Physics Letters B 708 (2012) 55–67

Contents lists available at SciVerse ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ ☆

LHCb Collaboration

ARTICLE INFO

Article history:
Received 8 December 2011
Received in revised form 12 January 2012
Accepted 13 January 2012
Available online 18 January 2012
Editor: W.-D. Schlatter

Keywords:
LHC
b-Hadron
FCNC
Rare decays
Leptonic decays

ABSTRACT

A search for the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ is performed with 0.37 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected by the LHCb experiment in 2011. The upper limits on the branching fractions are $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 1.6 \times 10^{-8}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.6 \times 10^{-9}$ at 95% confidence level. A combination of these results with the LHCb limits obtained with the 2010 dataset leads to $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 1.4 \times 10^{-8}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.2 \times 10^{-9}$ at 95% confidence level.

1. Introduction

Measurements of low-energy processes can provide indirect constraints on particles that are too heavy to be produced directly. This is particularly true for Flavour Changing Neutral Current (FCNC) processes which are highly suppressed in the Standard Model (SM) and can only occur through higher-order diagrams. The SM predictions for the branching fractions of the FCNC decays[1] $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (0.10 \pm 0.01) \times 10^{-9}$ [1]. However, contributions from new processes or new heavy particles can significantly enhance these values. For example, within Minimal Supersymmetric extensions of the SM (MSSM), in the large $\tan\beta$ regime, $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ is found to be approximately proportional to $\tan^6\beta$ [2], where $\tan\beta$ is the ratio of the vacuum expectation values of the two neutral *CP*-even Higgs fields. The branching fractions could therefore be enhanced by orders of magnitude for large values of $\tan\beta$.

The best published limits from the Tevatron are $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 5.1 \times 10^{-8}$ at 95% confidence level (CL) by the D0 Collaboration using 6.1 fb^{-1} of data [3], and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 6.0 \times 10^{-9}$ at 95% CL by the CDF Collaboration using 6.9 fb^{-1} of data [4]. In the same dataset the CDF Collaboration observes an excess of $B_s^0 \to \mu^+\mu^-$ candidates compatible with $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (1.8^{+1.1}_{-0.9}) \times 10^{-8}$ and with an upper limit of $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.0 \times 10^{-8}$ at 95% CL. The CMS Collaboration has recently published $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 1.9 \times 10^{-8}$ at 95% CL and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 4.6 \times 10^{-9}$ at 95% CL using 1.14 fb^{-1} of data [5]. The LHCb Collaboration has published the limits [6] $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 5.4 \times 10^{-8}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.5 \times 10^{-8}$ at 95% CL based on about 37 pb^{-1} of integrated luminosity collected in the 2010 run.

This Letter presents an analysis of the data recorded by LHCb in the first half of 2011 which correspond to an integrated luminosity of ~ 0.37 fb^{-1}. The results of this analysis are then combined with those published from the 2010 dataset.

2. The LHCb detector

The LHCb detector [7] is a single-arm forward spectrometer designed to study production and decays of hadrons containing b or c quarks. The detector consists of a vertex locator (VELO) providing precise locations of primary pp interaction vertices and detached vertices of long lived hadrons.

The momenta of charged particles are determined using information from the VELO together with the rest of the tracking system, composed of a large area silicon tracker located before a warm dipole magnet with a bending power of ~ 4 Tm, and a combination of silicon strip detectors and straw drift chambers located after the magnet. Two Ring Imaging Cherenkov (RICH) detectors are used for charged hadron identification in the momentum range 2–100 GeV/c. Photon, electron and hadron candidates are identified by electromagnetic and hadronic calorimeters. A muon system of alternating layers of iron and drift chambers provides muon identification. The two calorimeters and the muon system provide the energy and momentum information to implement a first level (L0) hardware trigger. An additional trigger level (HLT) is software

☆ © CERN for the benefit of the LHCb Collaboration.

[1] Inclusion of charged conjugated processes is implied throughout.

0370-2693/
doi:10.1016/j.physletb.2012.01.038

based, and its algorithms are tuned to the experimental operating condition.

Events with a muon final states are triggered using two L0 trigger decisions: the single-muon decision, which requires one muon candidate with a transverse momentum $p_{\rm T}$ larger than 1.5 GeV/c, and the di-muon decision, which requires two muon candidates with transverse momenta $p_{\rm T,1}$ and $p_{\rm T,2}$ satisfying the relation $\sqrt{p_{\rm T,1}\cdot p_{\rm T,2}} > 1.3$ GeV/c.

The single muon trigger decision in the second trigger level (HLT) includes a cut on the impact parameter (IP) with respect to the primary vertex, which allows for a lower $p_{\rm T}$ requirement ($p_{\rm T} > 1.0$ GeV/c, IP > 0.1 mm). The di-muon trigger decision requires muon pairs of opposite charge with $p_{\rm T} > 500$ MeV/c, forming a common vertex and with an invariant mass $m_{\mu\mu} > 4.7$ GeV/c^2. A second trigger decision, primarily to select J/ψ events, requires $2.97 < m_{\mu\mu} < 3.21$ GeV/c^2. The remaining region of the di-muon invariant mass range is also covered by trigger decisions that in addition require the di-muon secondary vertex to be well separated from the primary vertex.

Events with purely hadronic final states are triggered by the L0 trigger if there is a calorimeter cluster with transverse energy $E_{\rm T} > 3.6$ GeV. Other HLT trigger decisions select generic displaced vertices, providing high efficiency for purely hadronic decays.

3. Analysis strategy

Assuming the branching fractions predicted by the SM, and using the $b\bar{b}$ cross-section measured by LHCb in the pseudorapidity interval $2 < \eta < 6$ and integrated over all transverse momenta of $\sigma_{b\bar{b}} = 75 \pm 14$ μb [8], approximately 3.9 $B^0_s \to \mu^+\mu^-$ and 0.4 $B^0 \to \mu^+\mu^-$ events are expected to be triggered, reconstructed and selected in the analyzed sample embedded in a large background.

The general structure of the analysis is based upon the one described in Ref. [6]. First a very efficient selection removes the biggest amount of background while keeping most of the signal within the LHCb acceptance. The number of observed events is compared to the number of expected signal and background events in bins of two independent variables, the invariant mass and the output of a multi-variate discriminant. The discriminant is a Boosted Decision Tree (BDT) constructed using the TMVA package [9]. It supersedes the Geometrical Likelihood (GL) used in the previous analysis [6] as it has been found more performant in discriminating between signal and background events in simulated samples. No data were used in the choice of the multivariate discriminant in order not to bias the result.

The combination of variables entering the BDT discriminant is optimized using simulated events. The probability for a signal or background event to have a given value of the BDT output is obtained from data using $B^0_{(s)} \to h^+h'^-$ candidates (where $h^{(\prime)}$ can be a pion or a kaon) as signal and sideband $B^0_{(s)} \to \mu^+\mu^-$ candidates as background.

The invariant mass line shape of the signals is described by a Crystal Ball function [10] with parameters extracted from data control samples. The central values of the masses are obtained from $B^0 \to K^+\pi^-$ and $B^0_s \to K^+K^-$ samples. The B^0_s and B^0 mass resolutions are estimated by interpolating those obtained with di-muon resonances (J/ψ, $\psi(2S)$ and $\Upsilon(1S, 2S, 3S)$) and cross-checked with a fit to the invariant mass distributions of both inclusive $B^0_{(s)} \to h^+h'^-$ decays and exclusive $B^0 \to K^+\pi^-$ decays. The central values of the masses and the mass resolution are used to define the signal regions.

The number of expected signal events, for a given branching fraction hypothesis, is obtained by normalizing to channels of known branching fractions: $B^+ \to J/\psi K^+$, $B^0_s \to J/\psi\phi$ and $B^0 \to K^+\pi^-$. These channels are selected in a way as similar as possible to the signals in order to minimize the systematic uncertainty related to the different phase space accessible to each final state.

The BDT output and invariant mass distributions for combinatorial background events in the signal regions are obtained using fits of the mass distribution of events in the mass sidebands in bins of the BDT output.

The two-dimensional space formed by the invariant mass and the BDT output is binned. For each bin we count the number of candidates observed in the data, and compute the expected number of signal events and the expected number of background events. The binning is unchanged with respect to the 2010 analysis [6]. The compatibility of the observed distribution of events in all bins with the distribution expected for a given branching fraction hypothesis is computed using the $\rm CL_s$ method [11], which allows a given hypothesis to be excluded at a given confidence level.

4. Selection

The $B^0_{(s)} \to \mu^+\mu^-$ selections require two muon candidates of opposite charge. Tracks are required to be of good quality and to be displaced with respect to any primary vertex. The secondary vertex is required to be well fitted ($\chi^2/{\rm nDoF} < 9$) and must be separated from the primary vertex in the forward direction by a distance of flight significance ($L/\sigma(L)$) greater than 15. When more than one primary vertex is reconstructed, the one that gives the minimum impact parameter significance for the candidate is chosen. The reconstructed candidate has to point to this primary vertex ($\rm IP/\sigma(IP) < 5$).

Improvements have been made to the selection developed for 2010 data [6]. The RICH is used to identify kaons in the $B^0_s \to J/\psi\phi$ normalization channel and the Kullback–Leibler (KL) distance [12] is used to suppress duplicated tracks created by the reconstruction. This procedure compares the parameters and correlation matrices of the reconstructed tracks and where two are found to be similar, in this case with a symmetrized KL divergence less than 5000, only the one with the higher track fit quality is considered.

The inclusive $B^0_{(s)} \to h^+h'^-$ sample is the main control sample for the determination from data of the probability distribution function (PDF) of the BDT output. This sample is selected in exactly the same way as the $B^0_{(s)} \to \mu^+\mu^-$ signals apart from the muon identification requirement. The same selection is also applied to the $B^0 \to K^+\pi^-$ normalization channel.

The muon identification efficiency is uniform within $\sim 1\%$ in the considered phase space therefore no correction is added to the BDT PDF extracted from the $B^0_{(s)} \to h^+h'^-$ sample. The remaining phase space dependence of the muon identification efficiency is instead taken into account in the computation of the normalization factor when the $B^0 \to K^+\pi^-$ channel is considered.

The $J/\psi \to \mu\mu$ decay in the $B^+ \to J/\psi K^+$ and $B^0_s \to J/\psi\phi$ normalization channels is selected in a very similar way to the $B^0_{(s)} \to \mu^+\mu^-$ channels, apart from the pointing requirement. $K^\pm$ candidates are required to be identified by the RICH detector and to pass track quality and impact parameter cuts.

