Search for decays of neutral beauty mesons into four muons
LHCb Collaboration

DOI: 10.1007/JHEP03(2017)001
License: Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
LHCb Collaboration 2017, 'Search for decays of neutral beauty mesons into four muons', Journal of High Energy Physics, vol. 2017, no. 3, 1. https://doi.org/10.1007/JHEP03(2017)001

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 08/04/2019

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

Users may freely distribute the URL that is used to identify this publication.
Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.
Search for decays of neutral beauty mesons into four muons

The LHCb collaboration

E-mail: tobias.tekampe@cern.ch

ABSTRACT: A search for the non-resonant decays $B_s^0 \to \mu^+ \mu^- \mu^+ \mu^-$ and $B^0 \to \mu^+ \mu^- \mu^+ \mu^-$ is presented. The measurement is performed using the full Run 1 data set collected in proton-proton collisions by the LHCb experiment at the LHC. The data correspond to integrated luminosities of 1 and 2 fb$^{-1}$ collected at centre-of-mass energies of 7 and 8 TeV, respectively. No signal is observed and upper limits on the branching fractions of the non-resonant decays at 95% confidence level are determined to be

\begin{align*}
B(B_s^0 \to \mu^+ \mu^- \mu^+ \mu^-) &< 2.5 \times 10^{-9}, \\
B(B^0 \to \mu^+ \mu^- \mu^+ \mu^-) &< 6.9 \times 10^{-10}.
\end{align*}

KEYWORDS: B physics, Flavour Changing Neutral Currents, Hadron-Hadron scattering (experiments), Rare decay, Supersymmetry

ArXiv ePrint: 1611.07704

do:10.1007/JHEP03(2017)001
1 Introduction

The rare decays \( B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^- \) proceed through \( b \to d(s) \) flavour-changing neutral-current processes, which are strongly suppressed in the Standard Model (SM).\(^1\) In the main non-resonant SM amplitude, one muon pair is produced via amplitudes described by electroweak loop diagrams and the other is created by a virtual photon as shown in figure 1(a). The branching fraction of the non-resonant \( B^0_s \to \mu^+\mu^-\mu^+\mu^- \) decay is expected to be \( 3.5 \times 10^{-11} \) \([1]\).

Theories extending the SM can significantly enhance the \( B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^- \) decay rate by contributions of new particles. For example, in minimal supersymmetric models (MSSM), the decay can proceed via new scalar \( S \) and pseudoscalar \( P \) sgoldstino particles, which both decay into a dimuon final state as shown in figure 1(b). There are two types of couplings between sgoldstinos and SM fermions. Type-I couplings describe interactions between a sgoldstino and two fermions, where the coupling strength is proportional to the fermion mass. Type-II couplings describe a four-particle vertex, where a scalar and a pseudoscalar sgoldstino interact with two fermions. Branching fractions up to \( \mathcal{B}(B^0_s \to SP) \approx 10^{-4} \) and \( \mathcal{B}(B^0 \to SP) \approx 10^{-7} \) are possible \([2]\). Sgoldstinos can decay into a pair of photons or a pair of charged leptons \([3]\). In this analysis the muonic decay is considered, as the coupling to electrons is smaller and the large \( \tau \)-lepton mass limits the available phase space. The branching fractions of the sgoldstino decays strongly depend on the model parameters such as the sgoldstino mass and the supersymmetry breaking scale. In the search for \( \Sigma^+ \to \mu^+\mu^- \) decays the HyperCP collaboration found an excess of events,

\(^1\)The inclusion of charge-conjugate processes is implied throughout.
which is consistent with the decay $\Sigma^+ \to P p$ with $P \to \mu^+ \mu^-$ and a pseudoscalar mass of $m(P) = 214.3 \pm 0.5$ MeV [4].

