
<tr>
<td>$B^+ \to D^+_s J/\psi$</td>
<td>0.2 ± 1.1</td>
<td>0.129 ± 0.006</td>
</tr>
<tr>
<td>$B^0_s \to \bar{D}^0 J/\psi$</td>
<td>0.0 ± 1.1</td>
<td>0.098 ± 0.004</td>
</tr>
<tr>
<td>$B^+_c \to D^+_s J/\psi$</td>
<td>127 ± 12</td>
<td>0.076 ± 0.002</td>
</tr>
</tbody>
</table>

Table 2. Fitted yields and efficiencies for the decays $B \to D \mu^+ \mu^-$, $B \to D J/\psi$ and the normalisation modes (before S-wave subtraction). The quoted uncertainties arise solely from the limited size of the simulation samples.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>External [%]</th>
<th>Simulation [%]</th>
<th>PDF [%]</th>
<th>Normalisation procedure [%]</th>
<th>Total [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B \left( B^0 \to \bar{D}^0 \mu^+ \mu^- \right)$</td>
<td>4.0</td>
<td>10.3</td>
<td>0.4</td>
<td>-</td>
<td>11.1</td>
</tr>
<tr>
<td>$B \left( B^0 \to \bar{D}^0 J/\psi \right)$</td>
<td>4.0</td>
<td>11.8</td>
<td>0.3</td>
<td>-</td>
<td>12.5</td>
</tr>
<tr>
<td>$B \left( B^0_s \to \bar{D}^0 \mu^+ \mu^- \right)$</td>
<td>8.8</td>
<td>11.1</td>
<td>0.3</td>
<td>-</td>
<td>14.2</td>
</tr>
<tr>
<td>$B \left( B^+_c \to D^+_s \mu^+ \mu^- \right)$</td>
<td>8.8</td>
<td>12.2</td>
<td>0.3</td>
<td>-</td>
<td>15.0</td>
</tr>
<tr>
<td>$B \left( B^+ \to D^+_s \mu^+ \mu^- \right)$</td>
<td>4.3</td>
<td>7.9</td>
<td>0.5</td>
<td>3.0</td>
<td>9.5</td>
</tr>
<tr>
<td>$B \left( B^+ \to D^+_s J/\psi \right)$</td>
<td>4.3</td>
<td>9.9</td>
<td>0.7</td>
<td>3.0</td>
<td>11.2</td>
</tr>
<tr>
<td>$f_c/f_u \cdot B \left( B^+_c \to D^+_s \mu^+ \mu^- \right)$</td>
<td>12.2</td>
<td>9.1</td>
<td>1.9</td>
<td>-</td>
<td>15.3</td>
</tr>
<tr>
<td>$f_c/f_u \cdot B \left( B^+_c \to D^+_s J/\psi \right)$</td>
<td>4.3</td>
<td>5.9</td>
<td>1.2</td>
<td>3.0</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Table 3. Relative systematic uncertainties for the branching-fraction measurements defined in eq. (5.1), quoted separately for the different sources. The total is the sum in quadrature along the corresponding row.

Contributions remain, which are summarized in table 3. The first of these is the external uncertainty imported from the established branching fraction measurements [39]. For the $B^0_s$ modes, the uncertainty of $f_s/f_d$ is also added in quadrature. The $B^+_c \to D^+_s \mu^+ \mu^-$ signal channel has a larger uncertainty due to the low signal yield of the normalisation-mode channel, however the yields of the other normalisation samples incur negligible errors on the branching-ratio measurements. The second is the precision on the efficiencies used in eq. (5.1) which is a result of the uncertainties on the accuracy of the simulation’s replication of the data. This systematic uncertainty is made up from several contributions added in quadrature. The finite size of the simulated sample limits the accuracy of the correction.
derived from them (2.1–4.7% relative uncertainty). Only a subset of running conditions are simulated even though data from all data-taking conditions are used (2.7–5.0% relative uncertainty). Possible mismodelling of the data leads to a 2.7–7.7% relative systematic uncertainty. This uncertainty includes any imperfect BDT response estimation and also the efficiency variation over $q^2$ to account for different form-factor models and comparisons of angular distributions resulting from phase-space and spin-dependent decay of the di-muon system. The PID response in the simulation is corrected by resampling the calibration-channel data, and through the use of alternate sampling distributions, relative uncertainties of up to 3.1% are estimated. The trigger efficiencies are also corrected for differences between simulation and data (3.2–7.0% relative uncertainty). The third systematic contribution is a consequence of the low number of candidate events in the final fit requiring that most PDF parameters must be fixed from fits to simulation to ensure fit stability. Subsequently, these parameters are varied systematically by ±1 standard deviation and the fits repeated on data to assess the uncertainty from these shape parameters. The fits are also repeated with an alternative combinatorial background model (0.3–1.9% total relative uncertainty). The final systematic contribution is associated with the three modes where the corresponding normalisation channel has a different number of tracks from the signal. This includes the uncertainty in tracking efficiencies and differences in the interactions of charged pions and kaons with the detector material, added in quadrature (3.0% combined relative uncertainty).

