Prospects for studies of $D^{*0} \rightarrow \mu^+ \mu^-$ and $B^{*0}_{(s)} \rightarrow \mu^+ \mu^-$ decays

F. Abudinén, T. Blake, U. Egede, T. Gershon

1 Department of Physics, University of Warwick, Coventry, UK
2 School of Physics and Astronomy, Monash University, Melbourne, Australia

Received: 11 February 2022 / Accepted: 22 April 2022 © The Author(s) 2022

Abstract Weak decays of the vector $D^{*0}$ and $B^{*0}_{(s)}$ mesons to the $\mu^+ \mu^-$ final state provide novel potential to test the Standard Model of particle physics. Such processes have extremely small branching fractions as the vector mesons are able to decay through electromagnetic and (for the $D^{*0}$ meson) strong interactions. Nonetheless, the production of copious quantities of these particles in LHC collisions, and the ability to exploit experimental techniques that can suppress background to low levels, provides good potential to reach interesting sensitivity. The possibility to reconstruct these processes as part of the decay chain of $B^-$ or $B^+_s$ mesons appears particularly attractive due to the clean experimental signature of the displaced vertex. Indeed, published LHCb data on $B^- \rightarrow \pi^- \mu^+ \mu^-$ decays already implies a stringent limit on the branching fraction of $D^{*0} \rightarrow \mu^+ \mu^-$. Estimates are made on the achievable sensitivity to $D^{*0} \rightarrow \mu^+ \mu^-$ and $B^{*0}_{(s)} \rightarrow \mu^+ \mu^-$ decays with the LHCb experiment.

1 Introduction

Weak decays of heavy-flavoured hadrons provide a range of methods to test the Standard Model (SM) of particle physics. In particular, many such transitions are suppressed by particular features of the SM such as the GIM mechanism [1] and the CKM quark mixing matrix [2,3]. As a consequence, the SM predicts a distinctive pattern of decay rates to various different final states, which may be modified by contributions from physics beyond the SM. Experimental and theoretical investigations in this area are therefore a priority in the field.

Until now, experimental studies of weak decays have been almost completely limited to the ground-state hadrons; considering neutral heavy-flavoured mesons, these are the pseudoscalar $D^0$, $B^0$ and $B^0_s$ states. The leptonic decays, to the $\ell^+ \ell^-$ final state where $\ell = e, \mu$ and (for $B^{*0}_{(s)}$ decays) $\tau$, have branching fractions that are suppressed by the chiral structure of the SM weak interaction, and that can be predicted with small theoretical uncertainties [4]. These features together make them highly sensitive to potential contributions from physics beyond the SM. Intense activity on the $B^0 \rightarrow \mu^+ \mu^-$ channels has resulted in the observation of the $B^0 \rightarrow \mu^+ \mu^-$ decay by the LHCb, CMS and ATLAS experiments, and a combined limit on the $B^0 \rightarrow \mu^+ \mu^-$ branching fraction that approaches its SM value [5–10]. The experimental limits on $B (D^0 \rightarrow e^+ e^-)$ [11], $B (D^0 \rightarrow \mu^+ \mu^-)$ [12], $B (B^{*0}_{(s)} \rightarrow e^+ e^-)$ [13] and $B (B^{*0}_{(s)} \rightarrow \tau^+ \tau^-)$ [14] are still several orders of magnitude above their SM predictions.

It is also possible to consider weak decays of the excited counterparts of the pseudoscalar mesons, the vector $D^{*0}$, $B^{*0}$ and $B^{*0}_{(s)}$ resonances. In contrast to the situation for pseudoscalar mesons, the leptonic vector meson decays have no chiral suppression. Consequently the decay widths for each of the $\ell^+ \ell^-$ final states are expected to be the same, in the SM, up to effects related to the lepton mass (e.g., the available phase space), and will be larger compared to those for the pseudoscalar decays. However, the vector mesons can also decay via electromagnetic and (for the $D^{*0}$ meson) strong transitions, which have widths many orders of magnitude larger than those for the weak decays. Therefore, the branching fractions of the weak decays are suppressed to what would usually be considered unobservably small levels. As an illustration, one can compare the width of the $D^{*+}$ vector state, $\Gamma_{D^{*+}} = 83$ keV [15], with that of its $D^+$ pseudoscalar counterpart, $\Gamma_{D^+} = h/\tau_{D^+} \approx h/(1.0 \text{ ps}) = 0.7$ meV [16], a difference of over 8 orders of magnitude. This is in the ballpark that one would naively expect, given that the weak decays are suppressed by the Fermi constant, but have more phase space available compared to the electromagnetic and strong transitions.
transitions in which flavour is conserved. Given that weak decays lead to a plethora of different final states, the branching fractions for even the most favoured are unlikely to be above \(10^{-9}\), unless large enhancement factors due to physics beyond the SM are at play. For the \(D_s^0 \to \ell^+ \ell^-\) decays, further suppression by the GIM mechanism [1] reduces the Standard Model prediction for the branching fractions to the \(\mathcal{O}(10^{-20})\) level, while for \(B_s^0 \to \ell^+ \ell^-\) decays the prediction is in the range \((0.7\text{--}2.2) \times 10^{-11}\) [17, 18].

Hints of physics beyond the SM in \(B\) meson decays have, nevertheless, prompted theoretical activity on weak decays of excited heavy-flavoured states. Collectively referred to as the flavour anomalies, these hints include tensions between SM predictions and experimental measurements in branching fractions and angular observables in decays mediated by \(b \to s \ell^+ \ell^-\) transitions, including observables that are sensitive to violations of lepton universality (see Ref. [19] for a recent review). The leptonic \(D_s^0\), \(B_s^0\) and \(B_s^0\) decays are of most interest, since these are theoretically cleanest and, for the \(B_s^0\) case, are sensitive to the same operators which could be responsible for the flavour anomalies.\(^2\) These decays have been considered as a potential probe of physics beyond the SM in Refs. [17, 18], with further investigations of the impact of particular extensions of the SM considered in Refs. [20–24].