So far only limits on the SM and MSSM branching fractions at 95% confidence level have been measured by LHCb based on the data recorded in 2011 [5] to be

$$\mathcal{B}(B^0_s \to \mu^+ \mu^- \mu^+ \mu^-) < 1.6 \times 10^{-8},$$
$$\mathcal{B}(B^0 \to \mu^+ \mu^- \mu^+ \mu^-) < 6.6 \times 10^{-9},$$
$$\mathcal{B}(B^0_s \to S(-\to \mu^+ \mu^-)P(-\to \mu^+ \mu^-)) < 1.6 \times 10^{-8},$$
$$\mathcal{B}(B^0 \to S(-\to \mu^+ \mu^-)P(-\to \mu^+ \mu^-)) < 6.3 \times 10^{-9}.$$  

These limits are based on assumed sgoldstino masses of $m(S) = 2.5$ GeV/$c^2$, which is approximately the central value of the allowed mass range, and $m(P) = 214.3$ MeV/$c^2$.

The dominant SM decays of neutral $B$ mesons into four muons proceed through resonances like $\phi$, $J/\psi$ and $\psi(2S)$. The most frequent of these decays is $B^0_s \to J/\psi \phi$, where both the $J/\psi$ and the $\phi$ mesons decay into a pair of muons, as shown in figure 1(c). In the following, this decay is referred to as the resonant decay mode and treated as a background. From the product of the measured branching fractions of the underlying decays $\mathcal{B}(B^0_s \to J/\psi \phi)$, $\mathcal{B}(J/\psi \to \mu^+ \mu^-)$, and $\mathcal{B}(\phi \to \mu^+ \mu^-)$ [6] its branching fraction is calculated to be $(1.83 \pm 0.18) \times 10^{-8}$.

In this paper a search for the non-resonant SM process, and for the MSSM-induced $B^0_{(s)} \to \mu^+ \mu^- \mu^+ \mu^-$ decays is presented, using $pp$ collision data recorded by the LHCb detector during LHC Run 1. Potentially contributing sgoldstinos are assumed to be short lived, such that they do not form a displaced vertex. The analysed data correspond to integrated luminosities of 1 and 2 fb$^{-1}$ collected at centre-of-mass energies of 7 and 8 TeV, respectively. The branching fraction is measured relative to the decay $B^+ \to J/\psi (\to \mu^+ \mu^-) K^+$, which gives a clean signal with a precisely measured branching fraction [6]. This yields a significant improvement compared to the previous measurement, where the use of the $B^0 \to J/\psi K^{*0}$ decay as normalisation mode resulted in a large systematic uncertainty originating from the S-wave fraction and the less precisely measured branching fraction. The advantage of normalising to a well-known $B$ meson decay is that dominant systematic uncertainties originating mainly from the $b \bar{b}$ cross-section cancel in the ratio.
2 Detector and simulation

The LHCb detector [7, 8] 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 (PV), 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 (RICH) 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.

In the simulation, pp collisions are generated using PYTHIA [9, 10] with a specific LHCb configuration [11]. Decays of hadronic particles are described by EvtGen [12], in which final-state radiation is generated using PHOTOS [13]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [14, 15] as described in ref. [16].

3 Event selection

The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. In this analysis candidate events are first required to pass the hardware trigger, which for 7 TeV (8 TeV) data selects events with at least one muon with a transverse momentum of \( p_T > 1.48 \) GeV/c (\( p_T > 1.76 \) GeV/c) or at least one pair of muons with the product of the transverse momenta larger than \( (1.296)^2 \) GeV\(^2/c^2\) ((1.6)^2 GeV\(^2/c^2\)). In the subsequent software trigger, at least one of the final-state particles is required to have \( p_T > 1 \) GeV/c and an impact parameter larger than 100 \( \mu m \) with respect to all PVs in the event.