To avoid pathological events, all tracks from selected candidates are required to have a momentum less than 1 TeV/c. Only B candidates with decay times less than $5\,\tau_{B^0_{(s)}}$, where $\tau_{B^0_{(s)}}$ is the B lifetime [13], are accepted for further analysis. Di-muon candidates coming from elastic di-photon production are removed by requiring a minimum transverse momentum of the B candidate of 500 MeV/c.

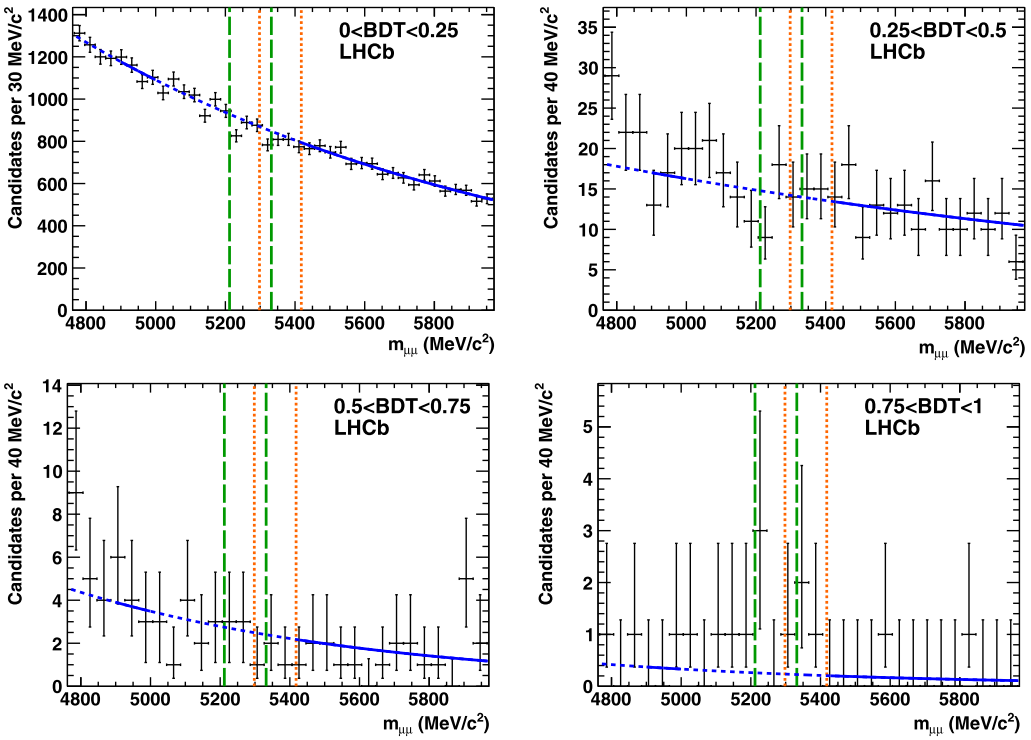

Fig. 1. Distribution of the $\mu^+\mu^-$ invariant mass for events in each BDT output bin. The curve shows the model used to fit the sidebands and extract the expected number of combinatorial background events in the B_s^0 and B^0 signal regions, delimited by the vertical dotted orange and dashed green lines respectively. Only events in the region in which the line is solid have been considered in the fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

5. Determination of the mass and BDT distributions

The variables entering the BDT discriminant are the six variables used as input to the GL in the 2010 analysis plus three new variables. The six variables used in the 2010 analysis are the B lifetime, impact parameter, transverse momentum, the minimum impact parameter significance (IP/σ(IP)) of the muons, the distance of closest approach between the two muons and the isolation of the two muons with respect to any other track in the event. The three new variables are:

1. the minimum p_{T} of the two muons;
2. the cosine of the angle between the muon momentum in the B rest frame and the vector perpendicular to the B momentum and the beam axis:

$$\cos P = \frac{p_{y,\mu 1} p_{x,B} - p_{x,\mu 1} p_{y,B}}{p_{\mathrm{T},B}(m_{\mu\mu}/2)} \tag{1}$$

where μ_1 labels one of the muons and $m_{\mu\mu}$ is the reconstructed B candidate mass[2];
3. the B isolation [14]

$$I_B = \frac{p_{\mathrm{T}}(B)}{p_{\mathrm{T}}(B) + \sum_i p_{\mathrm{T},i}}, \tag{2}$$

where $p_{\mathrm{T}}(B)$ is the B transverse momentum with respect to the beam line and the sum is over all the tracks, excluding the muon candidates, that satisfy $\sqrt{\delta\eta^2 + \delta\phi^2} < 1.0$, where $\delta\eta$ and $\delta\phi$ denote respectively the difference in pseudorapidity and azimuthal angle between the track and the B candidate.

The BDT output is found to be independent of the invariant mass for both signal and background and is defined such that the signal is uniformly distributed between zero and one and the background peaks at zero. The BDT range is then divided in four bins of equal width. The BDT is trained using simulated samples ($B^0_{(s)} \to \mu^+\mu^-$ for signals and $b\bar{b} \to \mu^+\mu^- X$ for background where X is any other set of particles) and the PDF obtained from data as explained below.

5.1. Combinatorial background PDFs

The BDT and invariant mass shapes for the combinatorial background inside the signal regions are determined from data by interpolating the number of expected events using the invariant mass sidebands for each BDT bin. The boundaries of the signal regions are defined as $m_{B^0} \pm 60$ MeV/c^2 and $m_{B_s^0} \pm 60$ MeV/c^2 and the mass sidebands as $[m_{B^0} - 600\ \mathrm{MeV}/c^2, m_{B^0} - 60\ \mathrm{MeV}/c^2]$ and $[m_{B_s^0} + 60\ \mathrm{MeV}/c^2, m_{B_s^0} + 600\ \mathrm{MeV}/c^2]$.

Fig. 1 shows the invariant mass distribution for events that lie in each BDT output bin. In each case the fit model used to estimate the expected number of combinatorial background events in the signal regions is superimposed.

Aside from combinatorial background, the low-mass sideband is potentially polluted by two other contributions: cascading $b \to c\mu\nu \to \mu\mu X$ decays below 4900 MeV/c^2 and peaking background from $B^0_{(s)} \to h^+h'^-$ candidates with the two hadrons misidentified as muons above 5000 MeV/c^2. To avoid these contaminations, the number of expected combinatorial background events is obtained by fitting a single exponential function to the events in the reduced low-mass sideband $[4900, 5000]$ MeV/c^2 and in the full

[2] As the B is a (pseudo)-scalar particle, this variable is uniformly distributed for signal candidates while is peaked at zero for $b\bar{b} \to \mu^+\mu^- X$ background candidates. In fact, muons from semi-leptonic decays are mostly emitted in the direction of the b's and, therefore, lie in a plane formed by the B momentum and the beam axis.

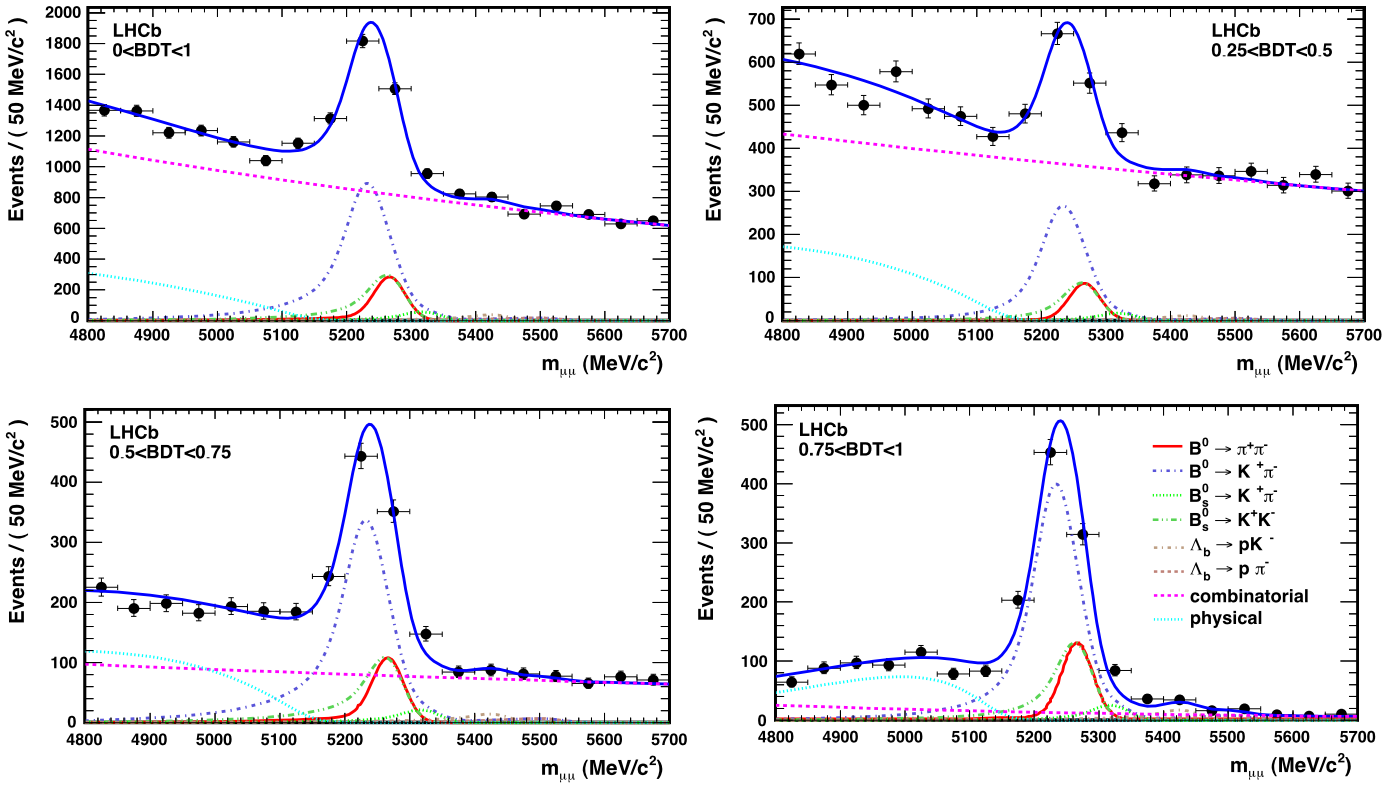

Fig. 2. Invariant mass distributions of $B^0_{(s)} \to h^+h'^-$ candidates in the $\mu^+\mu^-$ mass hypothesis for the whole sample (top left) and for the samples in the three highest bins of the BDT output (top right, bottom left, bottom right). The $B^0_{(s)} \to h^+h'^-$ exclusive decays, the combinatorial background and the physical background components are drawn under the fit to the data (solid blue line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

high-mass sideband. As a cross-check, two other models, a single exponential function and the sum of two exponential functions, have been used to fit the events in different ranges of sidebands providing consistent background estimates inside the signal regions.