There is extra motivation to search for leptonic \(D_s^0\) decays. The relevant operators for the \(D_s^0\) and \(B_s^0\) decays are already constrained from measurements of pseudoscalar \(B_{(s)} \to \ell^+ \ell^-\) and \(h \ell^+ \ell^-\) transitions (where \(h\) is a hadron), and can be used to set limits on \(\mathcal{B}\left(B_{(s)} \to \ell^+ \ell^-\right)\) that are below the current experimentally achievable sensitivity [18]. The interpretation of results from \(D_s^0 \to \ell^+ \ell^-\) and \(D_{(s)} \to h \ell^+ \ell^-\) decays, on the other hand, is challenging due to long-distance effects [25–28]. Correspondingly, constraints on the relevant operators are weaker, and the possibility of a large enhancement to \(\mathcal{B}\left(D_s^0 \to \ell^+ \ell^-\right)\) from physics beyond the Standard Model cannot be ruled out. Thus, searches for \(D_s^0 \to \ell^+ \ell^-\) decays could in principle provide an important complementary approach to probe the operators involved, if sufficient experimental precision can be obtained.

Since there is no suppression of the coupling of the heavy flavoured vector resonances to dielectron, compared to dimuon, states, it may be attractive experimentally to search for these interactions through production in \(e^+ e^-\) collisions, rather than in decay processes [17]. A search for the \(e^+ e^- \to D_s^0(2007)^0\) process has been carried out by the CMD-3 collaboration, resulting in an upper limit being set, \(\mathcal{B}\left(D_s^0 \to e^+ e^-\right) < 1.7 \times 10^{-6}\) at 90\% confidence level [29]. While it is likely that this result can be significantly improved through analyses of larger data samples with better background suppression, the limit is not at such a stringent level to suggest that other approaches to study these processes would be futile.

In particular, the copious production of heavy flavoured hadrons at the LHC makes it worthwhile to consider what sensitivity might be achievable with current and future data samples. As the dimuon signature allows for effective background suppression in hadron collider experiments, \(D_s^0\), \(B_s^0\) and \(B_s^0\) decays to the \(\mu^+ \mu^-\) final state are the most amenable. There are several possible avenues to investigate these processes using LHC data. In particular, both “prompt” and “displaced” production can be considered, where in the former the heavy flavoured state is produced at the primary vertex of the proton-proton collision while in the latter the signal hadron originates from the decay of another weakly decaying particle some distance from the primary vertex. Specifically, \(D_s^0\) mesons are produced at high rates in \(b\) hadron decays, and \(B_s^0\) mesons can be produced in \(B^+_c\) meson, and potentially in other doubly-heavy hadron, decays. While prompt production has the highest rate, the large numbers of tracks originating from the proton-proton collision vertices lead to large backgrounds that limit the sensitivity of any rare decay search. In displaced production it is unlikely to find random tracks appearing to come from the vertex position, so long as the vertex resolution is sufficient. If in addition the signal is part of an exclusive decay process, further constraints can be applied to suppress background. Displaced production therefore appears to provide the most promising approach.

In the main part of this paper, the potential sensitivity of the LHCb experiment to \(D_s^0\), \(B_s^0\) and \(B_s^0\) decays to the \(\mu^+ \mu^-\) final state, using displaced production in exclusive final states, is investigated. Possibilities with prompt production, and with inclusive and semi-inclusive search approaches with displaced production, are considered for completeness in Appendices A and B, respectively. In principle the ATLAS and CMS experiments could also make competitive measurements, but until now they have fewer relevant measurements making extrapolations difficult. Moreover, as their vertexing and charged hadron identification capability is not as good as that of LHCb, it is expected that they will suffer from larger backgrounds.

2 \(B^- \to D_s^0 \pi^- \to \mu^+ \mu^- \pi^-\) decays

The decay chain \(B^- \to D_s^0 \pi^- \to \mu^+ \mu^- \pi^-\) not only provides an excellent illustrative example, it also allows an estimated upper limit on the \(D_s^0 \to \mu^+ \mu^-\) branching fraction to be obtained from published data. In particular, LHCb has studied the \(B^- \to \mu^+ \mu^- \pi^-\) decay using a data sample corresponding to an integrated lumi-

\(^2\) The inclusion of charge conjugate processes is implied throughout this paper.
nosity of 3.0 fb$^{-1}$, collected at centre-of-mass energies of 7 and 8 TeV [30]. The measured differential branching fraction $dB/dq^2$, where $q^2$ is the square of the dimuon invariant mass, is shown in Fig. 1. An upper limit on $B(D^{\ast 0} \to \mu^+ \mu^-)$ can be estimated by assuming conservatively that the $B^- \to D^{\ast 0} \pi^- \to \mu^+ \mu^- \pi^-$ decay contributes not more than half of the $B^- \to \mu^+ \mu^- \pi^-$ signal in the two bins either side of $q^2 = 4 \text{GeV}^2/c^4$, i.e. a branching fraction of $B(B^- \to \mu^+ \mu^-) \approx 1.5 \times 10^{-9}$. Then, using the world average value of $B(B^- \to D^{\ast 0} \pi^-) [31]$, one obtains

$$B(D^{\ast 0} \to \mu^+ \mu^-) = \frac{B(B^- \to \mu^+ \mu^-)_{D^{\ast 0} \pi^-}}{B(B^- \to D^{\ast 0} \pi^-)} \approx 3 \times 10^{-7}.$$  

This limit, while already more stringent than the result on $D^{\ast 0} \to e^+ e^-$ from the CMD-3 collaboration [29], could most likely be improved by at least an order of magnitude by a dedicated LHCb analysis. In particular, the expected mass resolution, which we expect to be around 5 MeV/$c^2$, is much better than can be obtained from the coarse $q^2$ binning of Fig. 1. Moreover, LHCb has already collected a significantly larger data sample than was analysed in Ref. [30], and a somewhat higher selection efficiency could be anticipated in a dedicated analysis. It is therefore of interest to ask what sensitivity might ultimately be achieved by LHCb in such a search.