6 Results

As seen in figure 4, a clear signal is observed in the $B_c^+ \to D_s^+ J/\psi$ decay mode, and the corresponding branching fraction is measured to be

$$\frac{f_c}{f_u} \cdot B(B_c^+ \to D_s^+ J/\psi) = (1.63 \pm 0.15 \pm 0.13) \times 10^{-5},$$

where the first uncertainty is statistical and the second is systematic. This improves the precision of the first measurement of this decay, previously reported in ref. [14]. As there is absence of signal in all other decay modes (less than 2 standard deviations), the CLs method [58] is used to evaluate the compatibility of the observed invariant-mass distributions for each search mode with signal-and-background and background-only hypotheses. The distributions of $p$-values as a function of assumed branching fraction are used to derive upper limits on the branching fractions at the 90% and 95% CL, which are given in table 4.

Two measurements relating to $B_c^+ \to D_s^+ J/\psi$ decays are also extracted from the fit shown in figure 4. The ratio of yields $B_c^+ \to D_s^{*+} J/\psi$ to $B_c^+ \to D_s^+ J/\psi$ decays is measured to be

$$\mathcal{R}_{D_s^{*+}/D_s^+} = \frac{B(B_c^+ \to D_s^{*+} J/\psi)}{B(B_c^+ \to D_s^+ J/\psi)} = 1.91 \pm 0.20 \pm 0.07,$$

where the systematic error is associated with the assumption that the efficiencies for the partially-reconstructed $B_c^+ \to D_s^{*+} J/\psi$ decay and the $B_c^+ \to D_s^+ J/\psi$ decay are equal. The ratio of the number of $B_c^+ \to D_s^{*+} J/\psi$ decays described by the $A^{\pm\pm}$ helicity amplitude compared to the total number is

$$\Gamma_{\pm\pm}/\Gamma_{\text{tot}} = \frac{N_{A^{\pm\pm}}}{N_{A^{\pm\pm}} + N_{A^{00}}} = 0.50 \pm 0.11 \pm 0.05,$$
Table 4. Upper limits at the 90% and 95% confidence levels for $B \to D\mu^+\mu^-$ and $B \to DJ/\psi$ decays.

<table>
<thead>
<tr>
<th>Branching fraction</th>
<th>Upper limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B (B^0 \to D^0\mu^+\mu^-)$</td>
<td>$4.0 \times 10^{-8}$</td>
</tr>
<tr>
<td>$B (B^+ \to D_s^+\mu^+\mu^-)$</td>
<td>$2.4 \times 10^{-8}$</td>
</tr>
<tr>
<td>$f_c/f_u \cdot B (B^{+}_c \to D^{+}\mu^+\mu^-)$</td>
<td>$7.5 \times 10^{-8}$</td>
</tr>
<tr>
<td>$f_c/f_u \cdot B (B^{+}_s \to D^{+}\mu^+\mu^-)$</td>
<td>$7.5 \times 10^{-8}$</td>
</tr>
</tbody>
</table>

where the uncertainty is dominantly statistical and the systematic term results from considering alternative fit models. The values of $\mathcal{R}_{D^{+}\mu^+}$ and $\Gamma_{\pm}/\Gamma_{\text{tot}}$ supersede the previous LHCb results [14], and also are in agreement with the ATLAS measurements [16]. The ratio $\Gamma_{\pm}/\Gamma_{\text{tot}}$ is consistent with the naive expectation of 2/3 from spin-counting considerations.

7 Conclusions

A search for four rare $B \to D\mu^+\mu^-$ decays is performed using proton-proton collision data collected by the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV, corresponding to a total integrated luminosity of 9 fb$^{-1}$. No new signals are observed and upper limits are set. Additional limits are determined when the muon pair originates from a $J/\psi \to \mu^+\mu^-$ decay. All upper limits are either an improvement on existing results or are the first limits set by any experiment. Improved measurements of the previously observed $B^+ \to D_s^+J/\psi$ decay are also made.

