First Evidence for the Decay $B_s^0 \to \mu^+\mu^-$

R. Aaij *et al.**
(LHCb Collaboration)

(Received 12 November 2012; published 7 January 2013)

A search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ is performed with data collected in 2011 and 2012 with the LHCb experiment at the Large Hadron Collider. The data samples comprise 1.1 fb^{-1} of proton-proton collisions at $\sqrt{s} = 8$ TeV and 1.0 fb^{-1} at $\sqrt{s} = 7$ TeV. We observe an excess of $B_s^0 \to \mu^+\mu^-$ candidates with respect to the background expectation. The probability that the background could produce such an excess or larger is 5.3×10^{-4} corresponding to a signal significance of 3.5 standard deviations. A maximum-likelihood fit gives a branching fraction of $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2^{+1.5}_{-1.2}) \times 10^{-9}$, where the statistical uncertainty is 95% of the total uncertainty. This result is in agreement with the standard model expectation. The observed number of $B^0 \to \mu^+\mu^-$ candidates is consistent with the background expectation, giving an upper limit of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 9.4 \times 10^{-10}$ at 95% confidence level.

DOI: 10.1103/PhysRevLett.110.021801 PACS numbers: 13.20.He, 12.15.Mm, 12.60.Jv

The rare decays, $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$, are highly suppressed in the standard model (SM). Precise predictions of their branching fractions, $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.23 \pm 0.27) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.07 \pm 0.10) \times 10^{-10}$ [1], make these modes powerful probes in the search for deviations from the SM, especially in models with a nonstandard Higgs sector. Taking the measured finite width difference of the B_s^0 system [2] into account [3], the time integrated branching fraction of $B_s^0 \to \mu^+\mu^-$ that should be compared to the experimental value is $(3.54 \pm 0.30) \times 10^{-9}$.

Previous searches [4–8] already constrain possible deviations from the SM predictions. The lowest published limits are $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.5 \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.0 \times 10^{-9}$ at 95% confidence level (C.L.) from the LHCb Collaboration using 1.0 fb^{-1} of data collected in pp collisions in 2011 at $\sqrt{s} = 7$ TeV [8]. This Letter reports an update of this search with 1.1 fb^{-1} of data recorded in 2012 at $\sqrt{s} = 8$ TeV.

The analysis of 2012 data is similar to that described in Ref. [8] with two main improvements: the use of particle identification to select $B_{(s)}^0 \to h^+h'^-$ [with $h^{(\prime)} = K, \pi$] decays used to calibrate the geometrical and kinematic variables, and a refined estimate of the exclusive backgrounds. To avoid potential bias, the events in the signal region were not examined until all the analysis choices were finalized. The updated estimate of the exclusive backgrounds is also applied to the 2011 data [8] and the results reevaluated. The results obtained with the combined 2011 and 2012 data sets supersede those of Ref. [8].

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, and is described in detail in Ref. [9]. The simulated events used in this analysis are produced using the software described in Refs. [10–16].

Candidate $B_{(s)}^0 \to \mu^+\mu^-$ events are required to be selected by a hardware and a subsequent software trigger [17]. The candidates are predominantly selected by single and dimuon trigger and, to a smaller extent, by a generic b-hadron trigger. Candidate events in the $B^+ \to J/\psi K^+$ control channel, with $J/\psi \to \mu^+\mu^-$ (inclusion of charged conjugated processes is implied throughout this Letter), are selected in a very similar way, the only difference being a different dimuon mass requirement in the final software trigger. The $B_{(s)}^0 \to h^+h'^-$ decays are predominantly selected by a hardware trigger based on the calorimeter transverse energy and subsequently by a generic b-hadron software trigger.

The $B_{(s)}^0 \to \mu^+\mu^-$ candidates are selected by requiring two high quality muon candidates [18] displaced with respect to any pp interaction vertex [primary vertex (PV)], and forming a secondary vertex with a χ^2 per degree of freedom smaller than 9 and separated from the PV in the downstream direction by a flight distance significance greater than 15. Only candidates with an impact parameter χ^2, IPχ^2 (defined as the difference between the χ^2 of the PV formed with and without the considered tracks) less than 25 are considered. When more than one PV is reconstructed, that giving the smallest IPχ^2 for the B candidate is chosen. Tracks from selected candidates are required to have transverse momentum p_T satisfying $0.25 < p_T < 40$ GeV/c and $p < 500$ GeV/c. Only B candidates with decay times smaller than $9\tau(B_s^0)$ [19] and with invariant mass in the range [4900, 6000] MeV/c^2 are kept.

*Full author list given at the end of the article.

Dimuon candidates from elastic diphoton production are heavily suppressed by requiring $p_T(B) > 0.5\ \mathrm{GeV}/c$. The surviving background comprises mainly random combinations of muons from semileptonic decays of two different b hadrons ($b\bar{b} \to \mu^+\mu^- X$, where X is any other set of particles).

Two channels, $B^+ \to J/\psi K^+$ and $B^0 \to K^+\pi^-$, serve as normalization modes. The first mode has trigger and muon identification efficiencies similar to those of the signal, but a different number of tracks in the final state. The second mode has a similar topology, but is triggered differently. The selection of these channels is as close as possible to that of the signal to reduce the impact of potential systematic uncertainties.

The $B^0 \to K^+\pi^-$ selection is the same as for the $B^0_{(s)} \to \mu^+\mu^-$ signal except for muon identification. The two tracks are nevertheless required to be within the muon detector acceptance.

The $J/\psi \to \mu^+\mu^-$ decay in the $B^+ \to J/\psi K^+$ normalization channel is also selected similarly to the $B^0_{(s)} \to \mu^+\mu^-$ signals, except for the requirements on the IPχ^2 and mass. Kaon candidates are required to have IP$\chi^2 > 25$.

A two-stage multivariate selection, based on boosted decision trees (BDT) [20,21], is applied to the $B^0_{(s)} \to \mu^+\mu^-$ candidates. A cut on the first multivariate discriminant, unchanged from Ref. [8], removes 80% of the background while retaining 92% of the signal. The efficiencies of this cut for the signal and the normalization samples are equal within 0.2% as determined from simulation.