In LHCb experiments, it is advantageous to measure branching fractions relative to those of appropriate normalisation channels. It is anticipated that the $B^- \to J/\psi K^-$ decay will be used for this purpose and hence the signal yield is converted to a measurement of the branching fraction as

$$B(D^{\ast 0} \to \mu^+ \mu^-) = \frac{N_{B^- \to D^{\ast 0} \pi^-} \epsilon_{B^- \to J/\psi K^-} B(B^- \to J/\psi K^-)}{N_{B^- \to J/\psi K^-} \epsilon_{B^- \to D^{\ast 0} \pi^-} B(B^- \to D^{\ast 0} \pi^-)} B(J/\psi \to \mu^+ \mu^-),$$  

where $\alpha_{D^{\ast 0} \to \mu^+ \mu^-}$ is the single-event sensitivity, which corresponds to the branching fraction at which one signal event is expected in the data sample. In Eq. (1), $N$ and $\epsilon$ represent the yield and the efficiency for the decay indicated in the subscript, where the reconstruction of the $D^{\ast 0}$ or $J/\psi$ vector meson in the $\mu^+ \mu^-$ final state is implied.

The single-event sensitivity is estimated from the yield of $B^- \to J/\psi K^-$ decays, scaled from measurements with existing data [30] appropriately to each integrated luminosity value, and the branching fractions $B(B^- \to J/\psi K^-)$, $B(J/\psi \to \mu^+ \mu^-)$ and $B(B^- \to D^{\ast 0} \pi^-) [31]$. It is assumed that $\epsilon_{B^- \to J/\psi K^-} \approx \epsilon_{B^- \to D^{\ast 0} \pi^-}$, since the final state is the same (apart from $K \leftrightarrow \pi$) and the efficiency varies slowly with $q^2$. The single-event sensitivity is shown in Fig. 2 as a function of the data set size.

The achievable precision depends not only on the single-event sensitivity but also on the uncertainty on the signal yield, which is often limited by the background level. To investigate how the achievable limit may scale with integrated luminosity, pseudoexperiments are generated under a background-only hypothesis. Three background components are considered: combinatorial background from random combinations of tracks from two or more decays; background from nonresonant $B^- \to \mu^+ \mu^- \pi^-$ decays; and background from $B^- \to K^- \mu^+ \mu^-$ decays, where the $K^-$ meson is mistakenly identified as a $\pi^-$. The combinatorial background is assumed to be uniformly distributed in the dimuon mass, $m(\mu^+ \mu^-)$. The backgrounds from $B^- \to K^- \mu^+ \mu^-$ and $B^- \to \mu^+ \mu^- \pi^-$ decays are assumed to be uniform in $q^2 = m^2(\mu^+ \mu^-)$, consistent with both the expected [35] and the observed [36] shape of the differential $B^- \to K^- \mu^+ \mu^-$ decay rate in the $q^2$ range of interest. The backgrounds are generated over the interval $2 < q^2 < 6 \text{GeV}^2$, which covers the two bins of the LHCb $B^- \to \mu^+ \mu^- \pi^-$ analysis closest to the $D^{\ast 0}$ mass. The level of each of the three backgrounds is taken from Ref. [30] and is scaled to the considered integrated luminosity. Proton-proton collisions during LHC Run 2 (2015–18) were at 13 TeV centre-of-mass energy and those in future run periods are expected to be at similar or slightly higher energies. When extrapolating to future data sample sizes it is assumed that the $b\bar{b}$ production cross-section scales linearly with centre-of-mass energy [37].

Limits are set at 90% confidence level for each pseudoexperiment using the method described in Ref. [38] (as imple-
The rate of reduction of the expected limit as the sample size shown in Fig. 3. An upper limit at the level of 10^{−10} B\text{\scriptsize{kinematic fit to the}} regions are defined, a signal region uncertainty on the single-event sensitivity into account. Two mass regions are defined, a signal region. 

\text{TRolke} commented in the TRolke class in ROOT, taking the uncertainty on the single-event sensitivity into account. Two mass regions are defined, a signal region \pm 10\text{MeV}/c^2 around the known D^{*0} mass, m_{D^{*0}}, comprising a mixture of signal and background candidates and two sideband regions \(-35 < m(\mu^+\mu^-) - m_{D^{*0}} < -15\text{MeV}/c^2\) and \(15 < m(\mu^+\mu^-) - m_{D^{*0}} < 35\text{MeV}/c^2\), comprising only background candidates. The sideband regions are used to estimate the background in the signal region. The width of the signal region is taken as \pm 2 times the expected \(m(\mu^+\mu^-)\) resolution of \(\sim 5\text{MeV}/c^2\) [39], which is minimised by applying a kinematic fit to the \(B^-\rightarrow \mu^+\mu^-\pi^-\) process in which the \(B^-\) mass is constrained to its known value.