In the future, it is expected that a new measurement of these modes with a 50 fb$^{-1}$ dataset from the upgraded LHCb experiment, later increasing to 300 fb$^{-1}$ with Upgrade II, will probe down towards the 10$^{-9}$ level for the non-resonant modes, assuming luminosity scaling. These sensitivities will approach SM expectations, in particular for the $B^0$ modes studied in the present paper.

Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (Netherlands); MNiSW and NCN (Poland); MCID/IFA (Romania); MICINN (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); DOE NP and NSF (U.S.A.). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY.
(Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), Polish WLCG (Poland) and NERSC (U.S.A.). 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 ARC and ARDC (Australia); Minciencias (Colombia); AvH Foundation (Germany); EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union); A*MIDEX, ANR, IPhU and Labex P2IO, and Région Auvergne-Rhône-Alpes (France); Key Research Program of Frontier Sciences of CAS, CAS PIFI, CAS CCEPP, Fundamental Research Funds for the Central Universities, and Sci. & Tech. Program of Guangzhou (China); GVA, XuntaGal, GENCAT and Prog. Atracción Talento, CM (Spain); SRC (Sweden); the Leverhulme Trust, the Royal Society and UKRI (United Kingdom).

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] LHCb collaboration, Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb$^{-1}$ of integrated luminosity, JHEP 02 (2016) 104 [arXiv:1512.04442] [SPIRE].

[2] LHCb collaboration, Strong constraints on the $b \to s\gamma$ photon polarisation from $B^0 \to K^{*0}e^+e^-$ decays, JHEP 12 (2020) 081 [arXiv:2010.06011] [SPIRE].


[8] C.S. Kim, R.-H. Li and Y. Li, Study of $\bar{B}^0 \to D^0\mu^+\mu^-$ Decay in Perturbative QCD Approach, JHEP 10 (2011) 152 [arXiv:1106.2711] [SPIRE].


[10] BaBar collaboration, Search for the rare decay $\bar{B}^0 \to D^{*0}\gamma$, Phys. Rev. D 72 (2005) 051106 [hep-ex/0506070] [SPIRE].


[13] LHCb collaboration, Measurement of $B^+_s$ production in proton-proton collisions at $\sqrt{s} = 8$ TeV, 

[14] LHCb collaboration, Observation of $B^+_s \rightarrow J/\psi D^+_s$ and $B^+_s \rightarrow J/\psi D^{*+}_s$ decays, Phys. Rev. D 87
(2013) 112012 [Addendum ibid. 89 (2014) 019901] [arXiv:1304.4530] [insPIRE].

[15] ATLAS collaboration, Study of the $B^+_s \rightarrow J/\psi D^+_s$ and $B^+_s \rightarrow J/\psi D^{*+}_s$ decays with the ATLAS 

[16] ATLAS collaboration, Study of $B^+_s \rightarrow J/\psi D^+_s$ and $B^+_s \rightarrow J/\psi D^{*+}_s$ decays in $pp$ collisions at
$\sqrt{s} = 13$ TeV with the ATLAS detector, JHEP 08 (2022) 087 [arXiv:2203.01808] [insPIRE].

[arXiv:1701.01856] [insPIRE].

[18] LHCb collaboration, Branching fraction and CP asymmetry of the decays $B^+ \rightarrow K^0\pi^+$ and

[19] LHCb collaboration, Search for $B^+_s$ decays to the $p\bar{p}\pi^+$ final state, Phys. Lett. B 759 (2016) 313
[arXiv:1603.07037] [insPIRE].

[20] P. Maji, S. Biswas, P. Nayek and S. Sahoo, Charged Higgs contribution on $B_c \rightarrow (D_s,D_s^*)l^+l^-$, 

[21] LHCb collaboration, The LHCb Detector at the LHC, 2008 JINST 3 S08005 [insPIRE].

[arXiv:1412.6352] [insPIRE].

[arXiv:1405.7808] [insPIRE].

[arXiv:1311.3893] [insPIRE].

JINST 12 P11016 [arXiv:1708.00819] [insPIRE].

(2013) 2431 [arXiv:1211.6759] [insPIRE].

[27] A.A. Alves Jr. et al., Performance of the LHCb muon system, 2013 JINST 8 P02022
[arXiv:1211.1346] [insPIRE].

[arXiv:1211.3055] [insPIRE].