5.2. Peaking background PDFs

The peaking backgrounds due to $B^0_{(s)} \to h^+h'^-$ events in which both hadrons are misidentified as muons have been evaluated from data and simulated events to be $N_{B^0_s} = 1.0 \pm 0.4$ events and $N_{B^0} = 5.0 \pm 0.9$ events within the two mass windows and in the whole BDT output range. The mass line shape of the peaking background is obtained from a simulated sample of doubly-misidentified $B^0_{(s)} \to h^+h'^-$ events and normalized to the number of events expected in the two search windows from data, $N_{B^0_s}$ and N_{B^0}. The BDT PDF of the peaking background is assumed to be the same as for the signal.

5.3. Signal PDFs

The BDT PDF for signal events is determined using an inclusive $B^0_{(s)} \to h^+h'^-$ sample. Only events which are triggered independently on the signal candidates have been considered (TIS events).

The number of $B^0_{(s)} \to h^+h'^-$ signal events in each BDT output bin is determined by fitting the hh' invariant mass distribution under the $\mu\mu$ mass hypothesis [15]. Fig. 2 shows the fit to the mass distribution of the full sample and for the three highest BDT output bins for $B^0_{(s)} \to h^+h'^-$ TIS events. The $B^0_{(s)} \to h^+h'^-$ exclusive decays, the combinatorial background and the physical background components are drawn under the fit to the data; the physical background is due to the partial reconstruction of three-body B meson decays.

In order to cross-check this result, two other fits have been performed on the same dataset. The signal line shape is parametrized either by a single or a double Crystal Ball function [10], the combinatorial background by an exponential function and the physical background by an ARGUS function [16]. In addition, exclusive $B^0_{(s)} \to \pi^-K^+, \pi^-\pi^+, K^-K^+$ channels, selected using the K–π separation capability of the RICH system, are used to cross-check the calibration of the BDT output both using the π^-K^+, $\pi^-\pi^+$, K^-K^+ inclusive yields without separating B and B^0_s and using the $B^0 \to K^+\pi^-$ exclusive channel alone. The maximum spread in the fractional yield obtained among the different models has been used as a systematic uncertainty in the signal BDT PDF. The BDT PDFs for signals and combinatorial background are shown in Fig. 3.

The invariant mass shape for the signal is parametrized as a Crystal Ball function. The mean value is determined using the $B^0 \to K^+\pi^-$ and $B^0_s \to K^+K^-$ exclusive channels and the transition point of the radiative tail is obtained from simulated events [6]. The central values are

$$m_{B^0_s} = 5358.0 \pm 1.0\ \text{MeV}/c^2,$$
$$m_{B^0} = 5272.0 \pm 1.0\ \text{MeV}/c^2.$$

The measured values of m_{B^0} and $m_{B^0_s}$ are 7–8 MeV/c^2 below the PDG values [13] due to the fact that the momentum scale is uncalibrated in the dataset used in this analysis. The mass resolutions are extracted from data with a linear interpolation between the measured resolution of charmonium and bottomonium resonances decaying into two muons: J/ψ, $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$. The mass line shapes for quarkonium resonances are

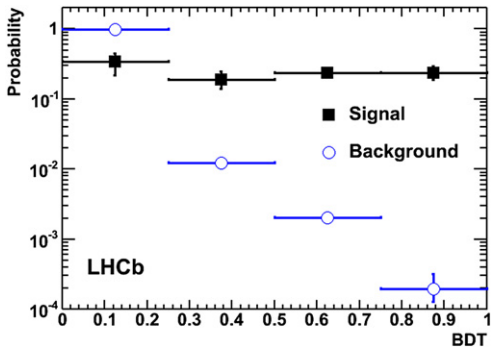

Fig. 3. BDT probability distribution functions of signal events (solid squares) and combinatorial background (open circles): the PDF for the signal is obtained from the inclusive sample of TIS $B^0_{(s)} \to h^+h'^-$ events, the PDF for the combinatorial background is obtained from the events in the mass sidebands.

shown in Fig. 4. Each resonance is fitted with two Crystal Ball functions with common mean value and common resolution but different parameterization of the tails. The background is fitted with an exponential function.

The results of the interpolation at the $m_{B^0_s}$ and m_{B^0} masses are

$$\sigma(m_{B^0_s}) = 24.6 \pm 0.2_{(\text{stat})} \pm 1.0_{(\text{syst})}\ \text{MeV}/c^2,$$

$$\sigma(m_{B^0}) = 24.3 \pm 0.2_{(\text{stat})} \pm 1.0_{(\text{syst})}\ \text{MeV}/c^2.$$

This result has been checked using both the fits to the $B^0_{(s)} \to h^+h'^-$ inclusive decay line shape and the $B^0 \to K^+\pi^-$ exclusive decay. The results are in agreement within the uncertainties.

6. Normalization

To estimate the signal branching fraction, the number of observed signal events is normalized to the number of events of a channel with a well-known branching fraction. Three complementary normalization channels are used: $B^+ \to J/\psi(\mu^+\mu^-)K^+$, $B^0_s \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ and $B^0 \to K^+\pi^-$. The first two channels have similar trigger and muon identification efficiencies to the signal but different number of particles in the final state. The third channel has a similar topology but is selected by different trigger lines.

The numbers of $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ candidates are translated into a branching fractions ($\mathcal{B}$) using the equation

$$\begin{aligned}\mathcal{B} &= \mathcal{B}_{\text{norm}} \times \frac{\epsilon^{\text{REC}}_{\text{norm}}\epsilon^{\text{SEL|REC}}_{\text{norm}}\epsilon^{\text{TRIG|SEL}}_{\text{norm}}}{\epsilon^{\text{REC}}_{\text{sig}}\epsilon^{\text{SEL|REC}}_{\text{sig}}\epsilon^{\text{TRIG|SEL}}_{\text{sig}}} \times \frac{f_{\text{norm}}}{f_{d(s)}} \times \frac{N_{B^0_{(s)}\to\mu^+\mu^-}}{N_{\text{norm}}} \\ &= \alpha^{\text{norm}}_{B^0_{(s)}\to\mu^+\mu^-} \times N_{B^0_{(s)}\to\mu^+\mu^-}, \end{aligned} \quad (3)$$

where $f_{d(s)}$ and f_{norm} are the probabilities that a b quark fragments into a $B^0_{(s)}$ and into the b hadron involved for the chosen normalization mode. LHCb has measured $f_s/f_d = 0.267^{+0.021}_{-0.020}$ [17]. $\mathcal{B}_{\text{norm}}$ is the branching fraction and N_{norm} is the number of selected events of the normalization channel. The efficiency is the product of three factors: ϵ^{REC} is the reconstruction efficiency of all the final state particles of the decay including the geometric acceptance of the detector; $\epsilon^{\text{SEL|REC}}$ is the selection efficiency for reconstructed events; $\epsilon^{\text{TRIG|SEL}}$ is the trigger efficiency for reconstructed and selected events. The subscript (sig, norm) indicates whether the efficiency refers to the signal or the normalization channel. Finally, $\alpha^{\text{norm}}_{B^0_{(s)}\to\mu^+\mu^-}$ is the normalization factor (or single event sensitivity) and $N_{B^0_{(s)}\to\mu^+\mu^-}$ the number of observed signal events.

For each normalization channel N_{norm} is obtained from a fit to the invariant mass distribution. The invariant mass distributions for reconstructed $B^+ \to J/\psi K^+$ and $B^0_s \to J/\psi\phi$ candidates are shown in Fig. 5, while the $B^0 \to K^+\pi^-$ yield is obtained from the full $B^0_{(s)} \to h^+h'^-$ fit as shown in the top left of Fig. 2.

The numbers used to calculate the normalization factors are summarized in Table 1. A weighted average of the three normalization channels, assuming the tracking and trigger uncertainties to be correlated between the two J/ψ normalization channels and the uncertainty on f_d/f_s to be correlated between the $B^+ \to J/\psi K^+$ and $B^0 \to K^+\pi^-$, gives

$$\alpha^{\text{norm}}_{B^0_s\to\mu^+\mu^-} = (8.38 \pm 0.74) \times 10^{-10},$$

$$\alpha^{\text{norm}}_{B^0\to\mu^+\mu^-} = (2.20 \pm 0.11) \times 10^{-10}.$$

These normalization factors are used to determine the limits.

7. Results

The results for $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are summarized in Tables 2 and 3 respectively and in each of the bins the expected number of combinatorial background, peaking background, signal events, with the SM prediction assumed, is shown together with the observations on the data. The uncertainties in the signal and

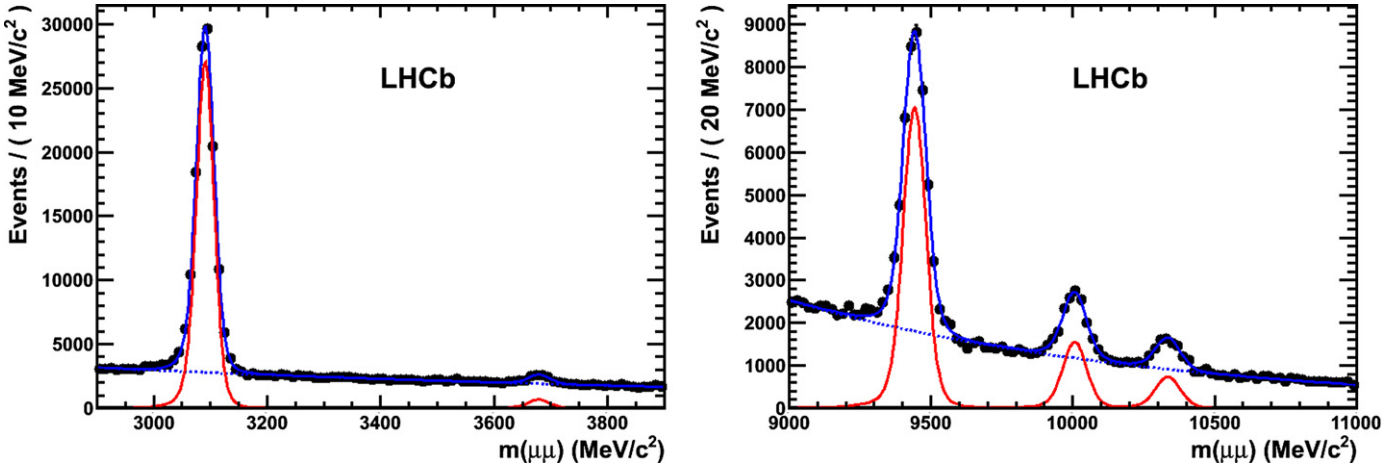

Fig. 4. Di-muon invariant mass spectrum in the ranges (2.9–3.9) GeV/c^2 (left) and (9–11) MeV/c^2 (right).