The results of this pseudoexperiment-based study are shown in Fig. 3. An upper limit at the level of \(10^{-8}\) appears to be possible with the current LHCb data set. This can be further improved to \(\mathcal{O}(10^{-9})\) with the total sample of \(300\text{fb}^{-1}\) anticipated with future LHCb upgrades [40,41]. The rate of reduction of the expected limit as the sample size increases slows markedly at around \(20\text{fb}^{-1}\) as background expected, with size depending on the relative phase between the two interfering amplitudes. Even though the D^{*0} width is below the experimental resolution, this effect could in principle be used, once the sample size is sufficiently large, to obtain a better limit than indicated in Fig. 3. Indeed, the LHCb collaboration has already demonstrated the possibility to measure a similar interference effect between \(B^-\rightarrow (J/\psi\rightarrow \mu^+\mu^-)K^-\) and nonresonant \(B^-\rightarrow K^-\mu^+\mu^-\) decays [39]. The results of this analysis allow the maximum size of the \(B^-\rightarrow D^{*0}K^-\) contribution to \(B^-\rightarrow K^-\mu^+\mu^-\) decays to be estimated, and hence a limit on \(\mathcal{B}(D^{*0}\rightarrow \mu^+\mu^-)\) can be derived. Assuming the yield of \(B^-\rightarrow \mu^+\mu^-\) \(K^-\) decays is not more than \(2 \times \sqrt{N}\), where \(N \sim 50\) is the yield in the \(m(\mu^+\mu^-)\) bin around \(m_{D^{*0}}\), taken from Fig. 3 of Ref. [39], and normalising to the \(B^-\rightarrow (J/\psi\rightarrow \mu^+\mu^-)K^-\) contribution which has a yield of \(\sim 900\text{,}000\) (around 90\% of the total \(B^-\rightarrow K^-\mu^+\mu^-\) yield), and with an efficiency ratio of 0.85 (from Fig. 2 of Ref. [39]), then with known branching fractions [31,42] one obtains

\[
\mathcal{B}(D^{*0}\rightarrow \mu^+\mu^-) = \frac{N_{B^-\rightarrow D^{*0}K^-}}{N_{B^-\rightarrow J/\psi K^-}} \frac{\epsilon_{B^-\rightarrow J/\psi K^-}}{\epsilon_{B^-\rightarrow D^{*0}K^-}} \frac{\mathcal{B}(B^-\rightarrow J/\psi K^-)}{\mathcal{B}(B^-\rightarrow D^{*0}K^-)} \mathcal{B}(J/\psi\rightarrow \mu^+\mu^-) \leq 3 \times 10^{-6}.
\]

starts to become limiting. While combinatorial background and misidentified \(B^-\rightarrow K^-\mu^+\mu^-\) decays can be further reduced with tighter selection requirements, the contribution from nonresonant \(B^-\rightarrow \mu^+\mu^-\pi^-\) decays is irreducible in this approach.

The studies resulting in Fig. 3 neglect interference effects between the \(B^-\rightarrow (D^{*0}\rightarrow \mu^+\mu^-)\pi^-\) signal and nonresonant \(B^-\rightarrow \pi^-\mu^+\mu^-\) decays, which is reasonable given the narrow width of the \(D^{*0}\) meson. Nonetheless, since the \(D^{*0}\) quantum numbers are the same as part of the nonresonant contribution, a net interference effect is

As expected, this mode is not as sensitive as \(B^-\rightarrow (D^{*0}\rightarrow \mu^+\mu^-)\pi^-\) due to the smaller \(B^-\) branching fraction and the larger background from nonresonant \(B^-\rightarrow K^-\mu^+\mu^-\) decays.

3. \(B^+_c \rightarrow B^{*0}_c\pi^+ \rightarrow \mu^+\mu^-\pi^+\) decays

A similar strategy can in principle also be employed to search for \(B^+_c \rightarrow B^{*0}_c\pi^+ \rightarrow \mu^+\mu^-\pi^+\) decays, where the \(B^+_c \rightarrow J/\psi\pi^+\) decay can be used for normalisation. In this case,
however, some of the necessary ingredients to convert the signal yield to the $B^{0}_{(s)} \rightarrow \mu^{+}\mu^{-}$ branching fraction are not currently available. Specifically, the equivalent expression to Eq. (1) is

$$E \left( B^{0}_{(s)} \rightarrow \mu^{+} \mu^{-} \right) = \frac{N_{B^{0}_{(s)} \rightarrow B^{0}_{(s)} \pi^{+}}}{N_{B^{0}_{(s)} \rightarrow J/\psi \pi^{+}}} \frac{E_{B^{0}_{(s)} \rightarrow J/\psi \pi^{+}} \mathbb{E} \left( B^{+}_{c} \rightarrow J/\psi \pi^{+} \right)}{E_{B^{0}_{(s)} \rightarrow J/\psi \pi^{+}}} \mathbb{E} \left( B^{+}_{c} \rightarrow B^{0}_{(s)} \pi^{+} \right) \mathbb{E} \left( J/\psi \rightarrow \mu^{+} \mu^{-} \right),$$