In the offline selection, the \( B_{(s)}^0 \) decay vertex is constructed from four good quality muon candidates that form a common vertex and have a total charge of zero. The vertex is required to be significantly displaced from any PV. Among the four final-state muons, there are four possible dimuon combinations with zero charge. In all four combinations, the mass windows corresponding to the \( \phi \) (950–1090 MeV/c\(^2\)), \( J/\psi \) (3000–3200 MeV/c\(^2\)) and \( \psi(2S) \) (3600–3800 MeV/c\(^2\)) resonances are vetoed. This efficiently suppresses any background from any of the three mentioned resonances to a negligible level. Contributions of other charmonium states are found to be negligible.
The MatrixNet (MN) [17], a multivariate classifier based on a Boosted Decision Tree [18, 19], is applied in order to remove combinatorial background, where a candidate $B^0_{(s)}$ vertex is constructed from four muons that do not originate from a single $B$ meson decay. The input variables are the following properties of the $B^0_{(s)}$ candidate: the decay time, the vertex quality, the momentum and transverse momentum, the cosine of the direction angle (DIRA), and the smallest impact parameter chisquare ($\chi^2_{IP}$) with respect to any PV, where $\chi^2_{IP}$ is defined as the difference between the vertex-fit $\chi^2$ of a PV reconstructed with and without the $B^0_{(s)}$ candidate. The DIRA is defined as the angle between the momentum of the reconstructed $B^0_{(s)}$ candidate and the vector from the PV with the smallest $\chi^2_{IP}$ to the $B^0_{(s)}$ decay vertex. As training samples, simulated $B^0_s \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to \mu^+\mu^-\mu^+\mu^-$ events, generated with a uniform probability across the decay phase space, are used as a signal proxy. Before training, the signal simulation is weighted to correct for known discrepancies between data and simulation as described later. The background sample is taken from the far and the near sidebands in data as defined in table 1. In order to verify that the classification of each event is unbiased, 10-fold cross-validation [20] is employed.

Background arising from misidentifying one or more particles is suppressed by applying particle identification (PID) requirements. Information from the RICH system, the calorimeters and the muon system is used to calculate the difference in log-likelihood between the hypothesis of a final-state particle being a pion or a muon, $\text{DLL}_{\mu\pi}$.

Events in the signal region are not examined until the analysis is finalised. Events outside the signal region are split into the far sidebands, used to calculate the expected background yield, and the near sidebands, used to optimise the cuts on the MN response and the minimum DLL$_{\mu\pi}$ values of the four muon candidates in the final state. The optimization of the cuts is performed on a two-dimensional grid maximising the figure of merit [21]

$$\text{FoM} = \frac{\varepsilon_{\text{signal}}}{\sigma/2 + \sqrt{N_{\text{bkg}}^{\text{expected}}} \times \varepsilon_{\text{bkg}}}.$$  

The intended significance in terms of standard deviations ($\sigma$) is set to three. Very similar selection criteria are found when using five. The expected background yield before applying the MN and PID selection, $N_{\text{bkg}}^{\text{expected}}$, is determined from a fit to the events in the near sidebands using an exponential function. For each grid point the background efficiency, $\varepsilon_{\text{bkg}}$, is measured using events from the near sidebands. The signal efficiency, $\varepsilon_{\text{signal}}$, is measured for each grid point using simulated $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays. Lacking a model for non-resonant $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ simulation, the selection of the preceding measurement was developed on $B^0_s \to J/\psi(\to \mu^+\mu^-)\phi(\to \mu^+\mu^-)$ data. Now that a suitable simulation model is available, significant improvements in terms of signal efficiency and background rejection are made by employing a multivariate classifier and being able to measure the selection efficiency from simulation.

4 Selection efficiencies and systematic uncertainties

The optimal working point corresponds to signal efficiencies of $(0.580 \pm 0.003)\%$ and
(0.568 ± 0.003)% for the $B_s^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ decay modes, respectively. Sources of peaking background such as $B^0 \rightarrow K^*0\mu^+\mu^-$, in which the kaon and the pion originating from the $K^*$ decay are misidentified as muons, are reduced to a negligible level by the optimised selection. The efficiencies for the MSSM processes are measured using simulated samples of the $B^0_s \rightarrow S(\rightarrow \mu^+\mu^-)P(\rightarrow \mu^+\mu^-)$ decays, where the $B^0_s$ meson decays into a pseudoscalar sgoldstino with a mass of $214.3\,\text{MeV}/c^2$ [4] and a scalar sgoldstino with a mass of $2.5\,\text{GeV}/c^2$. Both the $P$ and $S$ particles are simulated with a decay width of $\Gamma = 0.1\,\text{MeV}/c^2$, which corresponds to a prompt decay. The measured efficiencies are the same for the $B^0_s$ and the $B^0$ decays and amount to $(0.648 \pm 0.003)\%$. The difference between the SM and the MSSM efficiencies originates from the fact that in the case of the decay proceeding via $P$ and $S$ sgoldstinos, the decay products are more likely to be within the acceptance of the LHCb detector. In order to test the dependence of the measured $B^0_s \rightarrow S(\rightarrow \mu^+\mu^-)P(\rightarrow \mu^+\mu^-)$ branching fractions on the mass of the scalar sgoldstino, the selection efficiency is measured in bins of dimuon invariant mass while requiring the corresponding other dimuon mass to be between 200 and 950 MeV/$c^2$. An efficiency variation of $O(20\%)$ is observed.

The selection applied to the normalisation mode $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ differs from that applied to the signal modes in the PID criteria and that no multivariate analysis technique is applied. The total efficiency is $(1.495 \pm 0.006)\%$. The uncertainties on the efficiencies are driven by the limited number of simulated events and are treated as systematic uncertainties of 0.4–0.5%.

The total efficiency is calculated as the product of the efficiencies of the different stages of the selection. As an alternative to the trigger efficiency calculated on simulation, the value is calculated on $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ data [22] and a systematic uncertainty of 3% is assigned corresponding to the relative difference. The efficiency of the MN classifier to select the more frequent decay $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi(\rightarrow K^+K^-)$ is compared between data and simulation. The relative difference of 0.3% is assigned as a systematic uncertainty. Another source of systematic uncertainty arises from the track finding efficiency. Again, values obtained from data [23] and simulation are compared and the deviation is treated as a correction factor for the efficiency, while the uncertainty on the deviation, 1.7%, is assigned as a systematic uncertainty.

In general the agreement in the observables used in the selection between data and simulation is very good, although there are some distributions that are known to deviate.

| Near sidebands | [5020, 5220] and [5426, 5626] |
| Far sidebands  | [4360, 5020] and [5626, 6360] |
| Signal region  | $[m(B^0) - 60, m(B^0_s) + 60]$ |
| $B^0_s$ search region | $[m(B^0) - 40, m(B^0_s) + 40]$ |
| $B^0$ search region  | $[m(B^0) - 40, m(B^0) + 40]$ |

Table 1. Definitions of intervals in the $B^0$ and $B^0_s$ reconstructed invariant mass distributions.
Therefore, the gradient boosting reweighting technique [24] is used to calculate weights that correct for differences between data and simulation in $B^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$. The weighting is performed in the track multiplicity, the $B$ transverse momentum, the $\chi^2$ of the decay vertex fit and the $\chi^2_{BP}$. The first two are chosen because they are correlated with the PID variables and the latter two dominate the feature ranking obtained from the MN training. These weights are applied to the $B^0_{(s)} \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ simulation samples, and are used to calculate the MN and the PID efficiencies. In order to account for inaccuracies of this method resulting from the kinematic and topological differences of the signal yield on $B^0_{(s)} \rightarrow \mu^+ \mu^- K^+$ and $B^0_{s} \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$. For the $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ decay mode, the efficiencies are measured with and without weights and the observed difference of 2.3% is assigned as systematic uncertainty.

In order to determine accurate efficiencies of the applied PID requirements, calibration samples of muons from $J/\psi \rightarrow \mu^+ \mu^-$ and $\phi \rightarrow \mu^+ \mu^-$ decays and of kaons from $D^+ \rightarrow D^0 (\rightarrow K^- \pi^+) \pi^+$ decays are used. The relative frequency for kaons and muons to pass the PID criteria is calculated in bins of track multiplicity, particle momentum and pseudorapidity. Different binning schemes are tested and the observed differences in the efficiencies of 1% for $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ and 0.5% for $B^0_{(s)} \rightarrow \mu^+ \mu^- K^+$ are assigned as systematic uncertainties. Additionally, 3% of the simulated $B^0_{(s)} \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays contain muons with low transverse momentum outside the kinematic region covered by the calibration data. This fraction is assigned as a systematic uncertainty. Candidates that have a reconstructed invariant mass within $\pm 40 \text{MeV}/c^2$ around the known $B^0_{(s)}$ mass, which corresponds to $\pm 2\sigma$ of the mass resolution, are treated as signal candidates. The accuracy of the efficiency of this cut is evaluated on $B^0_{(s)} \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ data. A systematic uncertainty of 0.5% corresponding to the relative difference of the efficiency measured on data and simulation is assigned. Systematic uncertainties of 0.9% and 0.5% in the case of $B^0_{(s)} \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ originate from the imperfections of the efficiency of the event reconstruction due to soft photon radiation and 0.6% from mismatching of track segments between different tracking stations in the detector, which is measured using simulated events. All relevant sources of systematic uncertainty along with the total values are summarised in table 2. The most significant improvements with regard to the preceding measurement are the larger available data sample, and the choice of the $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ decay as normalisation mode, which has the advantage of a precisely measured branching fraction and the absence of an additional systematic uncertainty originating from the S-wave correction.