[29] R. Aaij et al., Design and performance of the LHCb trigger and full real-time reconstruction in 
Run 2 of the LHC, 2019 JINST 14 P04013 [arXiv:1812.10790] [insPIRE].

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


026 [hep-ph/0603175] [insPIRE].

[33] I. Belyaev et al., Handling of the generation of primary events in Gauss, the LHCb simulation 


The LHCb collaboration

26 INFN Sezione di Milano-Bicocca, Milano, Italy
27 INFN Sezione di Cagliari, Monserrato, Italy
28 Università degli Studi di Padova, Università e INFN, Padova, Padova, Italy
29 INFN Sezione di Pisa, Pisa, Italy
30 INFN Sezione di Roma La Sapienza, Roma, Italy
31 INFN Sezione di Roma Tor Vergata, Roma, Italy
32 Nikhef National Institute for Subatomic Physics, Amsterdam, Netherlands
33 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, Netherlands
34 AGH - University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
35 Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
36 National Center for Nuclear Research (NCBJ), Warsaw, Poland
37 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
38 Affiliated with an institute covered by a cooperation agreement with CERN
39 ICCUB, Universitat de Barcelona, Barcelona, Spain
40 Instituto Galego de Física de Altas Enerxías (IGFAE), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
41 Instituto de Física Corpuscular, Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain
42 European Organization for Nuclear Research (CERN), Geneva, Switzerland
43 Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
44 Physik-Institut, Universität Zürich, Zürich, Switzerland
45 NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
46 Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
47 University of Birmingham, Birmingham, United Kingdom
48 H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
49 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
50 Department of Physics, University of Warwick, Coventry, United Kingdom
51 STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
52 School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
53 School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
54 Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
55 Imperial College London, London, United Kingdom
56 Department of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
57 Department of Physics, University of Oxford, Oxford, United Kingdom
58 Massachusetts Institute of Technology, Cambridge, MA, United States
59 University of Cincinnati, Cincinnati, OH, United States
60 University of Maryland, College Park, MD, United States
61 Los Alamos National Laboratory (LANL), Los Alamos, NM, United States
62 Syracuse University, Syracuse, NY, United States
63 School of Physics and Astronomy, Monash University, Melbourne, Australia, associated to 50
64 Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to 2
65 Physics and Micro Electronic College, Hunan University, Changsha City, China, associated to 7
66 Guangdong Provincial Key Laboratory of Nuclear Science, Guangdong-Hong Kong Joint Laboratory of Quantum Matter, Institute of Quantum Matter, South China Normal University, Guangzhou, China, associated to 3
67 Lanzhou University, Lanzhou, China, associated to 4
68 School of Physics and Technology, Wuhan University, Wuhan, China, associated to 3
69 Departamento de Física, Universidad Nacional de Colombia, Bogota, Colombia, associated to 13
70 Universität Bonn - Helmholtz-Institut für Strahlen und Kernphysik, Bonn, Germany, associated to 17
71 Eotvos Lorand University, Budapest, Hungary, associated to 42
72 INFN Sezione di Perugia, Perugia, Italy, associated to 21
73 Van Swinderen Institute, University of Groningen, Groningen, Netherlands, associated to 32
74 Universiteit Maastricht, Maastricht, Netherlands, associated to 32
Tadeusz Kosciuszko Cracow University of Technology, Cracow, Poland, associated to
DS4DS, La Salle, Universitat Ramon Llull, Barcelona, Spain, associated to
Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden, associated to
University of Michigan, Ann Arbor, MI, United States, associated to
Department de Physique Nucleaire (SPhN), Gif-Sur-Yvette, France

Universidade de Brasília, Brasília, Brazil
Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
Central South U., Changsha, China
Hangzhou Institute for Advanced Study, UCAS, Hangzhou, China
Excellence Cluster ORIGINS, Munich, Germany
Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras
Università di Bari, Bari, Italy
Università di Bologna, Bologna, Italy
Università di Cagliari, Cagliari, Italy
Università di Ferrara, Ferrara, Italy
Università di Firenze, Firenze, Italy
Università di Genova, Genova, Italy
Università degli Studi di Milano, Milano, Italy
Università di Milano Bicocca, Milano, Italy
Università di Padova, Padova, Italy
Università di Perugia, Perugia, Italy
Scuola Normale Superiore, Pisa, Italy
Università di Pisa, Pisa, Italy
Università della Basilicata, Potenza, Italy
Università di Roma Tor Vergata, Roma, Italy
Università di Urbino, Urbino, Italy
Universidad de Alcalá, Alcalá de Henares, Spain
† Deceased