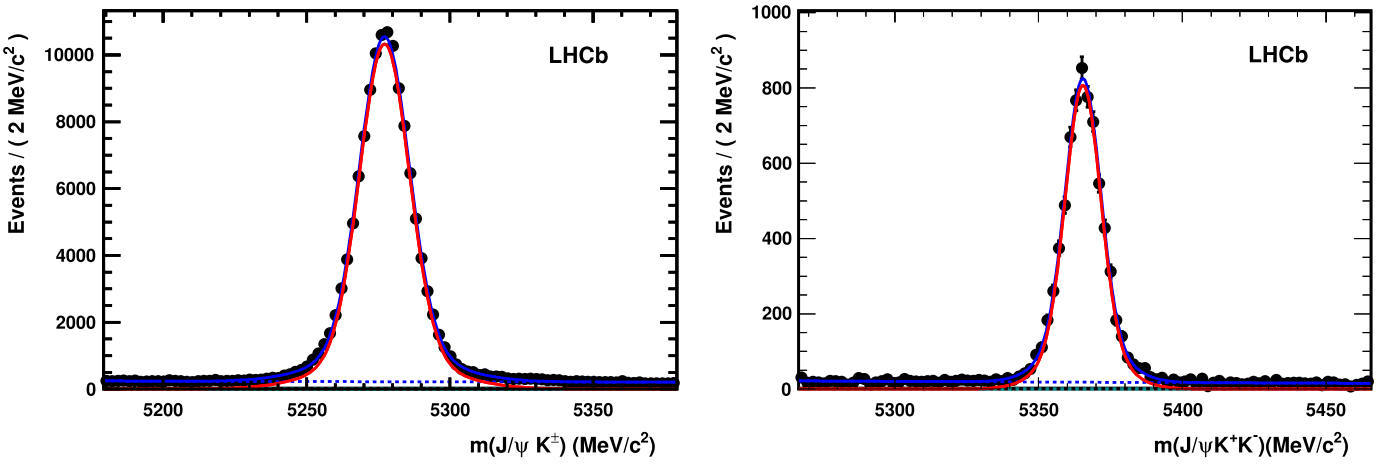

Fig. 5. Invariant mass distributions of the $B^+ \to J/\psi K^+$ (left) and $B_s^0 \to J/\psi\phi$ (right) candidates used in the normalization procedure.

Table 1
Summary of the quantities and their uncertainties required to calculate the normalization factors ($\alpha^{\rm norm}_{B^0_{(s)}\to\mu^+\mu^-}$) for the three normalization channels considered. The branching fractions are taken from Refs. [13,18]. The trigger efficiency and the number of $B^0 \to K^+\pi^-$ candidates correspond to TIS events.

	$\mathcal{B}$ ($\times 10^{-5}$)	$\frac{\epsilon^{\rm REC}_{\rm norm}\epsilon^{\rm SEL|REC}_{\rm norm}}{\epsilon^{\rm REC}_{\rm sig}\epsilon^{\rm SEL|REC}_{\rm sig}}$	$\frac{\epsilon^{\rm TRIG|SEL}_{\rm norm}}{\epsilon^{\rm TRIG|SEL}_{\rm sig}}$	$N_{\rm norm}$	$\alpha^{\rm norm}_{B^0\to\mu^+\mu^-}$ ($\times 10^{-10}$)	$\alpha^{\rm norm}_{B^0_s\to\mu^+\mu^-}$ ($\times 10^{-9}$)
$B^+ \to J/\psi K^+$	6.01 ± 0.21	0.48 ± 0.014	0.95 ± 0.01	$124\,518 \pm 2025$	2.23 ± 0.11	0.83 ± 0.08
$B_s^0 \to J/\psi\phi$	3.4 ± 0.9	0.24 ± 0.014	0.95 ± 0.01	6940 ± 93	2.96 ± 0.84	1.11 ± 0.30
$B^0 \to K^+\pi^-$	1.94 ± 0.06	0.86 ± 0.02	0.049 ± 0.004	4146 ± 608	1.98 ± 0.34	0.74 ± 0.14

Table 2
Expected combinatorial background events, expected peaking ($B^0_{(s)} \to h^+h'^-$) background events, expected signal events assuming the SM branching fraction prediction, and observed events in the $B_s^0 \to \mu^+\mu^-$ search window.

Invariant mass [MeV/c^2]		BDT			
		0–0.25	0.25–0.5	0.5–0.75	0.75–1
5298–5318	Expected comb. bkg	$575.5^{+6.5}_{-6.0}$	$6.96^{+0.63}_{-0.57}$	$1.19^{+0.39}_{-0.35}$	$0.111^{+0.083}_{-0.066}$
	Expected peak. bkg	$0.126^{+0.037}_{-0.030}$	$0.124^{+0.037}_{-0.030}$	$0.124^{+0.037}_{-0.030}$	$0.127^{+0.038}_{-0.031}$
	Expected signal	$0.059^{+0.023}_{-0.022}$	$0.0329^{+0.0128}_{-0.0095}$	$0.0415^{+0.0120}_{-0.0085}$	$0.0411^{+0.0135}_{-0.0099}$
	Observed	533	10	1	0
5318–5338	Expected comb. bkg	$566.8^{+6.3}_{-5.8}$	$6.90^{+0.61}_{-0.55}$	$1.16^{+0.38}_{-0.34}$	$0.109^{+0.079}_{-0.063}$
	Expected peak. bkg	$0.052^{+0.023}_{-0.018}$	$0.054^{+0.026}_{-0.019}$	$0.052^{+0.024}_{-0.018}$	$0.051^{+0.023}_{-0.018}$
	Expected signal	$0.205^{+0.073}_{-0.074}$	$0.114^{+0.040}_{-0.031}$	$0.142^{+0.036}_{-0.025}$	$0.142^{+0.042}_{-0.031}$
	Observed	525	9	0	1
5338–5358	Expected comb. bkg	$558.2^{+6.1}_{-5.6}$	$6.84^{+0.59}_{-0.54}$	$1.14^{+0.37}_{-0.33}$	$0.106^{+0.075}_{-0.060}$
	Expected peak. bkg	$0.024^{+0.028}_{-0.012}$	$0.025^{+0.026}_{-0.012}$	$0.024^{+0.027}_{-0.012}$	$0.025^{+0.025}_{-0.012}$
	Expected signal	$0.38^{+0.14}_{-0.14}$	$0.213^{+0.075}_{-0.058}$	$0.267^{+0.065}_{-0.047}$	$0.265^{+0.077}_{-0.058}$
	Observed	561	6	2	1
5358–5378	Expected comb. bkg	$549.8^{+6.0}_{-5.4}$	$6.77^{+0.57}_{-0.52}$	$1.11^{+0.36}_{-0.32}$	$0.103^{+0.073}_{-0.057}$
	Expected peak. bkg	$0.0145^{+0.0220}_{-0.0091}$	$0.0151^{+0.0230}_{-0.0091}$	$0.0153^{+0.0232}_{-0.0098}$	$0.015^{+0.023}_{-0.010}$
	Expected signal	$0.38^{+0.14}_{-0.14}$	$0.213^{+0.075}_{-0.057}$	$0.267^{+0.065}_{-0.047}$	$0.265^{+0.077}_{-0.057}$
	Observed	515	7	0	0
5378–5398	Expected comb. bkg	$541.5^{+5.8}_{-5.3}$	$6.71^{+0.55}_{-0.51}$	$1.09^{+0.34}_{-0.31}$	$0.101^{+0.070}_{-0.054}$
	Expected peak. bkg	$0.0115^{+0.0175}_{-0.0086}$	$0.0116^{+0.0177}_{-0.0090}$	$0.0118^{+0.0179}_{-0.0090}$	$0.0118^{+0.0179}_{-0.0088}$
	Expected signal	$0.204^{+0.073}_{-0.074}$	$0.114^{+0.040}_{-0.031}$	$0.142^{+0.036}_{-0.026}$	$0.141^{+0.042}_{-0.031}$
	Observed	547	10	1	1
5398–5418	Expected comb. bkg	$533.4^{+5.7}_{-5.2}$	$6.65^{+0.53}_{-0.49}$	$1.07^{+0.34}_{-0.30}$	$0.098^{+0.068}_{-0.051}$
	Expected peak. bkg	$0.0089^{+0.0136}_{-0.0065}$	$0.0088^{+0.0133}_{-0.0066}$	$0.0091^{+0.0138}_{-0.0070}$	$0.0090^{+0.0137}_{-0.0065}$
	Expected signal	$0.058^{+0.024}_{-0.021}$	$0.0323^{+0.0128}_{-0.0093}$	$0.0407^{+0.0120}_{-0.0087}$	$0.0402^{+0.0137}_{-0.0097}$
	Observed	501	4	1	0

Table 3
Expected combinatorial background events, expected peaking ($B^0_{(s)} \to h^+h'^-$) background events, expected $B^0 \to \mu^+\mu^-$ signal events assuming the SM branching fraction, expected cross-feed events from $B^0_s \to \mu^+\mu^-$ assuming the SM branching fraction and observed events in the $B^0 \to \mu^+\mu^-$ search window.