5 Normalisation

The $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ signal yield is determined by performing an unbinned extended maximum likelihood fit to the $K^+ \mu^+ \mu^-$ invariant mass distribution. In this fit the $J/\psi$ mass is constrained [25] to the world average [6]. The normalisation yield is found to be $N(B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+) = 687890 \pm 920$. The $J/\psi K^+$ mass spectrum along
with the fit result is shown in figure 2. A systematic uncertainty of 0.3% is assigned to the determined \( B^+ \rightarrow J/\psi (\rightarrow \mu^+\mu^-)K^+ \) yield by using an alternative fit model and performing a binned extended maximum likelihood fit.

The \( B^0_{(s)} \rightarrow \mu^+\mu^-\mu^+\mu^- \) branching fraction is calculated as

\[
B(B^0_{(s)} \rightarrow \mu^+\mu^-\mu^+\mu^-) = N(B^0_{(s)} \rightarrow \mu^+\mu^-\mu^+\mu^-) \times \eta_{d,s},
\]

with

\[
\eta_{d,s} = \frac{\varepsilon(B^+ \rightarrow J/\psi (\rightarrow \mu^+\mu^-)K^+) \times B(B^+ \rightarrow J/\psi (\rightarrow \mu^+\mu^-)K^+) \times f_s}{f_d},
\]

where \( N(B^+ \rightarrow J/\psi (\rightarrow \mu^+\mu^-)K^+) \) and \( N(B^0_{(s)} \rightarrow \mu^+\mu^-\mu^+\mu^-) \) are the observed yields of the normalisation and the signal channel, respectively. The ratio between the production rates of \( B^0_s \) and \( B^0 \) was measured by LHCb to be \( f_s/f_d = 0.259 \pm 0.015 \) [27]. The measurement was performed using \( pp \) collision data at \( \sqrt{s} = 7 \text{ TeV} \), but found to be stable between \( \sqrt{s} = 7 \text{ TeV} \) and 8 TeV by a previous LHCb measurement [28]. The ratio between the \( B^+ \) and \( B^0 \) production rates is assumed to be unity. As a consequence \( f_s/f_d \) is equal to \( f_s/f_d \).

The single event sensitivities, \( \eta_{d,s} \), amount

\[
\eta_s^{\text{SM}} = (8.65 \pm 0.80) \times 10^{-10},
\]
\[
\eta_d^{\text{SM}} = (2.29 \pm 0.16) \times 10^{-10},
\]
\[
\eta_s^{\text{MSSM}} = (7.75 \pm 0.72) \times 10^{-10},
\]
\[
\eta_d^{\text{MSSM}} = (2.01 \pm 0.14) \times 10^{-10},
\]
for the $B^0_s$ and the $B^0$ decay modes in the SM and in the MSSM scenario. Here, the uncertainties are the combined values of the statistical uncertainty on the $B^+ \to J/\psi (\to \mu^+\mu^-)K^+$ yield and the systematic uncertainty. In the case of $\eta_s$ the systematic uncertainty is dominated by the ratio of $f_u/f_s$ and in the case of $\eta_d$ by the weighting procedure applied to correct for the difference between data and simulation.

The individual sources of systematic uncertainties given in table 2 are assumed to be uncorrelated and are combined quadratically. The total systematic uncertainty is 9.2% for the $B^0_s$ decay and 7.2% for the $B^0$ decay. These values are small compared to the statistical uncertainty on the expected number of background events in the $B^0$ and $B^0_s$ search regions.

The whole analysis strategy is cross-checked by measuring the $B^0_s \to J/\psi (\to \mu^+\mu^-)\phi(\to \mu^+\mu^-)$ branching fraction. The obtained value has a precision of 20% and is compatible with the product of the branching fractions of the underlying decays taken from ref. [6].

The number of expected background events is determined by fitting an exponential function to the far sidebands of $m(B_{(s)} \to J/\psi \phi)$ width windows around the $B^0$ meson masses yields the number of expected background events,

$$N_{\text{bkg}}(B^0) = 0.55^{+0.24}_{-0.19} \text{ (stat)} \pm 0.20 \text{ (syst)}$$
$$N_{\text{bkg}}(B^0_s) = 0.47^{+0.23}_{-0.18} \text{ (stat)} \pm 0.18 \text{ (syst)}.$$

The statistical uncertainty is the combination of the Poissonian uncertainty originating from the limited size of the data sample and the uncertainty on the fit parameters. As an alternative fit model a second-order polynomial is used and the difference between these background expectations is assigned as a systematic uncertainty.

6 Results

The final distribution of the reconstructed mass of the four muon system is shown in figure 3. No candidates are observed in either the $B^0$ or the $B^0_s$ search region, which is consistent with the expected background yield.

The Hybrid CLs procedure [29–31], with log-normal priors to account for uncertainties of both background and efficiency estimations, is used to convert the observations into upper limits on the corresponding branching fractions. The exclusion at 95% confidence level assuming the SM single event sensitivities is shown in figure 4. The result for the corresponding MSSM values is presented in figure 5. The limits on the branching fractions of the $B^0_s$ and $B^0$ decays are anti-correlated. Replacing the log-normal priors by gamma distributions yields the same results.

Assuming negligible cross-feed between the $B^0_s$ and the $B^0$ search regions, the observed upper limits on the branching fractions at 95% confidence level are found to be

$$B(B^0_s \to \mu^+\mu^-\mu^+\mu^-) < 2.5 \times 10^{-9},$$
$$B(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 6.9 \times 10^{-10},$$
$$B(B^0_s \to S(\to \mu^+\mu^-)P(\to \mu^+\mu^-)) < 2.2 \times 10^{-9},$$
$$B(B^0 \to S(\to \mu^+\mu^-)P(\to \mu^+\mu^-)) < 6.0 \times 10^{-10}.$$
Table 2. Summary of systematic uncertainties affecting the single event sensitivities along with the total systematic uncertainty calculated by adding up the individual components in quadrature. The dominating uncertainty arising from $f_u/f_s$ only contributes to $\eta_s$. The uncertainty of the stated selection efficiencies arising from the limited number of simulated events is 0.5% for $B^0 \to \mu^+ \mu^- \mu^+ \mu^-$ and 0.4% for all other considered decay modes.

7 Conclusion

In summary, a search for non-resonant $B^{(s)}_s \to \mu^+ \mu^- \mu^+ \mu^-$ decays has been presented. In addition, the sensitivity to a specific MSSM scenario has been probed. The applied selection focuses on finding four muon tracks that form a common vertex. For the SM scenario and the MSSM decay through short-lived scalar and pseudoscalar new particles, the limits set by the previous measurement performed by LHCb on a subset of the present data, are improved by a factor of 6.4 (7.3) for the SM (MSSM) mode in the case of the $B^0_s$ decay and by a factor of 9.5 (10.5) in the case of the $B^0$ decay.
Figure 3. Mass distribution of selected $B^0_{(s)} \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ events observed in 3 fb$^{-1}$ of data in all considered $B$ mass regions. Background (red line) is modelled by an exponential function. Signal subregions for $B^0$ and $B^0_{s}$ searches are also shown. The error bars on the individual points with $n$ entries are $\pm \sqrt{n}$.

Figure 4. Expected and observed 95% CL exclusion in $B(B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-)$ vs. $B(B^0_s \rightarrow \mu^+ \mu^- \mu^+ \mu^-)$ parameters plane.

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...
Figure 5. Expected and observed 95% CL exclusion in $B(B_0 \to S(\to \mu^+\mu^-)P(\to \mu^+\mu^-))$ vs. $B(B_0 \to S(\to \mu^+\mu^-)P(\to \mu^+\mu^-))$ parameters plane with scalar and pseudoscalar $S$ and $P$ as described in section 3.

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 FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.A.). 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 (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 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), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] Y. Dincer and L.M. Sehgal, Electroweak effects in the double Dalitz decay $B_s \to \ell^+\ell^-\ell^+\ell^-$, Phys. Lett. B 556 (2003) 169 [hep-ph/0301056] [inSPIRE].
[2] S.V. Demidov and D.S. Gorbunov, Flavor violating processes with sgoldstino pair production, *Phys. Rev. D* **85** (2012) 077701 [arXiv:1112.5230] [inSPIRE].

[3] S.V. Demidov and D.S. Gorbunov, More about sgoldstino interpretation of HyperCP events, *JETP Lett.* **84** (2007) 479 [hep-ph/0610066] [inSPIRE].

[4] HyperCP collaboration, H. Park et al., Evidence for the decay $\Sigma^+ \rightarrow p\mu^+\mu^-$, *Phys. Rev. Lett.* **94** (2005) 021801 [hep-ex/0501014] [inSPIRE].

[5] LHCb collaboration, Search for rare $B^0_s \rightarrow \mu^+\mu^-\mu^-\mu^-$ decays, *Phys. Rev. Lett.* **110** (2013) 211801 [arXiv:1303.1092] [inSPIRE].

[6] Particle Data Group collaboration, C. Patrignani et al., Review of particle physics, *Chin. Phys. C* **40** (2016) 100001 [inSPIRE].

[7] LHCb collaboration, The LHCb detector at the LHC, *JINST* **3** S08005 [inSPIRE].

[8] LHCb collaboration, LHCb detector performance, *Int. J. Mod. Phys. A* **30** (2015) 1530022 [arXiv:1412.6352] [inSPIRE].

[9] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175] [inSPIRE].

[10] T. Sjöstrand, S. Mrenna and P.Z. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [arXiv:0710.3820] [inSPIRE].

[11] LHCb collaboration, Handling of the generation of primary events in Gauss, the LHCb simulation framework, *J. Phys. Conf. Ser.* **331** (2011) 032047 [inSPIRE].

[12] D.J. Lange, The EvtGen particle decay simulation package, *Nucl. Instrum. Meth. A* **462** (2001) 152 [inSPIRE].

[13] P. Golonka and Z. Was, PHOTOS Monte Carlo: a precision tool for QED corrections in $Z$ and $W$ decays, *Eur. Phys. J. C* **45** (2006) 97 [hep-ph/0506026] [inSPIRE].

[14] GEANT4 collaboration, J. Allison et al., GEANT4 developments and applications, *IEEE Trans. Nucl. Sci.* **53** (2006) 270.

[15] GEANT4 collaboration, S. Agostinelli et al., GEANT4: a simulation toolkit, *Nucl. Instrum. Meth. A* **506** (2003) 250 [inSPIRE].

[16] LHCb collaboration, The LHCb simulation application, Gauss: design, evolution and experience, *J. Phys. Conf. Ser.* **331** (2011) 032047 [inSPIRE].

[17] A. Gulin, I. Kuralenok and D. Pavlov, Winning the transfer learning track of Yahoo!’s learning to rank challenge with YetiRank, *JMLR Proc.* **14** (2011) 63.

[18] L. Breiman, J. H. Friedman, R. A. Olshen and C.J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont U.S.A. (1984).

[19] R.E. Schapire and Y. Freund, A decision-theoretic generalization of on-line learning and an application to boosting, *J. Comput. Syst. Sci.* **55** (1997) 119.

[20] A. Blum, A. Kalai and J. Langford, Beating the hold-out: bounds for $k$-fold and progressive cross-validation, in the proceedings of the Twelfth Annual Conference on Computational Learning Theory (COLT’99), July 7–9, New York, U.S.A. (1999).

[21] G. Punzi, Sensitivity of searches for new signals and its optimization, *Conf C* **030908** (2003) MODT002 [physics/0308063] [inSPIRE].
[22] S. Tolk, J. Albrecht, F. Dettori and A. Pellegrino, Data driven trigger efficiency determination at LHCb, LHCb-PUB-2014-039 (2014).

[23] LHCb collaboration, Measurement of the track reconstruction efficiency at LHCb, 2015 JINST 10 P02007 [arXiv:1408.1251] [nSPIRE].

[24] A. Rogozhnikov, Reweighting with boosted decision trees, J. Phys. Conf. Ser. 762 (2016) 012036 [arXiv:1608.05806] [nSPIRE].

[25] W.D. Hulsbergen, Decay chain fitting with a Kalman filter, Nucl. Instrum. Meth. A 552 (2005) 566 [physics/0503191] [nSPIRE].

[26] D. Martínez Santos and F. Dupertuis, Mass distributions marginalized over per-event errors, Nucl. Instrum. Meth. A 764 (2014) 150 [arXiv:1312.5000] [nSPIRE].

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

[28] LHCb collaboration, Measurement of the $B^{0}_{d} \rightarrow \mu^{+}\mu^{-}$ branching fraction and search for $B^{0} \rightarrow \mu^{+}\mu^{-}$ decays at the LHCb experiment, Phys. Rev. Lett. 111 (2013) 101805 [arXiv:1307.5024] [nSPIRE].

[29] R.D. Cousins and V.L. Highland, Incorporating systematic uncertainties into an upper limit, Nucl. Instrum. Meth. A 320 (1992) 331 [nSPIRE].

[30] A.L. Read, Presentation of search results: the CL(s) technique, J. Phys. G 28 (2002) 2693.

[31] T. Junk, Confidence level computation for combining searches with small statistics, Nucl. Instrum. Meth. A 434 (1999) 435 [hep-ex/9902006] [nSPIRE].
Sezione INFN di Bologna, Bologna, Italy
Sezione INFN di Cagliari, Cagliari, Italy
Sezione INFN di Ferrara, Ferrara, Italy
Sezione INFN di Firenze, Firenze, Italy
Laboratori Nazionali dell’INFN di Frascati, Frascati, Italy
Sezione INFN di Genova, Genova, Italy
Sezione INFN di Milano, Milano, Italy
Sezione INFN di Padova, Padova, Italy
Sezione INFN di Pisa, Pisa, Italy
Sezione INFN di Roma Tor Vergata, Roma, Italy
Sezione INFN di Roma La Sapienza, Roma, Italy
Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
AGH — University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
National Center for Nuclear Research (NCBJ), Warsaw, Poland
Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
Institute of Nuclear Physics, Moscow State University, Moscow, Russia
Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
Yandex School of Data Analysis, Moscow, Russia
Budker Institute of Nuclear Physics (SB RAS), Novosibirsk, Russia
Institute for High Energy Physics (IHEP), Protvino, Russia
ICCU, Universitat de Barcelona, Barcelona, Spain
Universidad de Santiago de Compostela, Santiago de Compostela, Spain
European Organization for Nuclear Research (CERN), Geneva, Switzerland
Institut de Physique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
Physik-Institut, Universität Zürich, Zürich, Switzerland
Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands
NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
University of Birmingham, Birmingham, United Kingdom
H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
Department of Physics, University of Warwick, Coventry, United Kingdom
STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
Imperial College London, London, United Kingdom
School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
Department of Physics, University of Oxford, Oxford, United Kingdom
Massachusetts Institute of Technology, Cambridge, MA, United States
University of Cincinnati, Cincinnati, OH, United States
University of Maryland, College Park, MD, United States
Syracuse University, Syracuse, NY, United States
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to

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 Genova, Genova, Italy
i Università di Milano Bicocca, Milano, Italy
j Università di Roma Tor Vergata, Roma, Italy
k Università di Roma La Sapienza, Roma, Italy
l AGH — University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
m LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
n Hanoi University of Science, Hanoi, Viet Nam
o Università di Padova, Padova, Italy
p Università di Pisa, Pisa, Italy
q Università degli Studi di Milano, Milano, Italy
r Università di Urbino, Urbino, Italy
s Università della Basilicata, Potenza, Italy
t Scuola Normale Superiore, Pisa, Italy
u Università di Modena e Reggio Emilia, Modena, Italy
v Iligan Institute of Technology (IIT), Iligan, Philippines
w Novosibirsk State University, Novosibirsk, Russia
x Moscow Institute of Physics and Technology, Moscow, Russia ¹Deceased