The output of the second multivariate discriminant, called BDT, and the dimuon invariant mass are used to classify the selected candidates. The nine variables entering the BDT are the B candidate IP, the minimum IPχ^2 of the two muons with respect to any PV, the sum of the degrees of isolation of the muons (the number of good two-track vertices a muon can make with other tracks in the event), the B candidate decay time, p_T, and isolation [22], the distance of closest approach between the two muons, the minimum p_T of the muons, and the cosine of the angle between the muon momentum in the dimuon rest frame and the vector perpendicular to both the B candidate momentum and the beam axis.

The BDT discriminant is trained with simulated samples consisting of $B^0_s \to \mu^+\mu^-$ for the signal and $b\bar{b} \to \mu^+\mu^- X$ for the background. The BDT response is defined such that it is approximately uniformly distributed between 0 and 1 for signal events and peaks at 0 for the background. The BDT response is independent of the invariant mass for the signal inside the search window. The probability for a $B^0_{(s)} \to \mu^+\mu^-$ event to have a given BDT value is obtained from data using $B^0 \to K^+\pi^-$, $\pi^+\pi^-$ and $B^0_s \to \pi^+K^-$, K^+K^- exclusive decays selected as the signal events and triggered independently of the tracks from $B^0_{(s)}$ candidates.

The invariant mass line shape of the signal events is described by a Crystal Ball function [23]. The peak values for the B^0_s and B^0 mesons, $m_{B^0_s}$ and m_{B^0}, are obtained from the $B^0_s \to K^+K^-$ and $B^0 \to K^+\pi^-$, $B^0 \to \pi^+\pi^-$ samples. The resolutions are determined by combining the results obtained with a power-law interpolation between the measured resolutions of charmonium and bottomonium resonances decaying into two muons with those obtained with a fit of the mass distributions of $B^0 \to K^+\pi^-$, $B^0 \to \pi^+\pi^-$, and $B^0_s \to K^+K^-$ samples. The results are $\sigma_{B^0_s} = 25.0 \pm 0.4\ \mathrm{MeV}/c^2$ and $\sigma_{B^0} = 24.6 \pm 0.4\ \mathrm{MeV}/c^2$, respectively. The transition point of the radiative tail is obtained from simulated $B^0_s \to \mu^+\mu^-$ events smeared to reproduce the mass resolution measured in the data.

The $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ yields are translated into branching fractions with

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = \frac{\mathcal{B}_{\mathrm{norm}}\epsilon_{\mathrm{norm}}f_{\mathrm{norm}}}{N_{\mathrm{norm}}\epsilon_{\mathrm{sig}}f_{d(s)}} \times N_{B^0_{(s)}\to\mu^+\mu^-}, \\ = \alpha^{\mathrm{norm}}_{B^0_{(s)}\to\mu^+\mu^-} \times N_{B^0_{(s)}\to\mu^+\mu^-}, \quad (1)$$

where $\mathcal{B}_{\mathrm{norm}}$ represents the branching fraction, N_{norm} the number of signal events in the normalization channel obtained from a fit to the invariant mass distribution, and $N_{B^0_{(s)}\to\mu^+\mu^-}$ is the number of observed signal events.

The factors $f_{d(s)}$ and f_{norm} indicate the probabilities that a b quark fragments into a $B^0_{(s)}$ meson and into the hadron involved in the given normalization mode, respectively. We assume $f_d = f_u$ and use $f_s/f_d = 0.256 \pm 0.020$ measured in pp collision data at $\sqrt{s} = 7$ TeV [24]. This value is in agreement within 1.5σ with that found at $\sqrt{s} = 8$ TeV by comparing the ratios of the yields of $B^0_s \to J/\psi\phi$ and $B^+ \to J/\psi K^+$ decays. The measured dependence of f_s/f_d on $p_T(B)$ [24] is found to be negligible for this analysis.

The efficiency $\epsilon_{\mathrm{sig(norm)}}$ for the signal (normalization channel) is the product of the reconstruction efficiency of the final state particles including the geometric detector acceptance, the selection efficiency, and the trigger efficiency. The ratio of acceptance, reconstruction, and selection efficiencies is computed with the use of simulation. Potential differences between data and simulation are accounted for as systematic uncertainties. Reweighting techniques are used for all the distributions in the simulation that do not match those from data. The trigger efficiency is evaluated with data-driven techniques [25]. The observed numbers of $B^+ \to J/\psi K^+$ and $B^0 \to K^+\pi^-$ candidates in the 2012 data set are $424\,200 \pm 1500$ and $14\,600 \pm 1100$, respectively. The two normalization factors $\alpha^{\mathrm{norm}}_{B^0_{(s)}\to\mu^+\mu^-}$ are in agreement within the uncertainties, and their weighted average, taking correlations into account, gives $\alpha_{B^0_s\to\mu^+\mu^-} = (2.52 \pm 0.23) \times 10^{-10}$ and $\alpha_{B^0\to\mu^+\mu^-} = (6.45 \pm 0.30) \times 10^{-11}$.

In total, 24 044 muon pairs with an invariant mass between 4900 and 6000 MeV/c^2 pass the trigger and selection requirements. Given the measured normalization factors and assuming the SM branching fractions, the data sample is expected to contain about 14.1 $B_s^0 \to \mu^+\mu^-$ and 1.7 $B^0 \to \mu^+\mu^-$ decays.

The BDT range is divided into eight bins with boundaries [0.0, 0.25, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]. For the 2012 data set, only one bin is considered in the BDT range 0.8–1.0 due to the lack of events in the mass sidebands for BDT > 0.9. The signal regions are defined by $m_{B^0_{(s)}} \pm 60$ MeV$/c^2$.

The expected number of combinatorial background events is determined by interpolating from the invariant mass sideband regions defined as [4900 MeV$/c^2$, $m_{B^0} -$ 60 MeV$/c^2$] and [$m_{B_s^0} + 60$ MeV$/c^2$, 6000 MeV$/c^2$]. The low-mass sideband and the B^0 and B_s^0 signal regions are potentially polluted by exclusive backgrounds with or without the misidentification of the muon candidates.

The first category includes $B^0 \to \pi^-\mu^+\nu_\mu$, $B^0_{(s)} \to h^+h'^-$, $B_s^0 \to K^-\mu^+\nu_\mu$, and $\Lambda_b^0 \to p\mu^-\bar{\nu}_\mu$ decays. The $B^0 \to \pi^-\mu^+\nu_\mu$ and $B^0_{(s)} \to h^+h'^-$ branching fractions are taken from Ref. [19]. The theoretical estimates of the $\Lambda_b^0 \to p\mu^-\bar{\nu}_\mu$ and $B_s^0 \to K^-\mu^+\nu_\mu$ branching fractions are taken from Refs. [26,27], respectively. The mass and BDT distributions of these modes are evaluated from simulated samples where the $K \to \mu$, $\pi \to \mu$ and $p \to \mu$ misidentification probabilities as a function of momentum and transverse momentum are those determined from $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$, and $\Lambda \to p\pi^-$ data samples. We use the Λ_b^0 fragmentation fraction $f_{\Lambda_b^0}$ measured by LHCb [28] and account for its p_T dependence.

The second category includes $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$, $B_s^0 \to \mu^+\mu^-\gamma$, and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ decays, evaluated assuming branching fraction values from Refs. [29–31], respectively. Apart from $B^0_{(s)} \to h^+h'^-$, all background modes are normalized relative to the $B^+ \to J/\psi K^+$ decay. The $B^0 \to \pi^-\mu^+\nu_\mu$, $B^0_{(s)} \to h^+h'^-$, and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ decays are the dominant exclusive modes in the range BDT > 0.8, which accounts for 70% of the sensitivity.

In the full BDT range, 8.6 ± 0.7 doubly misidentified $B^0_{(s)} \to h^+h'^-$ decays are expected in the full mass interval, $4.1^{+1.7}_{-0.8}$ in the B^0, and $0.76^{+0.26}_{-0.18}$ in the B_s^0signal region. The expected yields for $B^0 \to \pi^-\mu^+\nu_\mu$ and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ are 41.1 ± 0.4 and 11.9 ± 3.5, respectively, in the full mass and BDT ranges. The contributions of these two backgrounds above $m_{B^0} - 60$ MeV$/c^2$ are negligible. The fractions of these backgrounds with BDT > 0.8, in the full mass range, are $(19.0 \pm 1.4)\%$, $(11.1 \pm 0.5)\%$, and $(12.2 \pm 0.3)\%$ for $B^0_{(s)} \to h^+h'^-$, $B^0 \to \pi^-\mu^+\nu_\mu$, and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ decays, respectively.

A simultaneous unbinned maximum-likelihood fit to the mass projections in the BDT bins is performed on the mass sidebands to determine the number of expected combinatorial background events in the B^0 and B_s^0 signal regions used in the derivation of the branching fraction limit. In this fit, the parameters that describe the mass distributions of the exclusive backgrounds, their fractional yields in each BDT bin, and their overall yields are limited by Gaussian constraints according to their expected values and uncertainties. The combinatorial background is parametrized with an exponential function with slope and normalization allowed to vary. The systematic uncertainty on the estimated number of combinatorial background events in the signal regions is determined by fluctuating the number of events observed in the sidebands according to a Poisson distribution, and by varying the exponential slope according to its uncertainty. The same fit is then performed on the full mass range to determine the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions, which are free parameters of the fit. The $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ fractional yields in BDT bins are constrained to the BDT fractions calibrated with the $B^0_{(s)} \to h^+h'^-$ sample. The parameters of the Crystal Ball functions that describe the mass line shapes and the normalization factors are restricted by Gaussian constraints according to their expected values and uncertainties. The parameters of the Crystal Ball functions, the normalization factors, the parameters that describe the mass distributions of the exclusive backgrounds, the overall yields of the exclusive backgrounds, and the fractional yields in each BDT bin of the exclusive backgrounds and the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decay modes are considered as nuisance parameters in the maximum-likelihood fit.

The compatibility of the observed distribution of events with that expected for a given branching fraction hypothesis is computed with the CL_s method [32]. The method provides CL_{s+b} a measure of the compatibility of the observed distribution with the signal plus background hypothesis CL_b, a measure of the compatibility with the background-only hypothesis, and $\mathrm{CL}_s = \mathrm{CL}_{s+b}/\mathrm{CL}_b$.

The invariant mass signal regions are divided into nine bins with boundaries $m_{B^0_{(s)}} \pm 18, 30, 36, 48, 60$ MeV$/c^2$. In each bin of the two-dimensional space formed by the dimuon mass and the BDT output, we count the number of observed candidates, and compute the expected number of signal and background events.

The comparison of the distributions of observed events and expected background events in the 2012 data set results in p values $(1 - \mathrm{CL}_b)$ of 9×10^{-4} for the $B_s^0 \to \mu^+\mu^-$ and 0.16 for the $B^0 \to \mu^+\mu^-$ decay, computed at the branching fraction values corresponding to $\mathrm{CL}_{s+b} = 0.5$. We observe an excess of $B_s^0 \to \mu^+\mu^-$ candidates with respect to background expectation with a significance of 3.3 standard deviations. The simultaneous unbinned maximum-likelihood fit gives $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = [5.1^{+2.3}_{-1.9}(\mathrm{stat})^{+0.7}_{-0.4}(\mathrm{syst})] \times 10^{-9}$. The statistical uncertainty reflects the interval corresponding to a change of 0.5 with respect to the minimum of the log-likelihood after fixing all the fit parameters to their expected values except the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions and the slope and normalization of the combinatorial background. The systematic uncertainty is obtained by subtracting in quadrature the statistical uncertainty from the total uncertainty obtained from the likelihood with all nuisance parameters left to vary according to their uncertainties. An additional systematic uncertainty of 0.16×10^{-9} reflects the impact on the result of the change in the parametrization of the combinatorial background from a single to a double exponential, and is added in quadrature.

The expected and measured limits on the $B^0 \to \mu^+\mu^-$ branching fraction at 90% and 95% C.L. are shown in Table I. The expected limits are computed allowing for the presence of $B^0_{(s)} \to \mu^+\mu^-$ events according to the SM branching fractions, including cross feed between the two modes.