Invariant mass [MeV/c^2]		BDT			
		0–0.25	0.25–0.5	0.5–0.75	0.75–1
5212–5232	Expected comb. bkg	$614.2^{+7.5}_{-7.0}$	$7.23^{+0.77}_{-0.68}$	$1.31^{+0.46}_{-0.40}$	$0.123^{+0.107}_{-0.072}$
	Expected peak. bkg	$0.203^{+0.038}_{-0.034}$	$0.206^{+0.038}_{-0.034}$	$0.203^{+0.037}_{-0.034}$	$0.205^{+0.038}_{-0.034}$
	Cross-feed	$0.0056^{+0.0021}_{-0.0020}$	$0.00312^{+0.00119}_{-0.00087}$	$0.00391^{+0.00107}_{-0.00078}$	$0.00387^{+0.00122}_{-0.00092}$
	Expected signal	$0.0070^{+0.0027}_{-0.0026}$	$0.0039^{+0.0015}_{-0.0011}$	$0.0049^{+0.0014}_{-0.0010}$	$0.0048^{+0.0016}_{-0.0012}$
	Observed	554	6	0	2
5232–5252	Expected comb. bkg	$605.0^{+7.2}_{-6.8}$	$7.17^{+0.74}_{-0.65}$	$1.29^{+0.44}_{-0.39}$	$0.121^{+0.102}_{-0.072}$
	Expected peak. bkg	$0.281^{+0.056}_{-0.049}$	$0.279^{+0.056}_{-0.049}$	$0.280^{+0.056}_{-0.049}$	$0.280^{+0.058}_{-0.050}$
	Cross-feed	$0.0071^{+0.0027}_{-0.0026}$	$0.0039^{+0.0015}_{-0.0011}$	$0.00496^{+0.00134}_{-0.00099}$	$0.0049^{+0.0016}_{-0.0012}$
	Expected signal	$0.0241^{+0.0086}_{-0.0087}$	$0.0135^{+0.0048}_{-0.0037}$	$0.0169^{+0.0042}_{-0.0031}$	$0.0167^{+0.0050}_{-0.0037}$
	Observed	556	4	2	1
5252–5272	Expected comb. bkg	$595.9^{+7.0}_{-6.5}$	$7.10^{+0.71}_{-0.63}$	$1.26^{+0.42}_{-0.37}$	$0.119^{+0.097}_{-0.072}$
	Expected peak. bkg	$0.323^{+0.075}_{-0.061}$	$0.326^{+0.074}_{-0.061}$	$0.324^{+0.072}_{-0.060}$	$0.325^{+0.075}_{-0.062}$
	Cross-feed	$0.0097^{+0.0036}_{-0.0035}$	$0.0054^{+0.0021}_{-0.0015}$	$0.0068^{+0.0018}_{-0.0013}$	$0.0067^{+0.0021}_{-0.0016}$
	Expected signal	$0.045^{+0.016}_{-0.016}$	$0.0252^{+0.0088}_{-0.0067}$	$0.0317^{+0.0077}_{-0.0057}$	$0.0313^{+0.0093}_{-0.0068}$
	Observed	588	11	1	0
5272–5292	Expected comb. bkg	$586.9^{+6.7}_{-6.3}$	$7.04^{+0.68}_{-0.60}$	$1.23^{+0.41}_{-0.36}$	$0.117^{+0.092}_{-0.071}$
	Expected peak. bkg	$0.252^{+0.058}_{-0.047}$	$0.252^{+0.056}_{-0.046}$	$0.253^{+0.059}_{-0.048}$	$0.250^{+0.056}_{-0.046}$
	Cross-feed	$0.0154^{+0.0058}_{-0.0055}$	$0.0086^{+0.0033}_{-0.0024}$	$0.0108^{+0.0029}_{-0.0021}$	$0.0106^{+0.0033}_{-0.0025}$
	Expected signal	$0.045^{+0.016}_{-0.016}$	$0.0251^{+0.0089}_{-0.0067}$	$0.0317^{+0.0077}_{-0.0057}$	$0.0313^{+0.0092}_{-0.0069}$
	Observed	616	5	2	1
5292–5312	Expected comb. bkg	$578.1^{+6.5}_{-6.1}$	$6.98^{+0.66}_{-0.58}$	$1.20^{+0.39}_{-0.35}$	$0.114^{+0.087}_{-0.067}$
	Expected peak. bkg	$0.124^{+0.023}_{-0.021}$	$0.124^{+0.023}_{-0.021}$	$0.123^{+0.023}_{-0.021}$	$0.124^{+0.023}_{-0.021}$
	Cross-feed	$0.038^{+0.015}_{-0.014}$	$0.0214^{+0.0086}_{-0.0061}$	$0.0270^{+0.0080}_{-0.0056}$	$0.0266^{+0.0089}_{-0.0064}$
	Expected signal	$0.0241^{+0.0086}_{-0.0087}$	$0.0134^{+0.0048}_{-0.0036}$	$0.0169^{+0.0042}_{-0.0030}$	$0.0167^{+0.0050}_{-0.0037}$
	Observed	549	7	0	0
5312–5332	Expected comb. bkg	$569.3^{+6.3}_{-5.9}$	$6.92^{+0.63}_{-0.57}$	$1.18^{+0.38}_{-0.34}$	$0.111^{+0.083}_{-0.064}$
	Expected peak. bkg	$0.047^{+0.023}_{-0.012}$	$0.047^{+0.022}_{-0.012}$	$0.047^{+0.021}_{-0.012}$	$0.047^{+0.021}_{-0.012}$
	Cross-feed	$0.149^{+0.055}_{-0.054}$	$0.083^{+0.031}_{-0.022}$	$0.104^{+0.027}_{-0.019}$	$0.103^{+0.031}_{-0.023}$
	Expected signal	$0.0068^{+0.0028}_{-0.0026}$	$0.0038^{+0.0015}_{-0.0011}$	$0.0048^{+0.0014}_{-0.0010}$	$0.0048^{+0.0016}_{-0.0012}$
	Observed	509	10	1	1

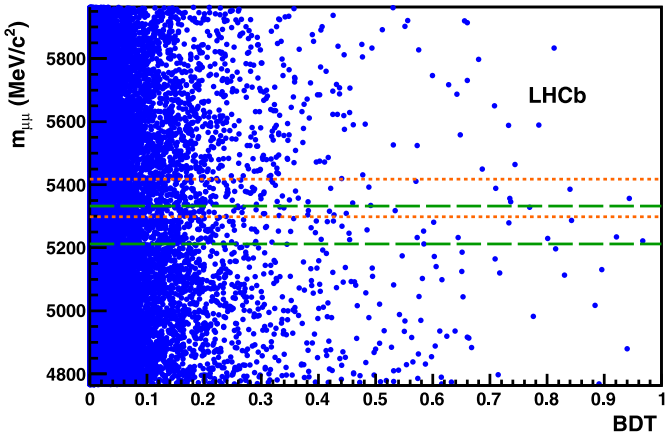

Fig. 6. Distribution of selected di-muon events in the invariant mass–BDT plane. The orange short-dashed (green long-dashed) lines indicate the ± 60 MeV/c^2 search window around the mean B^0_s (B^0) mass. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

background PDFs and normalization factors are used to compute the uncertainties on the background and signal predictions.

The two-dimensional (mass, BDT) distribution of selected events can be seen in Fig. 6. The distribution of the invariant mass in the four BDT bins is shown in Fig. 7 for $B^0_s \to \mu^+\mu^-$ and in Fig. 8 for $B^0 \to \mu^+\mu^-$ selected candidates.

The compatibility of the distribution of events inside the search window in the invariant mass–BDT plane with a given branching fraction hypothesis is evaluated using the CL_s method [11]. This method provides three estimators: CL_{s+b}, a measure of the compatibility of the observed distribution with the signal and background hypotheses, CL_b, a measure of the compatibility with the background-only hypothesis and CL_s, a measure of the compatibility of the observed distribution with the signal and background hypotheses normalized to the background-only hypothesis.

The expected CL_s values are shown in Fig. 9 for $B^0_s \to \mu^+\mu^-$ and for $B^0 \to \mu^+\mu^-$ as dashed black lines under the hypothesis that background and SM events are observed. The shaded areas cover the region of $\pm 1\sigma$ of compatible observations. The observed values of CL_s as a function of the assumed branching ratio is shown as dotted blue lines on both plots.

The expected limits and the measured limits for $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ at 90% and 95% CL are shown in Tables 4 and 5, respectively. For the $B^0_s \to \mu^+\mu^-$ decay, the expected limits are computed allowing the presence of $B^0_s \to \mu^+\mu^-$ events according to the SM branching fraction. For the $B^0 \to \mu^+\mu^-$ decay the expected limit is computed in the background-only hypothesis and also allowing the presence of $B^0 \to \mu^+\mu^-$ events with the SM rate: the two results are identical. In the determination of the

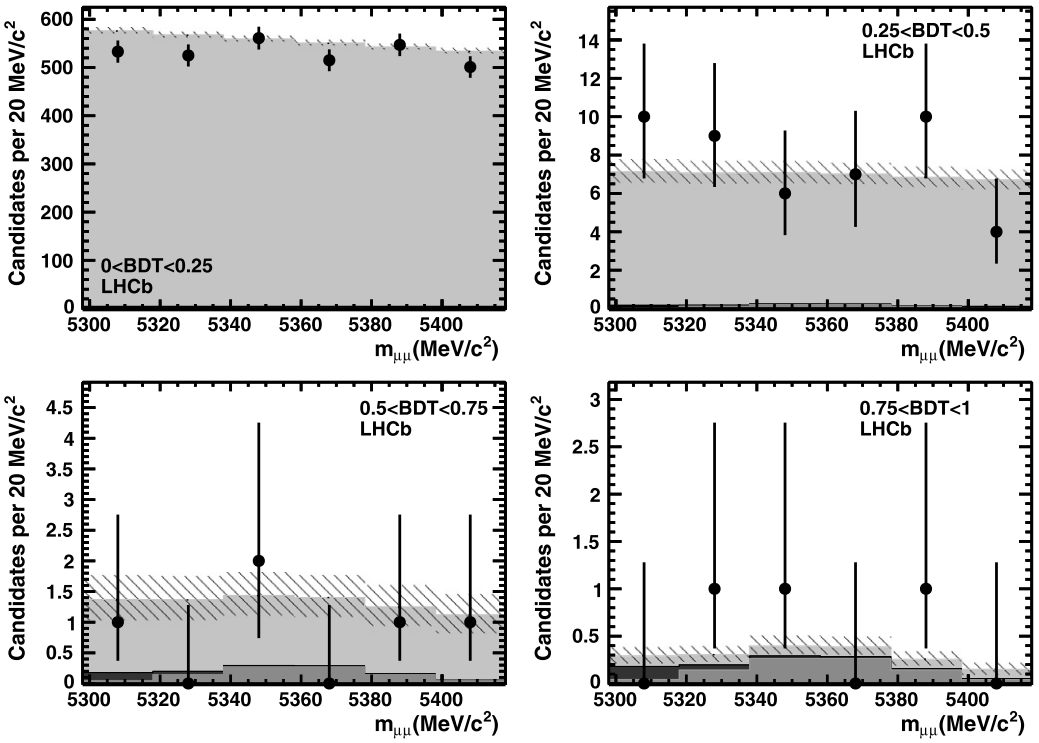

Fig. 7. Distribution of selected di-muon events in the $B_s^0 \to \mu^+\mu^-$ mass window for the four BDT output bins. The black dots are data, the light grey histogram shows the contribution of the combinatorial background, the black filled histogram shows the contribution of the $B_{(s)}^0 \to h^+h'^-$ background and the dark grey filled histogram the contribution of $B_s^0 \to \mu^+\mu^-$ signal events according to the SM rate. The hatched area depicts the uncertainty on the sum of the expected contributions.