(3)

and the branching fractions for $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$ and $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$ decays are present unmeasured. It seems possible, however, that at least the former decay could be observed with existing data. A previous LHCb analysis has observed $B^{+}_{s} \rightarrow B^{0}_{s} \pi^{+}$ with 3 fb$^{-1}$ of data [43] and this channel has also been used with the full current data sample of 9 fb$^{-1}$ to measure the $B^{+}_{s}$ mass [44]. The signal for $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$ would be expected in the same spectrum as a satellite peak shifted below the $B^{+}_{s}$ mass by approximately $m_{B^{0}_{s}} - m_{B^{0}}$ due to the soft photon that is not included in the reconstructed candidate. Similar “partial reconstruction” techniques have been used by LHCb to observe $B^{+}_{s} \rightarrow J/\psi D^{*+}_{s}$ [45] and $B^{-} \rightarrow D^{0} \pi^{-}$ [46] decays, where the soft neutral particles from the $D^{*+}$ and $D^{0}$ decays are not included in the reconstructed candidate. Assuming that $E(B^{+}_{s} \rightarrow B^{0}_{s} \pi^{+})$ is not much smaller than $E(B^{+}_{s} \rightarrow B^{0}_{s} \pi^{+})$, as is expected theoretically [47–51], it should be possible to observe partially reconstructed $B^{+}_{s} \rightarrow B^{0}_{s} \pi^{+}$ decays, and measure the corresponding branching fraction, in the existing LHCb data sample. The $B^{+}_{s} \rightarrow B^{0}_{s} \pi^{+}$ and $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$ decays could be searched for with a similar technique although, since these transitions are Cabibbo-suppressed relative to $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$, larger data samples may be required to observe them. A further, albeit, minor complication is that the $m_{B^{0}_{s}} - m_{B^{0}}$ mass difference is as-yet unmeasured, although it can be predicted rather reliably from the $m_{B^{+}_{s}} - m_{B^{+}}$ mass difference invoking isospin symmetry. (This can also be interpreted as a further motivation for the analysis since a first measurement of $m_{B^{0}_{s}}$ may be possible.)

Another apparent problem is that even for $B^{+}_{c}$ decay modes which have been observed, there are currently no measurements of absolute branching fractions. Rather, only the product of the branching fraction with a ratio of production fractions is known. However, the $B^{+}_{c}$ meson production rate cancels out in Eq. (3), allowing some simplification. With $E(B^{+}_{c} \rightarrow J/\psi \pi^{+})$, which appears in the numerator, measured relative to $B^{+} \rightarrow J/\psi K^{+}$ and $E(B^{+}_{c} \rightarrow B^{0}_{(s)} \pi^{+})$, which appears in the denominator, measured relative to $B^{0}_{(s)} \rightarrow J/\psi \phi$, it is only necessary to know the relative production rate of $B^{+}$ and $B^{0}_{(s)}$ mesons, which has been measured precisely [52]. Taking the yield of $25.2 \times 10^{3}$ of efficiencies in Eq. (3) is unity, a single-event sensitivity with 9 fb$^{-1}$ of $\alpha_{B^{0}_{(s)} \rightarrow \mu^{+}\mu^{-}} \approx 5.5 \times 10^{-3}$ is obtained. Further assuming the $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$ decay has a Cabibbo-suppression factor of $|V_{cd}/V_{cs}|^{2} \approx 5\%$ relative to $B^{+}_{c} \rightarrow B^{0}_{s} \pi^{+}$, the corresponding single-event sensitivity is found to be $\alpha_{B^{0}_{s} \rightarrow \mu^{+}\mu^{-}} \approx 1.1 \times 10^{-6}$. Scaling to a data sample of 300 fb$^{-1}$, the achievable single-event-sensitivities could reach $\sim 1.3 \times 10^{-9}$ and $2.7 \times 10^{-8}$ for $B^{0}_{(s)} \rightarrow \mu^{+}\mu^{-}$ and $B^{0}_{s} \rightarrow \mu^{+}\mu^{-}$ decays, respectively.

Interpretation of these single-event sensitivities must be made with care, however, since they assume a selection efficiency comparable to that for $B^{+}_{c} \rightarrow J/\psi \pi^{+}$ decays in Ref. [44]. In practice the selection requirements will be optimised to account for the level of background in the signal region for the $B^{+}_{c} \rightarrow B^{0}_{(s)} \pi^{+}$ search. Since there is no published search for $B^{+}_{c} \rightarrow \mu^{+}\mu^{-} \pi^{+}$ decays outside the mass regions where the dimuon pair originates from a $J/\psi$ or $\psi$(2S) charmonium state, it is hard to judge what the optimal requirements are likely to be. Moreover, a contribution from nonresonant $B^{+}_{c} \rightarrow \mu^{+}\mu^{-} \pi^{+}$ decays, which may occur in the SM through an annihilation diagram, may provide a limiting background if it is not negligible around $m(\mu^{+}\mu^{-}) \sim m_{B^{0}_{(s)}}$. Nevertheless, it appears possible that interesting sensitivity to $B^{0}_{(s)} \rightarrow \mu^{+}\mu^{-}$ and $B^{0}_{s} \rightarrow \mu^{+}\mu^{-}$ decays may be achievable.

4 Summary

In summary, the $B^{-} \rightarrow \mu^{+}\mu^{-} \pi^{-}$ and $B^{+}_{c} \rightarrow \mu^{+}\mu^{-} \pi^{+}$ decays provide interesting possibilities to search for the leptonic weak decays of $D^{*0}$, $B^{0}$ and $B^{0}_{s}$ vector mesons. Published data allow a world-leading limit of $\mathbb{E}(D^{*0} \rightarrow \mu^{+}\mu^{-}) \lesssim 3 \times 10^{-7}$ to be obtained, and this can be significantly improved upon with a dedicated analysis of existing data. Sensitivity at the level of $\mathcal{O}(10^{-9})$ is expected to be possible with the data sample to be collected by the end of HL-LHC operation with upgrades of the LHCb experiment. Competitive measurements could potentially also be made by the ATLAS and CMS experiments, but dedicated studies will be necessary to evaluate how well background can be suppressed. Good sensitivity to $B^{0}_{(s)} \rightarrow \mu^{+}\mu^{-}$ decays also appears achievable, although further experi-
mental investigations will be needed before a firm conclusion can be reached on this point. In particular, measurements of \( B \left( B_s^{0} \rightarrow \mu^+ \mu^- \pi^+ \right) \) and studies of nonresonant \( B_c^+ \rightarrow \mu^+ \mu^- \pi^+ \) decays are needed. Further studies of semi-inclusive search approaches, exploiting vector meson production through semileptonic \( b \)-hadron decays, will be necessary to understand if these may allow even better sensitivity to be reached.