The contribution of the exclusive background components is also evaluated for the 2011 data set, modifying the number of expected combinatorial background in the signal regions; a fraction of events populating the low-mass sideband [4.9–5.0] GeV/c^2 used in Ref. [8] to interpolate the combinatorial background in the signal regions, is now assigned to exclusive background components and, hence, not considered in the interpolation procedure. The results for the $B^0_{(s)} \to \mu^+\mu^-$ branching fractions have been updated accordingly. We obtain $\mathcal{B}(B^0_s \to \mu^+\mu^-) < 5.1 \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 13 \times 10^{-10}$ at 95% C.L. to be compared to the published limits $\mathcal{B}(B^0_s \to \mu^+\mu^-) < 4.5 \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 10.3 \times 10^{-10}$ at 95% C.L. [8], respectively. The $(1 - \mathrm{CL}_b)$ p value for $B^0_s \to \mu^+\mu^-$ changes from 18% to 11% and the $B^0_s \to \mu^+\mu^-$ branching fraction increases by $\sim 0.3\sigma$ from $(0.8^{+1.8}_{-1.3}) \times 10^{-9}$ to $(1.4^{+1.7}_{-1.3}) \times 10^{-9}$. This shift is compatible with the systematic uncertainty previously assigned to the background shape [8]. The values of the $B^0_s \to \mu^+\mu^-$ branching fraction obtained with the 2011 and 2012 data sets are compatible within 1.5σ.

The 2011 and 2012 results are combined by computing the CL_s and performing the maximum-likelihood fit simultaneously to the eight and seven BDT bins of the 2011 and 2012 data sets, respectively. The parameters that are considered 100% correlated between the two data sets are f_s/f_d, $\mathcal{B}(B^+ \to J/\psi K^+)$ and $\mathcal{B}(B^0 \to K^+\pi^-)$, the transition point of the Crystal Ball function describing the signal mass line shape, the mass distribution of the $B^0_{(s)} \to h^+h'^-$ background, the BDT and mass distributions of the $B^0 \to \pi^-\mu^+\nu_\mu$ and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ backgrounds, and the SM predictions of the $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions. The distribution of the expected and observed events in bins of BDT in the signal regions obtained from the simultaneous analysis of the 2011 and 2012 data sets, are available as supplemental material [33].

The expected and observed upper limits for the $B^0 \to \mu^+\mu^-$ channel obtained from the combined 2011 and 2012 data sets are summarized in Table I and the expected and observed CL_s values as a function of the branching fraction are shown in Fig. 1. The observed CL_b value at $\mathrm{CL}_{s+b} = 0.5$ is 89%. The probability that background processes can produce the observed number of $B^0_s \to \mu^+\mu^-$ candidates or more is 5×10^{-4} and corresponds to a statistical significance of 3.5σ. The value of the $B^0_s \to \mu^+\mu^-$ branching fraction obtained from the fit is

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = [3.2^{+1.4}_{-1.2}(\mathrm{stat})^{+0.5}_{-0.3}(\mathrm{syst})] \times 10^{-9}$$

and is in agreement with the SM expectation. The invariant mass distribution of the $B^0_{(s)} \to \mu^+\mu^-$ candidates with BDT > 0.7 is shown in Fig. 2.

The true value of the $B^0_s \to \mu^+\mu^-$ branching fraction is contained in the interval $[1.3, 5.8] \times 10^{-9}$ ($[1.1, 6.4] \times 10^{-9}$) at 90% C.L. (95% C.L.), where the lower and upper limit are the branching fractions evaluated at $\mathrm{CL}_{s+b} = 0.95$ ($\mathrm{CL}_{s+b} = 0.975$) and $\mathrm{CL}_{s+b} = 0.05$ ($\mathrm{CL}_{s+b} = 0.025$), respectively. These results are in good agreement with the lower and upper limits derived from integrating the profile likelihood obtained from the unbinned fit.

TABLE I. Expected and observed limits on the $B^0 \to \mu^+\mu^-$ branching fractions for the 2012 and for the combined 2011 and 2012 data sets.

Data set	Limit at	90% C.L.	95% C.L.
2012	Exp.bkg + SM	8.5×10^{-10}	10.5×10^{-10}
	Exp. bkg	7.6×10^{-10}	9.6×10^{-10}
	Observed	10.5×10^{-10}	12.5×10^{-10}
2011 and 2012	Exp.bkg + SM	5.8×10^{-10}	7.1×10^{-10}
	Exp. bkg	5.0×10^{-10}	6.0×10^{-10}
	Observed	8.0×10^{-10}	9.4×10^{-10}

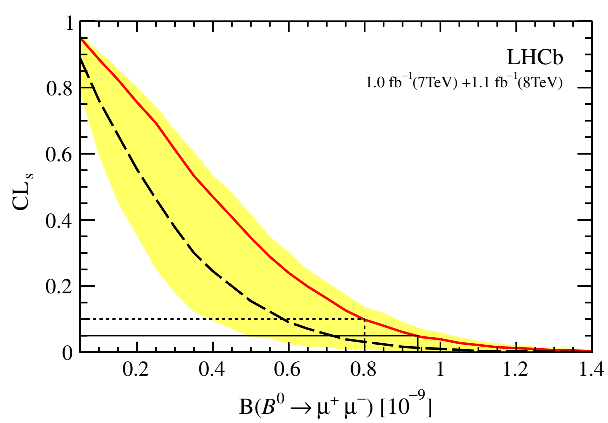

FIG. 1 (color online). CL_s as a function of the assumed $B^0 \to \mu^+\mu^-$ branching fraction for the combined 2011 and 2012 data sets. The dashed curve is the median of the expected CL_s distribution if the background and SM signal were observed. The shaded yellow area covers, for each branching fraction value, 34% of the expected CL_s distribution on each side of its median. The solid red curve is the observed CL_s.

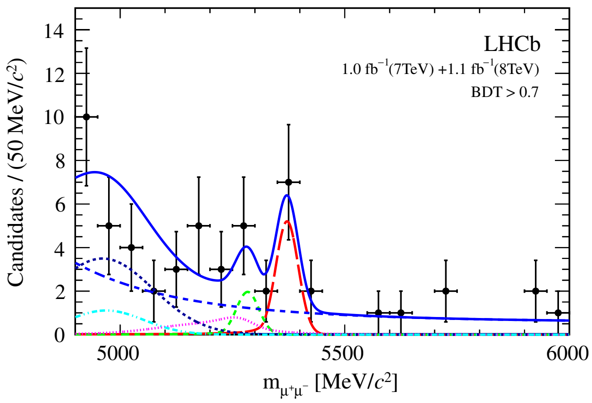

FIG. 2 (color online). Invariant mass distribution of the selected $B_s^0 \to \mu^+\mu^-$ candidates (black dots) with BDT > 0.7 in the combined 2011 and 2012 data sets. The result of the fit is overlaid (blue solid line) and the different components detailed: $B_s^0 \to \mu^+\mu^-$ (red long dashed curve), $B^0 \to \mu^+\mu^-$ (green medium dashed curve), $B^0_{(s)} \to h^+h'^-$ (pink dotted curve), $B^0 \to \pi^-\mu^+\nu_\mu$ (black short dashed curve), and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ (light blue dash-dotted curve), and the combinatorial background (blue medium dashed).

In summary, a search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ is performed with 1.0 fb^{-1} and 1.1 fb^{-1} of pp collision data collected at $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV, respectively. The data in the B^0 search window are consistent with the background expectation and an improved upper limit of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 9.4 \times 10^{-10}$ at 95% C.L. is obtained. The data in the B_s^0 search window show an excess of events with respect to the background-only prediction with a statistical significance of 3.5σ. A fit to the data leads to $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2^{+1.5}_{-1.2}) \times 10^{-9}$ which is in agreement with the SM prediction. This is the first evidence for the decay $B_s^0 \to \mu^+\mu^-$.

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ, and FINEP (Brazil); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF, and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); SCSR (Poland); ANCS/IFA (Romania); MinES, Rosatom, RFBR, and NRC "Kurchatov Institute" (Russia); MinECo, XuntaGal, and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); NSF (U.S.). We also acknowledge the support received from the ERC under FP7. The Tier1 computing centers are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), and GridPP (U.K.). We are thankful for the computing resources put at our disposal by Yandex LLC (Russia), as well as to the communities behind the multiple open source software packages that we depend on.

[1] A. J. Buras, J. Girrbach, D. Guadagnoli, and G. Isidori, Eur. Phys. J. C **72**, 2172 (2012).
[2] R. Aaij *et al.* (LHCb Collaboration), Report No. LHCb-CONF-2012-002.
[3] K. de Bruyn, R. Fleischer, R. Knegjens, P. Koppenburg, M. Merk, A. Pellegrino, and N. Tuning, Phys. Rev. Lett. **109**, 041801 (2012).
[4] V. M. Abazov *et al.* (DO Collaboration), Phys. Lett. B **693**, 539 (2010).
[5] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **107**, 191801 (2011).
[6] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 04 (2012) 033.
[7] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **713**, 387 (2012).
[8] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **108**, 231801 (2012).
[9] A. A. Alves, Jr. *et al.* (LHCb Collaboration), JINST **3**, S08005 (2008).
[10] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.
[11] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[12] J. Allison *et al.* (GEANT4 Collaboration) IEEE Trans. Nucl. Sci. **53**, 270 (2006).
[13] S. Agostinelli *et al.* (GEANT4 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[14] P. Golonka and Z. Was, Eur. Phys. J. C **45**, 97 (2006).
[15] I. Belyaev *et al.*, in *Nuclear Science Symposium Conference Record (NSS/MIC)* (IEEE, New York, 2010), p. 1155.
[16] M. Clemencic, G. Corti, S. Easo, C. R. Jones, S. Miglioranzi, M. Pappagallo, and P. Robbe, J. Phys. Conf. Ser. **331**, 032023 (2011).
[17] R. Aaij *et al.*, arXiv:1211.3055.
[18] G. Lanfranchi *et al.*, Report No. LHCb-2009-013.
[19] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).
[20] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees* (Wadsworth International Group, Belmont, CA, 1984).
[21] R. E. Schapire and Y. Freund, J. Comput. Syst. Sci. **55**, 119 (1997).
[22] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **95**, 221805 (2005).
[23] T. Skwarnicki, Ph.D. thesis, Institute of Nuclear Physics, Krakow, 1986; Report No. DESY-F31-86-02.
[24] R. Aaij *et al.* (LHCb Collaboration), Report No. LHCb-PAPER-2012-037 (unpublished).
[25] J. H. Morata *et al.*, Report No. LHCb-2008-073.
[26] A. Datta, arXiv:hep-ph/9504429.
[27] W. Wang and Z. Xiao, arXiv:1207.0265.
[28] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **85**, 032008 (2012).
[29] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **81**, 2432 (1998).
[30] D. Melikhov and N. Nikitin, Phys. Rev. D **70**, 114028 (2004).
[31] R. Aaij *et al.* (LHCb Collaboration), arXiv:1210.2645 [J. High Energy Phys. (to be published)].

[32] A. L. Read, J. Phys. G 28, 2693 (2002).
[33] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.110.021801 for the expected combinatorial background, peaking background, cross-feed, and signal events assuming the SM prediction, together with the number of observed candidates in the signal regions, in bins of BDT for the 2011 and the 2012 data samples.

R. Aaij,[38] C. Abellan Beteta,[33,n] A. Adametz,[11] B. Adeva,[34] M. Adinolfi,[43] C. Adrover,[6] A. Affolder,[49] Z. Ajaltouni,[5] J. Albrecht,[35] F. Alessio,[35] M. Alexander,[48] S. Ali,[38] G. Alkhazov,[27] P. Alvarez Cartelle,[34] A. A. Alves, Jr.,[22] S. Amato,[2] Y. Amhis,[36] L. Anderlini,[17,f] J. Anderson,[37] R. Andreassen,[57] R. B. Appleby,[51] O. Aquines Gutierrez,[10] F. Archilli,[18,35] A. Artamonov,[32] M. Artuso,[53] E. Aslanides,[6] G. Auriemma,[22,m] S. Bachmann,[11] J. J. Back,[45] C. Baesso,[54] W. Baldini,[16] R. J. Barlow,[51] C. Barschel,[35] S. Barsuk,[7] W. Barter,[44] A. Bates,[48] Th. Bauer,[38] A. Bay,[36] J. Beddow,[48] I. Bediaga,[1] S. Belogurov,[28] K. Belous,[32] I. Belyaev,[28] E. Ben-Haim,[8] M. Benayoun,[8] G. Bencivenni,[18] S. Benson,[47] J. Benton,[43] A. Berezhnoy,[29] R. Bernet,[37] M.-O. Bettler,[44] M. van Beuzekom,[38] A. Bien,[11] S. Bifani,[12] T. Bird,[51] A. Bizzeti,[17,h] P. M. Bjørnstad,[51] T. Blake,[35] F. Blanc,[36] C. Blanks,[50] J. Blouw,[11] S. Blusk,[53] A. Bobrov,[31] V. Bocci,[22] A. Bondar,[31] N. Bondar,[27] W. Bonivento,[15] S. Borghi,[51,48] A. Borgia,[53] T. J. V. Bowcock,[49] E. Bowen,[37] C. Bozzi,[16] T. Brambach,[9] J. van den Brand,[39] J. Bressieux,[36] D. Brett,[51] M. Britsch,[10] T. Britton,[53] N. H. Brook,[43] H. Brown,[49] A. Büchler-Germann,[37] I. Burducea,[26] A. Bursche,[37] J. Buytaert,[35] S. Cadeddu,[15] O. Callot,[7] M. Calvi,[20,j] M. Calvo Gomez,[33,n] A. Camboni,[33] P. Campana,[18,35] A. Carbone,[14,c] G. Carboni,[21,k] R. Cardinale,[19,i] A. Cardini,[15] H. Carranza-Mejia,[47] L. Carson,[50] K. Carvalho Akiba,[2] G. Casse,[49] M. Cattaneo,[35] Ch. Cauet,[9] M. Charles,[52] Ph. Charpentier,[35] P. Chen,[3,36] N. Chiapolini,[37] M. Chrzaszcz,[23] K. Ciba,[35] X. Cid Vidal,[34] G. Ciezarek,[50] P. E. L. Clarke,[47] M. Clemencic,[35] H. V. Cliff,[44] J. Closier,[35] C. Coca,[26] V. Coco,[38] J. Cogan,[6] E. Cogneras,[5] P. Collins,[35] A. Comerma-Montells,[33] A. Contu,[15,52] A. Cook,[43] M. Coombes,[43] G. Corti,[35] B. Couturier,[35] G. A. Cowan,[36] D. Craik,[45] S. Cunliffe,[50] R. Currie,[47] C. D'Ambrosio,[35] P. David,[8] P. N. Y. David,[38] I. De Bonis,[4] K. De Bruyn,[38] S. De Capua,[51] M. De Cian,[37] J. M. De Miranda,[1] L. De Paula,[2] P. De Simone,[18] D. Decamp,[4] M. Deckenhoff,[9] H. Degaudenzi,[36,35] L. Del Buono,[8] C. Deplano,[15] D. Derkach,[14] O. Deschamps,[5] F. Dettori,[39] A. Di Canto,[11] J. Dickens,[44] H. Dijkstra,[35] P. Diniz Batista,[1] M. Dogaru,[26] F. Domingo Bonal,[33,n] S. Donleavy,[49] F. Dordei,[11] P. Dornan,[50] A. Dosil Suárez,[34] D. Dossett,[45] A. Dovbnya,[40] F. Dupertuis,[36] R. Dzhelyadin,[32] A. Dziurda,[23] A. Dzyuba,[27] S. Easo,[46,35] U. Egede,[50] V. Egorychev,[28] S. Eidelman,[31] D. van Eijk,[38] S. Eisenhardt,[47] R. Ekelhof,[9] L. Eklund,[48] I. El Rifai,[5] Ch. Elsasser,[37] D. Elsby,[42] A. Falabella,[14,e] C. Färber,[11] G. Fardell,[47] C. Farinelli,[38] S. Farry,[12] V. Fave,[36] V. Fernandez Albor,[34] F. Ferreira Rodrigues,[1] M. Ferro-Luzzi,[35] S. Filippov,[30] C. Fitzpatrick,[35] M. Fontana,[10] F. Fontanelli,[19,i] R. Forty,[35] O. Francisco,[2] M. Frank,[35] C. Frei,[35] M. Frosini,[17,f] S. Furcas,[20] A. Gallas Torreira,[34] D. Galli,[14,c] M. Gandelman,[2] P. Gandini,[52] Y. Gao,[3] J. Garofoli,[53] P. Garosi,[51] J. Garra Tico,[44] L. Garrido,[33] C. Gaspar,[35] R. Gauld,[52] E. Gersabeck,[11] M. Gersabeck,[51] T. Gershon,[45,35] Ph. Ghez,[4] V. Gibson,[44] V. V. Gligorov,[35] C. Göbel,[54] D. Golubkov,[28] A. Golutvin,[50,28,35] A. Gomes,[2] H. Gordon,[52] M. Grabalosa Gándara,[33] R. Graciani Diaz,[33] L. A. Granado Cardoso,[35] E. Graugés,[33] G. Graziani,[17] A. Grecu,[26] E. Greening,[52] S. Gregson,[44] O. Grünberg,[55] B. Gui,[53] E. Gushchin,[30] Yu. Guz,[32] T. Gys,[35] C. Hadjivasiliou,[53] G. Haefeli,[36] C. Haen,[35] S. C. Haines,[44] S. Hall,[50] T. Hampson,[43] S. Hansmann-Menzemer,[11] N. Harnew,[52] S. T. Harnew,[43] J. Harrison,[51] P. F. Harrison,[45] T. Hartmann,[55] J. He,[7] V. Heijne,[38] K. Hennessy,[49] P. Henrard,[5] J. A. Hernando Morata,[34] E. van Herwijnen,[35] E. Hicks,[49] D. Hill,[52] M. Hoballah,[5] C. Hombach,[51] P. Hopchev,[4] W. Hulsbergen,[38] P. Hunt,[52] T. Huse,[49] N. Hussain,[52] D. Hutchcroft,[49] D. Hynds,[48] V. Iakovenko,[41] P. Ilten,[12] J. Imong,[43] R. Jacobsson,[35] A. Jaeger,[11] E. Jans,[38] F. Jansen,[38] P. Jaton,[36] F. Jing,[3] M. John,[52] D. Johnson,[52] C. R. Jones,[44] B. Jost,[35] M. Kaballo,[9] S. Kandybei,[40] M. Karacson,[35] T. M. Karbach,[35] I. R. Kenyon,[42] U. Kerzel,[35] T. Ketel,[39] A. Keune,[36] B. Khanji,[20] O. Kochebina,[7] V. Komarov,[36,29] R. F. Koopman,[39] P. Koppenburg,[38] M. Korolev,[29] A. Kozlinskiy,[38] L. Kravchuk,[30] K. Kreplin,[11] M. Kreps,[45] G. Krocker,[11] P. Krokovny,[31] F. Kruse,[9] M. Kucharczyk,[20,23,j] V. Kudryavtsev,[31] T. Kvaratskheliya,[28,35] V. N. La Thi,[36] D. Lacarrere,[35] G. Lafferty,[51] A. Lai,[15] D. Lambert,[47] R. W. Lambert,[39] E. Lanciotti,[35] G. Lanfranchi,[18,35] C. Langenbruch,[35] T. Latham,[45] C. Lazzeroni,[42] R. Le Gac,[6] J. van Leerdam,[38] J.-P. Lees,[4] R. Lefèvre,[5] A. Leflat,[29,35] J. Lefrançois,[7] O. Leroy,[6] T. Lesiak,[23] Y. Li,[3] L. Li Gioi,[5] M. Liles,[49] R. Lindner,[35] C. Linn,[11] B. Liu,[3] G. Liu,[35] J. von Loeben,[20] J. H. Lopes,[2] E. Lopez Asamar,[33] N. Lopez-March,[36] H. Lu,[3] J. Luisier,[36] H. Luo,[47] A. Mac Raighne,[48] F. Machefert,[7] I. V. Machikhiliyan,[4,28] F. Maciuc,[26] O. Maev,[27,35] M. Maino,[20] S. Malde,[52] G. Manca,[15,d] G. Mancinelli,[6] N. Mangiafave,[44] U. Marconi,[14] R. Märki,[36] J. Marks,[11] G. Martellotti,[22] A. Martens,[8] L. Martin,[52] A. Martín Sánchez,[7] M. Martinelli,[38] D. Martinez Santos,[34] D. Martins Tostes,[2] A. Massafferri,[1] R. Matev,[35]

Z. Mathe,[35] C. Matteuzzi,[20] M. Matveev,[27] E. Maurice,[6] A. Mazurov,[16,30,35,e] J. McCarthy,[42] R. McNulty,[12] B. Meadows,[57] M. Meissner,[11] M. Merk,[38] D. A. Milanes,[13] M.-N. Minard,[4] J. Molina Rodriguez,[54] S. Monteil,[5] D. Moran,[51] P. Morawski,[23] R. Mountain,[53] I. Mous,[38] F. Muheim,[47] K. Müller,[37] R. Muresan,[26] B. Muryn,[24] B. Muster,[36] P. Naik,[43] T. Nakada,[36] R. Nandakumar,[46] I. Nasteva,[1] M. Needham,[47] N. Neufeld,[35] A. D. Nguyen,[36] T. D. Nguyen,[36] C. Nguyen-Mau,[36,o] M. Nicol,[7] V. Niess,[5] N. Nikitin,[29] T. Nikodem,[11] S. Nisar,[56] A. Nomerotski,[52,35] A. Novoselov,[32] A. Oblakowska-Mucha,[24] V. Obraztsov,[32] S. Oggero,[38] S. Ogilvy,[48] O. Okhrimenko,[41] R. Oldeman,[15,35,d] M. Orlandea,[26] J. M. Otalora Goicochea,[2] P. Owen,[50] B. K. Pal,[53] A. Palano,[13,b] M. Palutan,[18] J. Panman,[35] A. Papanestis,[46] M. Pappagallo,[48] C. Parkes,[51] C. J. Parkinson,[50] G. Passaleva,[17] G. D. Patel,[49] M. Patel,[50] G. N. Patrick,[46] C. Patrignani,[19,i] C. Pavel-Nicorescu,[26] A. Pazos Alvarez,[34] A. Pellegrino,[38] G. Penso,[22,l] M. Pepe Altarelli,[35] S. Perazzini,[14,c] D. L. Perego,[20,j] E. Perez Trigo,[34] A. Pérez-Calero Yzquierdo,[33] P. Perret,[5] M. Perrin-Terrin,[6] G. Pessina,[20] K. Petridis,[50] A. Petrolini,[19,i] A. Phan,[53] E. Picatoste Olloqui,[33] B. Pietrzyk,[4] T. Pilař,[45] D. Pinci,[22] S. Playfer,[47] M. Plo Casasus,[34] F. Polci,[8] G. Polok,[23] A. Poluektov,[45,31] E. Polycarpo,[2] D. Popov,[10] B. Popovici,[26] C. Potterat,[33] A. Powell,[52] J. Prisciandaro,[36] V. Pugatch,[41] A. Puig Navarro,[36] W. Qian,[4] J. H. Rademacker,[43] B. Rakotomiaramanana,[36] M. S. Rangel,[2] I. Raniuk,[40] N. Rauschmayr,[35] G. Raven,[39] S. Redford,[52] M. M. Reid,[45] A. C. dos Reis,[1] S. Ricciardi,[46] A. Richards,[50] K. Rinnert,[49] V. Rives Molina,[33] D. A. Roa Romero,[5] P. Robbe,[7] E. Rodrigues,[51,48] P. Rodriguez Perez,[34] G. J. Rogers,[44] S. Roiser,[35] V. Romanovsky,[32] A. Romero Vidal,[34] J. Rouvinet,[36] T. Ruf,[35] H. Ruiz,[33] G. Sabatino,[22,k] J. J. Saborido Silva,[34] N. Sagidova,[27] P. Sail,[48] B. Saitta,[15,d] C. Salzmann,[37] B. Sanmartin Sedes,[34] M. Sannino,[19,i] R. Santacesaria,[22] C. Santamarina Rios,[34] E. Santovetti,[21,k] M. Sapunov,[6] A. Sarti,[18,l] C. Satriano,[22,m] A. Satta,[21] M. Savrie,[16,e] P. Schaack,[50] M. Schiller,[39] H. Schindler,[35] S. Schleich,[9] M. Schlupp,[9] M. Schmelling,[10] B. Schmidt,[35] O. Schneider,[36] A. Schopper,[35] M.-H. Schune,[7] R. Schwemmer,[35] B. Sciascia,[18] A. Sciubba,[18,l] M. Seco,[34] A. Semennikov,[28] K. Senderowska,[24] I. Sepp,[50] N. Serra,[37] J. Serrano,[6] P. Seyfert,[11] M. Shapkin,[32] I. Shapoval,[40,35] P. Shatalov,[28] Y. Shcheglov,[27] T. Shears,[49,35] L. Shekhtman,[31] O. Shevchenko,[40] V. Shevchenko,[28] A. Shires,[50] R. Silva Coutinho,[45] T. Skwarnicki,[53] N. A. Smith,[49] E. Smith,[52,46] M. Smith,[51] K. Sobczak,[5] M. D. Sokoloff,[57] F. J. P. Soler,[48] F. Soomro,[18,35] D. Souza,[43] B. Souza De Paula,[2] B. Spaan,[9] A. Sparkes,[47] P. Spradlin,[48] F. Stagni,[35] S. Stahl,[11] O. Steinkamp,[37] S. Stoica,[26] S. Stone,[53] B. Storaci,[38] M. Straticiuc,[26] U. Straumann,[37] V. K. Subbiah,[35] S. Swientek,[9] M. Szczekowski,[25] P. Szczypka,[36,35] T. Szumlak,[24] S. T'Jampens,[4] M. Teklishyn,[7] E. Teodorescu,[26] F. Teubert,[35] C. Thomas,[52] E. Thomas,[35] J. van Tilburg,[11] V. Tisserand,[4] M. Tobin,[37] S. Tolk,[39] D. Tonelli,[35] S. Topp-Joergensen,[52] N. Torr,[52] E. Tournefier,[4,50] S. Tourneur,[36] M. T. Tran,[36] M. Tresch,[37] A. Tsaregorodtsev,[6] P. Tsopelas,[38] N. Tuning,[38] M. Ubeda Garcia,[35] A. Ukleja,[25] D. Urner,[51] U. Uwer,[11] V. Vagnoni,[14] G. Valenti,[14] R. Vazquez Gomez,[33] P. Vazquez Regueiro,[34] S. Vecchi,[16] J. J. Velthuis,[43] M. Veltri,[17,g] G. Veneziano,[36] M. Vesterinen,[35] B. Viaud,[7] D. Vieira,[2] X. Vilasis-Cardona,[33,n] A. Vollhardt,[37] D. Volyanskyy,[10] D. Voong,[43] A. Vorobyev,[27] V. Vorobyev,[31] C. Voß,[55] H. Voss,[10] R. Waldi,[55] R. Wallace,[12] S. Wandernoth,[11] J. Wang,[53] D. R. Ward,[44] N. K. Watson,[42] A. D. Webber,[51] D. Websdale,[50] M. Whitehead,[45] J. Wicht,[35] D. Wiedner,[11] L. Wiggers,[38] G. Wilkinson,[52] M. P. Williams,[45,46] M. Williams,[50,p] F. F. Wilson,[46] J. Wishahi,[9] M. Witek,[23] W. Witzeling,[35] S. A. Wotton,[44] S. Wright,[44] S. Wu,[3] K. Wyllie,[35] Y. Xie,[47,35] F. Xing,[52] Z. Xing,[53] Z. Yang,[3] R. Young,[47] X. Yuan,[3] O. Yushchenko,[32] M. Zangoli,[14] M. Zavertyaev,[10,a] F. Zhang,[3] L. Zhang,[53] W. C. Zhang,[12] Y. Zhang,[3] A. Zhelezov,[11] L. Zhong,[3] and A. Zvyagin[35]