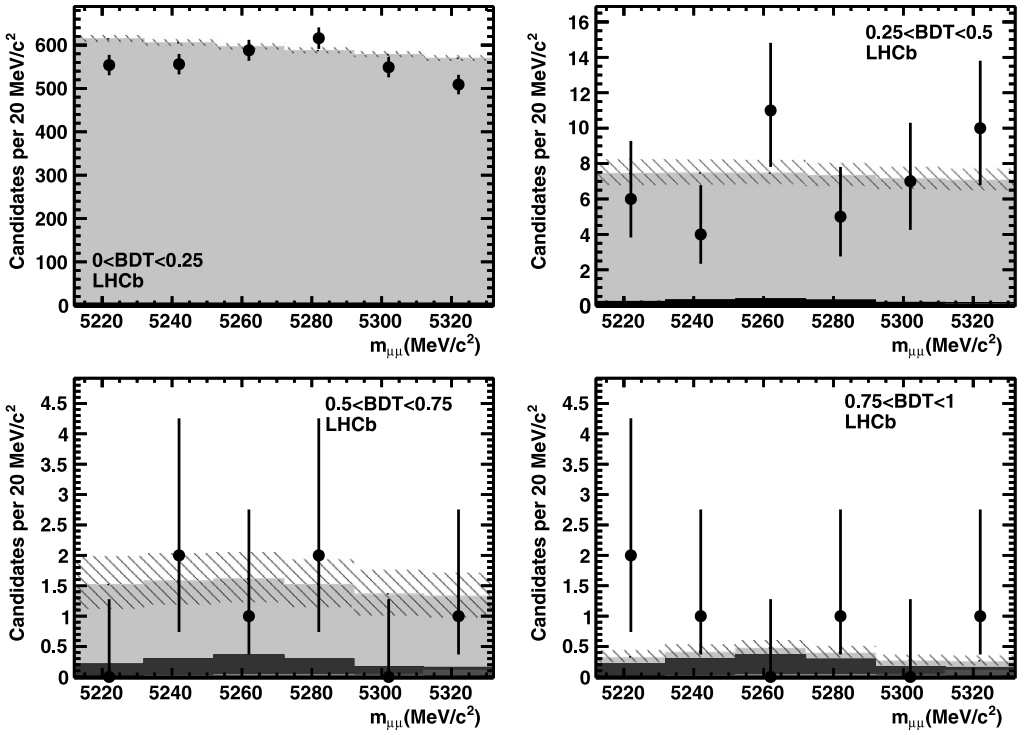

Fig. 8. Distribution of selected di-muon events in the $B^0 \to \mu^+\mu^-$ mass window for the four BDT output bins. The black dots are data, the light grey histogram shows the contribution of the combinatorial background, the black filled histogram shows the contribution of the $B_{(s)}^0 \to h^+h'^-$ background and the dark grey filled histogram shows the cross-feed of $B_s^0 \to \mu^+\mu^-$ events in the B^0 mass window assuming the SM rate. The hatched area depicts the uncertainty on the sum of the expected contributions.

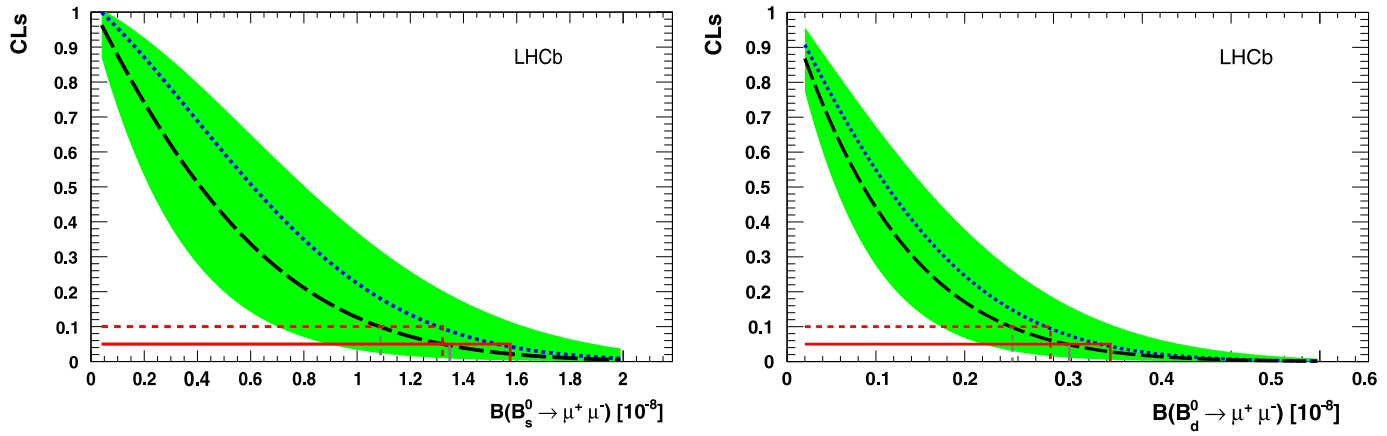

Fig. 9. CL_s as a function of the assumed $\mathcal{B}$. Expected (observed) values are shown by dashed black (dotted blue) lines. The expected CL_s values have been computed assuming a signal yield corresponding to the SM branching fractions. The green (grey) shaded areas cover the region of $\pm 1\sigma$ of compatible observations. The measured upper limits at 90% and 95% CL are also shown. Left: $B_s^0 \to \mu^+\mu^-$, right: $B^0 \to \mu^+\mu^-$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

Table 4
Expected and observed limits on the $B_s^0 \to \mu^+\mu^-$ branching fraction for the 2011 data and for the combination of 2010 and 2011 data. The expected limits are computed allowing the presence of $B_s^0 \to \mu^+\mu^-$ events according to the SM branching fraction.

		at 90% CL	at 95% CL	CL_b
2011	expected limit	1.1×10^{-8}	1.4×10^{-8}	
	observed limit	1.3×10^{-8}	1.6×10^{-8}	0.95
2010 + 2011	expected limit	1.0×10^{-8}	1.3×10^{-8}	
	observed limit	1.2×10^{-8}	1.4×10^{-8}	0.93

Table 5
Expected and observed limits on the $B^0 \to \mu^+\mu^-$ branching fraction for 2011 data and for the combination of 2010 and 2011 data. The expected limits are computed in the background only hypothesis.

		at 90% CL	at 95% CL	CL_b
2011	expected limit	2.5×10^{-9}	3.2×10^{-9}	
	observed limit	3.0×10^{-9}	3.6×10^{-9}	0.68
2010 + 2011	expected limit	2.4×10^{-9}	3.0×10^{-9}	
	observed limit	2.6×10^{-9}	3.2×10^{-9}	0.61

limits, the cross-feed of $B_s^0 \to \mu^+\mu^-$ ($B^0 \to \mu^+\mu^-$) events in the B^0 (B_s^0) mass window has been taken into account assuming the SM rates.

The observed CL_b values are shown in the same tables. The comparison of the observed distribution of events with the expected background distribution results in a p-value $(1 - CL_b)$ of 5% for the $B_s^0 \to \mu^+\mu^-$ and 32% for the $B^0 \to \mu^+\mu^-$ decay. For the $B_s^0 \to \mu^+\mu^-$ decay, the probability that the observed events are compatible with the sum of expected background events and signal events according to the SM rate is measured by $1 - CL_{s+b}$ and it is 33%.

The result obtained in 2011 with 0.37 fb^{-1} has been combined with the published result based on ~ 37 pb^{-1} [6]. The expected and observed limits for 90% and 95% CL for the combined results are shown in Table 4 for the $B_s^0 \to \mu^+\mu^-$ decay and in Table 5 for the $B^0 \to \mu^+\mu^-$ decay.

8. Conclusions

With 0.37 fb^{-1} of integrated luminosity, a search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ has been performed and sensitivities better than the existing limits have been obtained. The observed events in the B_s^0 and in the B^0 mass windows are compatible with the background expectations at 5% and 32% confidence level, respectively. For the $B_s^0 \to \mu^+\mu^-$ decay, the probability that the observed events are compatible with the sum of expected background events and signal events according to the SM rate is 33%. The upper limits for the branching fractions are evaluated to be

$$\mathcal{B}\left(B_s^0 \to \mu^+\mu^-\right) < 1.3\ (1.6) \times 10^{-8} \quad \text{at 90\% (95\%) CL},$$
$$\mathcal{B}\left(B^0 \to \mu^+\mu^-\right) < 3.0\ (3.6) \times 10^{-9} \quad \text{at 90\% (95\%) CL}.$$

The $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ upper limits have been combined with those published previously by LHCb [6] and the results are

$$\mathcal{B}\left(B_s^0 \to \mu^+\mu^-\right)(2010 + 2011) < 1.2\ (1.4) \times 10^{-8} \quad \text{at 90\% (95\%) CL},$$
$$\mathcal{B}\left(B^0 \to \mu^+\mu^-\right)(2010 + 2011) < 2.6\ (3.2) \times 10^{-9} \quad \text{at 90\% (95\%) CL}.$$

The above 90% (95%) CL upper limits are still about 3.8 (4.4) times the SM branching fractions for the B_s^0 and 26 (32) times for the B^0. These results represent the best upper limits to date.

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at CERN and at the LHCb institutes, and acknowledge support from the National Agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); CERN; NSFC (China); CNRS/IN2P3 (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); SCSR (Poland); ANCS (Romania); MinES of Russia and Rosatom (Russia); MICINN, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); NSF (USA). We also acknowledge the support received from the ERC under FP7 and the Region Auvergne.

References

[1] A.J. Buras, M.V. Carlucci, S. Gori, G. Isidori, JHEP 1010 (2010) 009, arXiv:1005.5310;
A.J. Buras, Acta Phys. Polon. B 41 (2010) 2487, arXiv:1012.1447.

[2] S. Choudhury, N. Gaur, Phys. Lett. B 451 (1999) 86, arXiv:hep-ph/9810307;
C. Hamzaoui, M. Pospelov, M. Toharia, Phys. Rev. D 59 (1999) 095005, arXiv:hep-ph/9807350;
K. Babu, C.F. Kolda, Phys. Rev. Lett. 84 (2000) 228, arXiv:hep-ph/9909476;
L.J. Hall, R. Rattazzi, U. Sarid, Phys. Rev. D 50 (1994) 7048, arXiv:hep-ph/9306309;
C.-S. Huang, W. Liao, Q.-S. Yan, Phys. Rev. D 59 (1999) 011701, arXiv:hep-ph/9803460.