Acknowledgements The authors wish to thank their colleagues on the LHCb experiment for the fruitful and enjoyable collaboration that inspired this study. In particular, they would like to thank Niels Tuning and Mike Williams for helpful comments. Additionally, the authors acknowledge stimulating discussions with Benjamin Grinstein, Alex Khodjamirian and Alexey Petrov. This work is supported by the Royal Society and the Science and Technology Facilities Council (UK) and the Monash Warwick Alliance.

Data Availability Statement This manuscript has no associated data. [Authors’ comment: There is no associated data as the work is phenomenological in nature.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP3.

Appendix A: Sensitivity of searches exploiting prompt production

The predominant fraction of heavy-flavoured mesons are produced “promptly” in LHC collisions, i.e. they originate directly from the \( pp \) interaction vertex (the so-called primary vertex). Thus one might wonder whether the large production rate could overcome the sizeable background from random combinations of tracks produced at the primary vertex, and allow competitive results on \( D^0 \), \( B^0 \) and \( B_s^{0} \) dimuon decays to be obtained.

Fortuitously, the LHCb collaboration has published, as part of a search for dark photons decaying to \( \mu^+ \mu^- \) [54], the prompt dimuon spectrum obtained from a data sample corresponding to 1.6 fb\(^{-1}\) of proton-proton collisions recorded at a centre-of-mass energy of 13 TeV, as shown in Fig. 4.3 This can be used to estimate the sensitivity to the dimuon decays for each of the \( D^0 \), \( B^0 \) and \( B_s^{0} \) meson. As an example, for the \( D^0 \) case,

\[
\begin{align*}
B \left( D^0 \rightarrow \mu^+ \mu^- \right) &= \frac{N_{D^0 \rightarrow \mu^+ \mu^-}}{N_{J/\psi \rightarrow \mu^+ \mu^-}} \frac{\epsilon_{J/\psi \rightarrow \mu^+ \mu^-} \sigma_{J/\psi}}{\epsilon_{D^0 \rightarrow \mu^+ \mu^-} \sigma_{D^0} J/\psi} B \left( J/\psi \rightarrow \mu^+ \mu^- \right), \\
&= \alpha_{D^0 \rightarrow \mu^+ \mu^-} N_{D^0 \rightarrow \mu^+ \mu^-},
\end{align*}
\]

where \( N \) and \( \epsilon \) represent the yields for the decays given in the subscripts. Since analysis of data from the LHCb experiment is being considered, \( \sigma_{J/\psi} \) and \( \sigma_{D^0} \) refer to the inclusive prompt \( pp \rightarrow J/\psi \) and \( pp \rightarrow D^0 \) production cross-sections in the LHCb acceptance.

The \( J/\psi \) production cross-section at \( \sqrt{s} = 13 \) TeV has been measured to be \( \sigma_{J/\psi} = 15.03 \pm 0.03 \pm 0.94 \) \( \mu b \) [57]. The cross-section for the \( D^0 \) meson is not measured, but the approximation \( \sigma_{D^0} \sim \sigma_{D^{+}} \) can be made, so that \( \sigma_{D^{+}} = 784 \pm 4 \pm 87 \) \( \mu b \) at 13 TeV [58] can be used. Hence, estimating \( N_{J/\psi \rightarrow \mu^+ \mu^-} \sim 2 \times 10^7 \) from Fig. 4, and assuming that the ratio of efficiencies is approximately unity, one obtains a single-event sensitivity with 1.6 fb\(^{-1}\) of 13 TeV data of \( \alpha_{D^0 \rightarrow \mu^+ \mu^-} \sim 6 \times 10^{-11} \).

The single-event sensitivities for dimuon decays of \( B^0 \) and \( B_s^{0} \) mesons can be estimated by replacing \( \sigma_{D^0} \) with the corresponding \( B^{0} \) production cross-section. Since these are not measured, rough estimates need to be made: \( \sigma_{B^{0}} \) is taken to be one-tenth of \( \sigma_{D^{0}} \), with \( \sigma_{B^{0}} \) suppressed by a further factor of four, consistent with the relative production rates of beauty and charm pseudoscalar mesons [52, 58, 59]. This gives \( \alpha_{B^0 \rightarrow \mu^+ \mu^-} \sim 6 \times 10^{-10} \) and \( \alpha_{B_s^{0} \rightarrow \mu^+ \mu^-} \sim 2 \times 10^{-9} \).

As is clear from Fig. 4, searches for rare processes in prompt production will suffer from large backgrounds. The number of dimuon candidates per bin in the region around \( m_{D^0} \) is just over 10\(^5\), and since the bin width is one half of the mass resolution, this corresponds to a background yield of around 5 \times 10\(^5\) in a signal window of \pm 2 \) times the mass resolution around the signal peak position. Hence the uncertainty on \( N_{D^0 \rightarrow \mu^+ \mu^-} \) arising from a fit to this distribution is expected to be around 700. Similar uncertainties are estimated also for \( N_{B^0 \rightarrow \mu^+ \mu^-} \) and \( N_{B_s^{0} \rightarrow \mu^+ \mu^-} \). Assuming that the upper limit corresponds to twice this uncertainty this then gives

\[
\frac{\text{signal}}{\text{background}} = \frac{\alpha}{\sigma_{D^0} \sim 6 \times 10^{-11}}
\]

which is based on a larger data sample. A spectrum of prompt dimuons without any additional requirements, appropriate for the search proposed here, is available in Ref. [56]; however peaks from the \( J/\psi \) and other resonances have been vetoed so that no normalisation sample is available in the same spectrum.
Fig. 4 Spectrum of prompt dimuon pairs recorded by LHCb, with background contributions indicated, from Ref. [54]

Fig. 5 Upper limits at 90% confidence level on the production cross-section times branching fraction \(\sigma(X) B(X \to \mu^+\mu^-)\) obtained from an inclusive search for displaced \(X \to \mu^+\mu^-\) decays, from Ref. [56]. Limits are shown as a function of the mass of the hypothesised \(X\) particle and its transverse momentum.

\[
\begin{align*}
B(D^{*0} \to \mu^+\mu^-) & \approx 10^{-7}, \\
B(B^{*0} \to \mu^+\mu^-) & \approx 10^{-6}, \\
B(B^{*0}_{s} \to \mu^+\mu^-) & \approx 3 \times 10^{-6},
\end{align*}
\]

where the uncertainty on the single-event sensitivity has not been included.

Appendix B: Sensitivity of (semi-)inclusive searches exploiting displaced production

A possible strategy to avoid the large background in prompt searches could be to make an inclusive search for dimuon \(D^{*0}, B^{*0}\) and \(B^{*0}_{s}\) decays originating from a displaced vertex (i.e. from \(b\)-hadron decays). Considering for example the \(D^{*0} \to \mu^+\mu^-\) case, the number of signal decays in the sample will be given by

\[
N_{D^{*0} \to \mu^+\mu^-} = \mathcal{L} \times \mathcal{B} (D^{*0} \to \mu^+\mu^-) \sum_i \sigma_{B_i} \times \mathcal{B} (B_i \to D^{*0} X) \times \epsilon_{B_i \to D^{*0} X},
\]  

(7)

where \(\mathcal{L}\) is the integrated luminosity of the sample, \(\sigma_{B_i}\) is the production cross-section for \(b\)-hadron species \(B_i\) that has an inclusive branching fraction to decay to a final state containing a \(D^{*0}\) meson of \(\mathcal{B} (B_i \to D^{*0} X)\), where \(X\) indicates any set of additional particles.\(^4\) The corresponding efficiency is \(\epsilon_{B_i \to D^{*0} X}\), which will depend on the requirements imposed to select displaced vertices and on the lifetime and kinematics of the \(B_i\) hadron. Considering every possible \(B_i\) species would be challenging, but for the \(D^{*0}\) case production through \(B^+\) and \(B^0\) decays can be assumed to dominate, and knowledge exists of their production cross-sections [59] and inclusion of double charm baryons, for example. However, the additional contributions to the total inclusive rate are negligible.

---

These values are rough estimates of the sensitivity that could be obtained from the 1.6 fb\(^{-1}\) data sample of Ref. [54]. Real measurements with this approach would require the relevant production cross-sections to be determined, rather than approximated as done above. Such measurements require the reconstruction of the soft neutral particles emitted in the \(D^{*0} \to D^0\pi^0\) or \(D^0\gamma\), \(B^{*0} \to B^0\gamma\) and \(B^{*0}_{s} \to B^0_{s}\gamma\) decays, for which the LHCb detector is not optimised. These challenges are not insurmountable, as demonstrated by LHCb measurements of \(\chi_c\) and \(\chi_{b}\) production with photons reconstructed either in the calorimeter or through \(\gamma \to e^+e^-\) conversions in detector material [60–62]. Nonetheless, given that the limits are background dominated and will therefore improve only slowly with larger data samples, this approach looks less attractive compared to production through \(B\)-meson decays.

---

\(^4\) Most generally one should consider all weakly decaying hadrons with non-zero branching fractions for decays including \(D^{*0}\) mesons in the final state. This would include double charm baryons, for example. However, the additional contributions to the total inclusive rate are negligible.
sive branching fractions [63]. As usual for measurements in hadron collider experiments, it is experimentally convenient to use a normalisation channel in order to cancel the luminosity and avoid dependence on the absolute efficiencies. The $J/\psi \rightarrow \mu^+\mu^-$ channel is again well-suited for this purpose. Contributions from $B_{s}^{0}$ and $\Lambda_{b}^{0}$ decays are not negligible in this case but measurements exist of all needed quantities [31]. This approach however is not currently viable for dimuon $B_{s}^{0}$ and $B_{s}^{0}$ decay searches, due to lack of knowledge of the relevant cross-sections and branching fractions.

The LHCb collaboration has published the results of a search for inclusive displaced $X \rightarrow \mu^+\mu^-$ decays, without any requirement that the $X$ particle is produced at the primary $pp$ collision vertex [56]. This search was motivated by the possible existence of beyond Standard Model particles (generically denoted $X$) and includes selection requirements to suppress background from dimuon pairs produced in $b$-hadron decay chains. Moreover, only the region below the $J/\psi$ mass is included in the search so the normalisation channel is not available. Therefore, it is not useful to attempt to use the results of this search, shown in Fig. 5, to obtain information about $D_{s}^{0} \rightarrow \mu^+\mu^-$ decays; nonetheless, Ref. [56] demonstrates that a dedicated analysis using this approach could be carried out. However, compared to the exclusive approach emphasised in this paper, the mass resolution will be worse and higher irreducible backgrounds are expected as all $b$-hadron decays producing a nonresonant $\mu^+\mu^-$ pair plus one or more extra particles will contribute. As such, this approach appears unlikely to be competitive with large data samples.

Background from nonresonant $\mu^+\mu^-$ pairs produced in $b$-hadron decays is irreducible in a fully inclusive search. However, a semi-inclusive approach could be used to remove it, in which an additional muon (or electron) is required to originate from the same vertex. In this scenario, the source of $D_{s}^{0}$ mesons is through semileptonic $b$-hadron (dominated by $B_{s}$ and $B_{s}^{0}$) decays. Also in this case there is a relevant prior LHCb publication, in the form of a search for $B_{s}^{+} \rightarrow \mu^+\mu^-\nu\tau^-\tau^+$ decays [64]. This analysis, however, includes a requirement that one $\mu^+\mu^-$ pair has mass below 900 MeV/c$^2$, and therefore cannot be used directly to obtain a limit on $B_{s}^{+} \rightarrow (D_{s}^{0} \rightarrow \mu^+\mu^-) \mu^+\nu$. Nonetheless, it appears likely that a dedicated analysis, focusing on the $m(\mu^+\mu^-) \sim m_{D_{s}^{0}}$ mass range, could achieve a limit on $B (D_{s}^{0} \rightarrow \mu^+\mu^-)$ of $O(10^{-7})$. The possibility to relax selection requirements to allow additional particles, beyond the neutrino, in the decay could also be exploited if the corresponding improvement in single-event sensitivity is sufficient to overcome the increase in background. Thus, if background levels can be kept very low, this approach may in the long-term be competitive with, or even better than, the method with exclusive reconstruction of hadronic decays. Unfortunately, with the currently available information the achievable sensitivity cannot be reliably quantified. Similarly, the use of $B_{l} \rightarrow B_{s}^{0}(\mu\chi)$ production, with the most important contribution expected to be from semileptonic $B_{s}^{\ast}$ decays, to search for $B_{s}^{0}(\mu\chi) \rightarrow \mu^+\mu^-$ transitions appears an interesting possibility, but in the absence of relevant experimental results the achievable sensitivity cannot be estimated.

References

5. LHCb collaboration, R. Aaij et al., Measurement of the $B_{s}^{0} \rightarrow \mu^+\mu^-$ branching fraction and effective lifetime and search for $B_{s}^{0} \rightarrow \mu^+\mu^-$ decays. Phys. Rev. Lett. 118, 191801 (2017). arXiv:1703.05747
6. CMS collaboration, A.M. Sirunyan et al., Measurement of properties of $B_{s}^{0} \rightarrow \mu^+\mu^-$ decays and search for $B_{s}^{0} \rightarrow \mu^+\mu^-$ with the CMS experiment. JHEP 04, 188 (2020). arXiv:1910.12127
7. ATLAS collaboration, M. Aaboud et al., Study of the rare decays of $B_{s}^{0}$ and $B_{s}^{0}$ mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector. JHEP 04, 098 (2019). arXiv:1812.03017
9. LHCb collaboration, R. Aaij et al., Measurement of the $B_{s}^{0} \rightarrow \mu^+\mu^-$ decay properties and search for the $B_{s}^{0} \rightarrow \mu^+\mu^-$ and $B_{s}^{0} \rightarrow \mu^+\mu^-\gamma$ decays. Phys. Rev. D 105, 012010 (2022). arXiv:2108.09283
10. ATLAS, CMS and LHCb collaborations, Combination of the ATLAS, CMS and LHCb results on the $B_{s}^{0} \rightarrow \mu^+\mu^-$ decays. ATLAS-CONF-2020-049, CMS PAS BPH-20-003, LHCB-CONF-2020-002 (2020)
14. LHCb collaboration, R. Aaij et al., Search for the decays $B_{s}^{0} \rightarrow \tau^+\tau^-$ and $B_{s}^{0} \rightarrow \tau^+\tau^-$. Phys. Rev. Lett. 118, 251802 (2017). arXiv:1703.02508


28. A. Bharucha, D. Boito, C. Meaux, Disentangling QCD and new physics in $D \rightarrow \pi \ell \ell$. JHEP 04, 158 (2021).

29. CMD-3 collaboration, D.N. Shemyakin, Search for the process $e^{+}e^{-} \rightarrow D^{(*)}(2007)^{0}$ with the CMD-3 Detector. Phys. Atom. Nucl. 83, 954 (2020).

30. LHCb collaboration, R. Aaij et al., First measurement of the differential branching fraction and CP asymmetry of the $B^{0}_{s} \rightarrow \pi^{+}\mu^{-}\mu^{+}$ decay. JHEP 10, 034 (2015). arXiv:1509.00414


40. LHCb collaboration, Physics case for an LHCb Upgrade II—Opportunities in flavour physics, and beyond, in the HL-LHC era. arXiv:1808.08865

41. LHCb collaboration, Expression of interest for a phase-II LHCb upgrade: opportunities in flavour physics, and beyond, in the HL-LHC era. CERN-LHCC-2017-003

42. LHCb collaboration, R. Aaij et al., Measurement of CP observables in $B^{\pm} \rightarrow D^{(*)}\bar{K}^{(*)}$ and $B^{\pm} \rightarrow D^{(*)}\pi^{\pm}$ decays using two-body D final states. JHEP 04, 081 (2021). arXiv:2012.09903

43. LHCb collaboration, R. Aaij et al., Observation of the decay $B_{s}^{0} \rightarrow B_{s}^{0}\pi^{0}$. Phys. Rev. Lett. 111, 181801 (2013). arXiv:1308.4544

44. LHCb collaboration, R. Aaij et al., Precision measurement of the $B_{s}^{0}$ meson mass. JHEP 07, 123 (2020). arXiv:2004.08163

45. LHCb collaboration, R. Aaij et al., Observation of $B_{s}^{0} \rightarrow J/\psi D_{s}^{0}$ and $B_{s}^{+} \rightarrow J/\psi D_{s}^{+}$ decays. Phys. Rev. D 87, 112012 (2013). arXiv:1304.4530