(LHCb Collaboration)

[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]Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
[24]AGH University of Science and Technology, Kraków, Poland
[25]National Center for Nuclear Research (NCBJ), Warsaw, Poland
[26]Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
[27]Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
[28]Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
[29]Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
[30]Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
[31]Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia
[32]Institute for High Energy Physics (IHEP), Protvino, Russia
[33]Universitat de Barcelona, Barcelona, Spain
[34]Universidad de Santiago de Compostela, Santiago de Compostela, Spain
[35]European Organization for Nuclear Research (CERN), Geneva, Switzerland
[36]Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[37]Physik-Institut, Universität Zürich, Zürich, Switzerland
[38]Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
[39]Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands
[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]University of Birmingham, Birmingham, United Kingdom
[43]H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
[44]Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
[45]Department of Physics, University of Warwick, Coventry, United Kingdom
[46]STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
[47]School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
[48]School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
[49]Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
[50]Imperial College London, London, United Kingdom
[51]School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
[52]Department of Physics, University of Oxford, Oxford, United Kingdom
[53]Syracuse University, Syracuse, New York, USA
[54]Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil [associated with Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil]
[55]Institut für Physik, Universität Rostock, Rostock, Germany (associated with Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany)
[56]Institute of Information Technology, COMSATS, Lahore, Pakistan (associated with Syracuse University, Syracuse, New York, USA)
[57]University of Cincinnati, Cincinnati, Ohio, USA (associated Syracuse University, Syracuse, New York, USA)

[a]Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
[b]Also at Università di Bari, Bari, Italy.
[c]Also at Università di Bologna, Bologna, Italy.
[d]Also at Università di Cagliari, Cagliari, Italy.
[e]Also at Università di Ferrara, Ferrara, Italy.
[f]Also at Università di Firenze, Firenze, Italy.
[g]Also at Università di Urbino, Urbino, Italy.
[h]Also at Università di Modena e Reggio Emilia, Modena, Italy.
[i]Also at Università di Genova, Genova, Italy.
[j]Also at Università di Milano Bicocca, Milano, Italy.
[k]Also at Università di Roma Tor Vergata, Roma, Italy.

[l]Also at Università di Roma La Sapienza, Roma, Italy.
[m]Also at Università della Basilicata, Potenza, Italy.
[n]Also at LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
[o]Also at Hanoi University of Science, Hanoi, Vietnam.
[p]Also at Massachusetts Institute of Technology, Cambridge, MA, USA.