[3] DØ Collaboration, V.M. Abazov, et al., Phys. Lett. B 693 (2010) 539, arXiv:1006.3469.

[4] CDF Collaboration, T. Aaltonen, et al., Phys. Rev. Lett. 107 (2011) 191801, arXiv:1107.2304.

[5] CMS Collaboration, S. Chatrchyan, et al., Phys. Rev. Lett. 107 (2011) 191802, arXiv:1107.5834.

[6] LHCb Collaboration, R. Aaij, et al., Phys. Lett. B 699 (2011) 330, arXiv:1103.2465.

[7] LHCb Collaboration, A. Alves, et al., JINST 3 (2008) S08005, and references therein.

[8] LHCb Collaboration, R. Aaij, et al., Phys. Lett. B 694 (2010) 209, arXiv:1009.2731.

[9] P. Speckmayer, A. Hocker, J. Stelzer, H. Voss, J. Phys. Conf. Ser. 219 (2010) 032057.

[10] T. Skwarnicki, A study of the radiative cascade transitions between the Υ and Υ' resonances, PhD thesis, DESY F31-86-02, 1986, Appendix E.

[11] A. Read, J. Phys. G 28 (2002) 2693.

[12] S. Kullback, R.A. Leibler, Ann. Math. Statist. 22 (1951) 79;
S. Kullback, Amer. Statist. 41 (1987) 340, The use of the Kullback–Leibler distance at LHCb is described in M. Needham, Clone track identification using the Kullback–Leibler distance, LHCb-2008-002.

[13] Particle Data Group, K. Nakamura, et al., J. Phys. G 37 (2010) 075021.

[14] CDF Collaboration, A. Abulencia, et al., Phys. Rev. Lett. 95 (2005) 221805, arXiv:hep-ex/0508036.

[15] A. Carbone, et al., Invariant mass line shape of $B \to h^+h^-$ decays at LHCb, LHCb-PUB-2009-031.

[16] ARGUS Collaboration, H. Albrecht, et al., Phys. Lett. B 241 (1990) 278.

[17] LHCb Collaboration, R. Aaij, et al., Phys. Rev. Lett. 107 (2011) 211801, arXiv:1106.4435.

[18] Belle Collaboration, R. Louvot, $\Upsilon(5S)$ results at Belle, arXiv:0905.4345.

LHCb Collaboration

R. Aaij [23], C. Abellan Beteta [35,n], B. Adeva [36], M. Adinolfi [42], C. Adrover [6], A. Affolder [48], Z. Ajaltouni [5], J. Albrecht [37], F. Alessio [37], M. Alexander [47], G. Alkhazov [29], P. Alvarez Cartelle [36], A.A. Alves Jr. [22], S. Amato [2], Y. Amhis [38], J. Anderson [39], R.B. Appleby [50], O. Aquines Gutierrez [10], F. Archilli [18,37], L. Arrabito [53], A. Artamonov [34], M. Artuso [52,37], E. Aslanides [6], G. Auriemma [22,m], S. Bachmann [11], J.J. Back [44], D.S. Bailey [50], V. Balagura [30,37], W. Baldini [16], R.J. Barlow [50], C. Barschel [37], S. Barsuk [7], W. Barter [43], A. Bates [47], C. Bauer [10], Th. Bauer [23], A. Bay [38], I. Bediaga [1], S. Belogurov [30], K. Belous [34], I. Belyaev [30,37], E. Ben-Haim [8], M. Benayoun [8], G. Bencivenni [18], S. Benson [46], J. Benton [42], R. Bernet [39], M.-O. Bettler [17,*], M. van Beuzekom [23], A. Bien [11], S. Bifani [12], T. Bird [50], A. Bizzeti [17,h], P.M. Bjørnstad [50], T. Blake [37], F. Blanc [38], C. Blanks [49], J. Blouw [11], S. Blusk [52], A. Bobrov [33], V. Bocci [22], A. Bondar [33], N. Bondar [29], W. Bonivento [15], S. Borghi [47,50], A. Borgia [52], T.J.V. Bowcock [48], C. Bozzi [16], T. Brambach [9], J. van den Brand [24], J. Bressieux [38], D. Brett [50], M. Britsch [10], T. Britton [52], N.H. Brook [42], H. Brown [48], A. Büchler-Germann [39], I. Burducea [28], A. Bursche [39], J. Buytaert [37], S. Cadeddu [15], O. Callot [7], M. Calvi [20,j], M. Calvo Gomez [35,n], A. Camboni [35], P. Campana [18,37], A. Carbone [14], G. Carboni [21,k], R. Cardinale [19,37,i], A. Cardini [15], L. Carson [49], K. Carvalho Akiba [2], G. Casse [48], M. Cattaneo [37], Ch. Cauet [9], M. Charles [51], Ph. Charpentier [37], N. Chiapolini [39], K. Ciba [37], X. Cid Vidal [36], G. Ciezarek [49], P.E.L. Clarke [46,37], M. Clemencic [37], H.V. Cliff [43], J. Closier [37], C. Coca [28], V. Coco [23], J. Cogan [6], P. Collins [37], A. Comerma-Montells [35], F. Constantin [28], G. Conti [38], A. Contu [51], A. Cook [42], M. Coombes [42], G. Corti [37], G.A. Cowan [38], R. Currie [46], B. D'Almagne [7], C. D'Ambrosio [37], P. David [8], P.N.Y. David [23], I. De Bonis [4], S. De Capua [21,k], M. De Cian [39], F. De Lorenzi [12], J.M. De Miranda [1], L. De Paula [2], P. De Simone [18], D. Decamp [4], M. Deckenhoff [9], H. Degaudenzi [38,37], M. Deissenroth [11], L. Del Buono [8], C. Deplano [15], D. Derkach [14,37], O. Deschamps [5], F. Dettori [24], J. Dickens [43], H. Dijkstra [37], P. Diniz Batista [1], F. Domingo Bonal [35,n], S. Donleavy [48], F. Dordei [11], P. Dornan [49], A. Dosil Suárez [36], D. Dossett [44], A. Dovbnya [40], F. Dupertuis [38], R. Dzhelyadin [34], A. Dziurda [25], S. Easo [45], U. Egede [49], V. Egorychev [30], S. Eidelman [33], D. van Eijk [23], F. Eisele [11], S. Eisenhardt [46], R. Ekelhof [9], L. Eklund [47], Ch. Elsasser [39], D. Elsby [55], D. Esperante Pereira [36], L. Estève [43], A. Falabella [16,14,e], E. Fanchini [20,j], C. Färber [11], G. Fardell [46], C. Farinelli [23], S. Farry [12], V. Fave [38], V. Fernandez Albor [36], M. Ferro-Luzzi [37], S. Filippov [32], C. Fitzpatrick [46], M. Fontana [10], F. Fontanelli [19,i], R. Forty [37], M. Frank [37], C. Frei [37], M. Frosini [17,37,f], S. Furcas [20], A. Gallas Torreira [36], D. Galli [14,c], M. Gandelman [2], P. Gandini [51], Y. Gao [3], J.-C. Garnier [37], J. Garofoli [52], J. Garra Tico [43], L. Garrido [35], D. Gascon [35], C. Gaspar [37], N. Gauvin [38], M. Gersabeck [37], T. Gershon [44,37], Ph. Ghez [4], V. Gibson [43], V.V. Gligorov [37], C. Göbel [54], D. Golubkov [30], A. Golutvin [49,30,37], A. Gomes [2], H. Gordon [51], M. Grabalosa Gándara [35], R. Graciani Diaz [35], L.A. Granado Cardoso [37], E. Graugés [35], G. Graziani [17], A. Grecu [28], E. Greening [51], S. Gregson [43], B. Gui [52], E. Gushchin [32], Yu. Guz [34], T. Gys [37], G. Haefeli [38], C. Haen [37], S.C. Haines [43], T. Hampson [42], S. Hansmann-Menzemer [11], R. Harji [49], N. Harnew [51], J. Harrison [50], P.F. Harrison [44], J. He [7], V. Heijne [23], K. Hennessy [48], P. Henrard [5],

J.A. Hernando Morata [36], E. van Herwijnen [37], E. Hicks [48], K. Holubyev [11], P. Hopchev [4], W. Hulsbergen [23], P. Hunt [51], T. Huse [48], R.S. Huston [12], D. Hutchcroft [48], D. Hynds [47], V. Iakovenko [41], P. Ilten [12], J. Imong [42], R. Jacobsson [37], A. Jaeger [11], M. Jahjah Hussein [5], E. Jans [23], F. Jansen [23], P. Jaton [38], B. Jean-Marie [7], F. Jing [3], M. John [51], D. Johnson [51], C.R. Jones [43], B. Jost [37], M. Kaballo [9], S. Kandybei [40], M. Karacson [37], T.M. Karbach [9], J. Keaveney [12], I.R. Kenyon [55], U. Kerzel [37], T. Ketel [24], A. Keune [38], B. Khanji [6], Y.M. Kim [46], M. Knecht [38], P. Koppenburg [23], A. Kozlinskiy [23], L. Kravchuk [32], K. Kreplin [11], M. Kreps [44], G. Krocker [11], P. Krokovny [11], F. Kruse [9], K. Kruzelecki [37], M. Kucharczyk [20,25,37,j], T. Kvaratskheliya [30,37], V.N. La Thi [38], D. Lacarrere [37], G. Lafferty [50], A. Lai [15], D. Lambert [46], R.W. Lambert [24], E. Lanciotti [37], G. Lanfranchi [18,*], C. Langenbruch [11], T. Latham [44], C. Lazzeroni [55], R. Le Gac [6], J. van Leerdam [23], J.-P. Lees [4], R. Lefèvre [5], A. Leflat [31,37], J. Lefrançois [7], O. Leroy [6], T. Lesiak [25], L. Li [3], L. Li Gioi [5], M. Lieng [9], M. Liles [48], R. Lindner [37], C. Linn [11], B. Liu [3], G. Liu [37], J.H. Lopes [2], E. Lopez Asamar [35], N. Lopez-March [38], H. Lu [38,3], J. Luisier [38], A. Mac Raighne [47], F. Machefert [7], I.V. Machikhiliyan [4,30], F. Maciuc [10], O. Maev [29,37], J. Magnin [1], S. Malde [51], R.M.D. Mamunur [37], G. Manca [15,d], G. Mancinelli [6], N. Mangiafave [43], U. Marconi [14], R. Märki [38], J. Marks [11], G. Martellotti [22], A. Martens [8], L. Martin [51], A. Martín Sánchez [7], D. Martinez Santos [37], A. Massafferri [1], Z. Mathe [12], C. Matteuzzi [20], M. Matveev [29], E. Maurice [6], B. Maynard [52], A. Mazurov [16,32,37], G. McGregor [50], R. McNulty [12], C. Mclean [14], M. Meissner [11], M. Merk [23], J. Merkel [9], R. Messi [21,k], S. Miglioranzi [37], D.A. Milanes [13,37], M.-N. Minard [4], J. Molina Rodriguez [54], S. Monteil [5], D. Moran [12], P. Morawski [25], R. Mountain [52], I. Mous [23], F. Muheim [46], K. Müller [39], R. Muresan [28,38], B. Muryn [26], B. Muster [38], M. Musy [35], J. Mylroie-Smith [48], P. Naik [42], T. Nakada [38], R. Nandakumar [45], I. Nasteva [1], M. Nedos [9], M. Needham [46], N. Neufeld [37], C. Nguyen-Mau [38,o], M. Nicol [7], V. Niess [5], N. Nikitin [31], A. Nomerotski [51], A. Novoselov [34], A. Oblakowska-Mucha [26], V. Obraztsov [34], S. Oggero [23], S. Ogilvy [47], O. Okhrimenko [41], R. Oldeman [15,d], M. Orlandea [28], J.M. Otalora Goicochea [2], P. Owen [49], K. Pal [52], J. Palacios [39], A. Palano [13,b], M. Palutan [18], J. Panman [37], A. Papanestis [45], M. Pappagallo [47], C. Parkes [47,37], C.J. Parkinson [49], G. Passaleva [17], G.D. Patel [48], M. Patel [49], S.K. Paterson [49], G.N. Patrick [45], C. Patrignani [19,i], C. Pavel-Nicorescu [28], A. Pazos Alvarez [36], A. Pellegrino [23], G. Penso [22,l], M. Pepe Altarelli [37], S. Perazzini [14,c], D.L. Perego [20,j], E. Perez Trigo [36], A. Pérez-Calero Yzquierdo [35], P. Perret [5], M. Perrin-Terrin [6], G. Pessina [20], A. Petrella [16,37], A. Petrolini [19,i], A. Phan [52], E. Picatoste Olloqui [35], B. Pie Valls [35], B. Pietrzyk [4], T. Pilař [44], D. Pinci [22], R. Plackett [47], S. Playfer [46], M. Plo Casasus [36], G. Polok [25], A. Poluektov [44,33], E. Polycarpo [2], D. Popov [10], B. Popovici [28], C. Potterat [35], A. Powell [51], T. du Pree [23], J. Prisciandaro [38], V. Pugatch [41], A. Puig Navarro [35], W. Qian [52], J.H. Rademacker [42], B. Rakotomiaramanana [38], M.S. Rangel [2], I. Raniuk [40], G. Raven [24], S. Redford [51], M.M. Reid [44], A.C. dos Reis [1], S. Ricciardi [45], K. Rinnert [48], D.A. Roa Romero [5], P. Robbe [7], E. Rodrigues [47,50], F. Rodrigues [2], P. Rodriguez Perez [36], G.J. Rogers [43], S. Roiser [37], V. Romanovsky [34], M. Rosello [35,n], J. Rouvinet [38], T. Ruf [37], H. Ruiz [35], G. Sabatino [21,k], J.J. Saborido Silva [36], N. Sagidova [29], P. Sail [47], B. Saitta [15,d], C. Salzmann [39], M. Sannino [19,i], R. Santacesaria [22], C. Santamarina Rios [36], R. Santinelli [37], E. Santovetti [21,k], M. Sapunov [6], A. Sarti [18,l], C. Satriano [22,m], A. Satta [21], M. Savrie [16,e], D. Savrina [30], P. Schaack [49], M. Schiller [24], S. Schleich [9], M. Schlupp [9], M. Schmelling [10], B. Schmidt [37], O. Schneider [38], A. Schopper [37], M.-H. Schune [7], R. Schwemmer [37], B. Sciascia [18], A. Sciubba [18,l], M. Seco [36], A. Semennikov [30], K. Senderowska [26], I. Sepp [49], N. Serra [39], J. Serrano [6], P. Seyfert [11], B. Shao [3], M. Shapkin [34], I. Shapoval [40,37], P. Shatalov [30], Y. Shcheglov [29], T. Shears [48], L. Shekhtman [33], O. Shevchenko [40], V. Shevchenko [30], A. Shires [49], R. Silva Coutinho [44], T. Skwarnicki [52], A.C. Smith [37], N.A. Smith [48], E. Smith [51,45], K. Sobczak [5], F.J.P. Soler [47], A. Solomin [42], F. Soomro [18], B. Souza De Paula [2], B. Spaan [9], A. Sparkes [46], P. Spradlin [47], F. Stagni [37], S. Stahl [11], O. Steinkamp [39], S. Stoica [28], S. Stone [52,37], B. Storaci [23], M. Straticiuc [28], U. Straumann [39], V.K. Subbiah [37], S. Swientek [9], M. Szczekowski [27], P. Szczypka [38], T. Szumlak [26], S. T'Jampens [4], E. Teodorescu [28], F. Teubert [37], C. Thomas [51], E. Thomas [37], J. van Tilburg [11], V. Tisserand [4], M. Tobin [39], S. Topp-Joergensen [51], N. Torr [51], E. Tournefier [4,49], M.T. Tran [38], A. Tsaregorodtsev [6], N. Tuning [23], M. Ubeda Garcia [37], A. Ukleja [27], P. Urquijo [52], U. Uwer [11], V. Vagnoni [14], G. Valenti [14], R. Vazquez Gomez [35], P. Vazquez Regueiro [36], S. Vecchi [16], J.J. Velthuis [42], M. Veltri [17,g], B. Viaud [7], I. Videau [7], X. Vilasis-Cardona [35,n], J. Visniakov [36], A. Vollhardt [39], D. Volyanskyy [10], D. Voong [42], A. Vorobyev [29], H. Voss [10], S. Wandernoth [11], J. Wang [52], D.R. Ward [43], N.K. Watson [55], A.D. Webber [50], D. Websdale [49], M. Whitehead [44], D. Wiedner [11], L. Wiggers [23], G. Wilkinson [51], M.P. Williams [44,45],

M. Williams [49], F.F. Wilson [45], J. Wishahi [9], M. Witek [25], W. Witzeling [37], S.A. Wotton [43], K. Wyllie [37], Y. Xie [46], F. Xing [51], Z. Xing [52], Z. Yang [3], R. Young [46], O. Yushchenko [34], M. Zavertyaev [10,a], F. Zhang [3], L. Zhang [52], W.C. Zhang [12], Y. Zhang [3], A. Zhelezov [11], L. Zhong [3], E. Zverev [31], A. Zvyagin [37]

[1] *Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil*
[2] *Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil*
[3] *Center for High Energy Physics, Tsinghua University, Beijing, China*
[4] *LAPP, Université de Savoie, CNRS/IN2P3, Annecy-Le-Vieux, France*
[5] *Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, Clermont-Ferrand, France*
[6] *CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*
[7] *LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France*
[8] *LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris, France*
[9] *Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany*
[10] *Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany*
[11] *Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany*
[12] *School of Physics, University College Dublin, Dublin, Ireland*
[13] *Sezione INFN di Bari, Bari, Italy*
[14] *Sezione INFN di Bologna, Bologna, Italy*
[15] *Sezione INFN di Cagliari, Cagliari, Italy*
[16] *Sezione INFN di Ferrara, Ferrara, Italy*
[17] *Sezione INFN di Firenze, Firenze, Italy*
[18] *Laboratori Nazionali dell'INFN di Frascati, Frascati, Italy*
[19] *Sezione INFN di Genova, Genova, Italy*
[20] *Sezione INFN di Milano Bicocca, Milano, Italy*
[21] *Sezione INFN di Roma Tor Vergata, Roma, Italy*
[22] *Sezione INFN di Roma La Sapienza, Roma, Italy*
[23] *Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands*
[24] *Nikhef National Institute for Subatomic Physics and Vrije Universiteit, Amsterdam, The Netherlands*
[25] *Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraców, Poland*
[26] *AGH University of Science and Technology, Kraców, Poland*
[27] *Soltan Institute for Nuclear Studies, Warsaw, Poland*
[28] *Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania*
[29] *Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia*
[30] *Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia*
[31] *Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia*
[32] *Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia*
[33] *Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia*
[34] *Institute for High Energy Physics (IHEP), Protvino, Russia*
[35] *Universitat de Barcelona, Barcelona, Spain*
[36] *Universidad de Santiago de Compostela, Santiago de Compostela, Spain*
[37] *European Organization for Nuclear Research (CERN), Geneva, Switzerland*
[38] *Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[39] *Physik-Institut, Universität Zürich, Zürich, Switzerland*
[40] *NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine*
[41] *Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine*
[42] *H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom*
[43] *Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom*
[44] *Department of Physics, University of Warwick, Coventry, United Kingdom*
[45] *STFC Rutherford Appleton Laboratory, Didcot, United Kingdom*
[46] *School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom*
[47] *School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*
[48] *Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom*
[49] *Imperial College London, London, United Kingdom*
[50] *School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*
[51] *Department of Physics, University of Oxford, Oxford, United Kingdom*
[52] *Syracuse University, Syracuse, NY, United States*
[53] *CC-IN2P3, CNRS/IN2P3, Lyon-Villeurbanne, France* [p]
[54] *Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil* [q]
[55] *University of Birmingham, Birmingham, United Kingdom*

* Corresponding authors.
E-mail addresses: Marc-Olivier.Bettler@cern.ch (M.-O. Bettler), Gaia.Lanfranchi@lnf.infn.it (G. Lanfranchi).
[a] P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
[b] Università di Bari, Bari, Italy.
[c] Università di Bologna, Bologna, Italy.
[d] Università di Cagliari, Cagliari, Italy.
[e] Università di Ferrara, Ferrara, Italy.
[f] Università di Firenze, Firenze, Italy.
[g] Università di Urbino, Urbino, Italy.
[h] Università di Modena e Reggio Emilia, Modena, Italy.
[i] Università di Genova, Genova, Italy.
[j] Università di Milano Bicocca, Milano, Italy.
[k] Università di Roma Tor Vergata, Roma, Italy.
[l] Università di Roma La Sapienza, Roma, Italy.
[m] Università della Basilicata, Potenza, Italy.
[n] LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.

[o] Hanoi University of Science, Hanoi, Viet Nam.
[p] Associated member.
[q] Associated to Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil.