Search for the Lepton-Flavor-Violating Decays $B[0 \over s] e^{\pm}$ and $B[0] e^{\pm}$

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Search for the Lepton-Flavor-Violating Decays $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ and $B^{0} \rightarrow e^{\pm} \mu^{\mp}$

R. Aaij et al.*
(LHCb Collaboration)
(Received 18 July 2013; published 30 September 2013)

A search for the lepton-flavor-violating decays $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ and $B^{0} \rightarrow e^{\pm} \mu^{\mp}$ is performed with a data sample, corresponding to an integrated luminosity of 1.0 fb$^{-1}$ of $p p$ collisions at $\sqrt{s} = 7$ TeV, collected by the LHCb experiment. The observed number of $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ and $B^{0} \rightarrow e^{\pm} \mu^{\mp}$ candidates is consistent with background expectations. Upper limits on the branching fractions of both decays are determined to be $\mathcal{B}(B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}) < 1.1(1.4) \times 10^{-8}$ and $\mathcal{B}(B^{0} \rightarrow e^{\pm} \mu^{\mp}) < 2.8(3.7) \times 10^{-9}$ at 90% (95%) confidence level (C.L.). These limits are a factor of 20 lower than those set by previous experiments. Lower bounds on the Pati-Salam leptoquark masses are also calculated, $M_{LQ}(B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}) > 101$ TeV/$c^2$ and $M_{LQ}(B^{0} \rightarrow e^{\pm} \mu^{\mp}) > 126$ TeV/$c^2$ at 95% C.L., and are a factor of 2 higher than the previous bounds.


Rare decays that are forbidden in the standard model (SM) probe potential contributions from new processes and particles at a scale beyond the reach of direct searches. The decays $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ and $B^{0} \rightarrow e^{\pm} \mu^{\mp}$ and their charged conjugate processes (inclusion of charge conjugate processes are implied throughout this Letter) are forbidden within the SM, in which lepton flavor is conserved. These decays are allowed in some scenarios beyond the SM that include models with heavy singlet Dirac neutrinos [1], supersymmetric models [2], and the Pati-Salam model [3]. The latter predicts a new interaction to mediate transitions between leptons and quarks via exchange of spin-1 gauge bosons, called Pati-Salam leptoquarks (LQ), that carry both color and lepton quantum numbers.

Current limits from ATLAS [4–6] and CMS [7–9] on the masses of first, second, or third generation leptoquarks are in the range [0.4, 0.9] TeV/$c^2$, depending on the value of the couplings and the decay channel. These leptoquarks arise from a coupling between a quark and lepton of the same generation. The decays $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ and $B^{0} \rightarrow e^{\pm} \mu^{\mp}$ can be mediated by other leptoquarks that couple leptons and quarks that are not necessarily from the same generation [10,11], such as when the $\tau$ lepton couples to a first or second quark generation.

The previous best upper limits on the branching fraction of these decays come from the CDF Collaboration [12], $\mathcal{B}(B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}) < 2.0(2.6) \times 10^{-7}$ and $\mathcal{B}(B^{0} \rightarrow e^{\pm} \mu^{\mp}) < 6.4(7.9) \times 10^{-8}$ at 90% (95%) confidence level (C.L.). These limits correspond to bounds on the masses of the corresponding Pati-Salam leptoquarks of $M_{LQ}(B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}) > 47.8(44.9)$ TeV/$c^2$ and $M_{LQ}(B^{0} \rightarrow e^{\pm} \mu^{\mp}) > 59.3(56.3)$ TeV/$c^2$ at 90% (95%) C.L. [12].

This Letter presents a search for the $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ and $B^{0} \rightarrow e^{\pm} \mu^{\mp}$ lepton-flavor-violating (LFV) decays performed with a data sample, corresponding to an integrated luminosity of 1.0 fb$^{-1}$ of $p p$ collisions at $\sqrt{s} = 7$ TeV, collected by the LHCb experiment in 2011 at the Large Hadron Collider. To avoid potential bias, events in the signal mass region [5.1, 5.5] GeV/$c^2$ were not examined until all analysis choices were finalized.

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, and is described in detail in Ref. [13]. Events were simulated for this analysis using the software described in Refs. [14–20].

The trigger [21] consists of a hardware stage (L0), based on information from the calorimeter and muon systems, followed by a software stage (HLT) that applies a full event reconstruction, and is split into two stages called HLT1 and HLT2. Candidate $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ decays considered in this analysis must satisfy a hardware decision that requires the presence of a muon candidate with transverse momentum $p_T > 1.5$ GeV/$c$.

All tracks considered in the HLT1 are required to have $p_T > 0.5$ GeV/$c$. The muon track of the $B_{s}^{0} \rightarrow e^{\pm} \mu^{\mp}$ candidates is required to have $p_T > 1.0$ GeV/$c$ and impact parameter, IP $> 0.1$ mm. The HLT2 consists of exclusive, cut-based triggers for $B_{s}^{0}$ two-body decays, and inclusive multivariate [21,22] $b$-hadron triggers.

The $B^{0} \rightarrow K^{+} \pi^{-}$ decay is used as the normalization channel and $B_{s}^{0} \rightarrow h^{+} h^{-}$ ($h^{0} = K, \pi$) decays are used as a control channel, since both have the same event topology as the signal. The $B^{0} \rightarrow K^{+} \pi^{-}$ yield is computed from the yield of $B_{s}^{0} \rightarrow h^{+} h^{-}$ decays, and the fraction of $B^{0} \rightarrow K^{+} \pi^{-}$ in the $B_{s}^{0} \rightarrow h^{+} h^{-}$ sample, as described in Ref. [23]. In order to minimize the bias introduced by the trigger requirements, only $B_{s}^{0} \rightarrow h^{+} h^{-}$ candidates that

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

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

0031-9007/13/111(14)/141801(9) 141801-1 © 2013 CERN, for the LHCb Collaboration
are triggered independently of the presence of either of the two signal hadrons at L0 and HLT1 are considered.

The $B^+_s \rightarrow e^+ \mu^-$ candidates that pass the trigger selection criteria are further required to have well identified electron and muon [24] candidates. The measured momenta of the electrons are corrected for loss of momentum by bremsstrahlung in the detector using the photon energy deposition in the electromagnetic calorimeter [25]. The signal candidates are required to be displaced with respect to any $p p$ collision vertex (PV), and form a secondary vertex with $\chi^2$ per degree of freedom smaller than 9 and separated from the PV in the direction by a flight distance significance greater than 15.

Only $B^0_s$ candidates with an impact parameter $\chi^2 (\chi^2_{IP})$ less than 25 are considered. The $\chi^2_{IP}$ of a $B^0_s$ candidate is defined as the difference between the $\chi^2$ of the PV reconstructed with and without the considered candidate. When more than one PV is reconstructed, that giving the smallest $\chi^2_{IP}$ for the $B^0_s$ candidate is chosen. Only $B^0_s$ candidates with invariant mass in the range [4.9, 5.9] GeV/c$^2$ are kept for further analysis. The selection criteria for the $B^0_s \rightarrow h^+ h^-$ and $B^0 \rightarrow K^+ \pi^-$ candidates are identical to those of the signal, apart from those used for particle identification.

A two-stage multivariate selection based on boosted decision trees (BDTs) [26,27] is applied to the $B^0_s \rightarrow e^+ \mu^-$ candidates following the same strategy as Ref. [28]. The two multivariate discriminants are trained using simulated samples, $B^0_s \rightarrow e^+ \mu^-$ for signal and $b \bar{b} \rightarrow t\bar{t} h^+ h^- X$ for background (where $t(0)$ can either be a $\mu$ or an $e$ and $X$ is any other set of particles), which is dominated by simultaneous semileptonic decays of both $b$ and $\bar{b}$ hadrons within the same event.

The requirement on the first multivariate discriminant [28] removes 75% of the background while retaining 93% of signal, as determined from simulation using half of the available samples to train and the other half to evaluate the efficiencies. The same selection is applied to the $B^0 \rightarrow K^+ \pi^-$ normalization channel and the efficiencies of this requirement for the signal and normalization channel are equal within 1.2%, as determined from simulation.

The surviving background mainly comprises random combinations of electrons and muons from semileptonic $b \bar{b} \rightarrow e^+ \mu^- X$ decays. In total 5766 electron-muon pairs pass the trigger, the off-line selection, and the first multivariate discriminant requirements. The selected candidates are classified in a binned two-dimensional space formed by the electron-muon invariant mass and the output of a second BDT, for which nine variables are employed [28]. The BDT output is independent of the invariant mass for signal inside the search window. The output is transformed such that the signal is approximately uniformly distributed between zero and one, while the background peaks at zero.

The probability for a signal event to have a given BDT value is obtained from data using the $B^0_s \rightarrow h^+ h^-$ sample [29,30]. Simulated samples of $B^0_s \rightarrow e^+ \mu^-$ and $B^0_s \rightarrow h^+ h^-$ decays have been used to check that the distributions of the variables entering in the BDT that do not depend on the bremsstrahlung radiation are in good agreement. Corrections to the BDT shape due to the presence of the radiation emitted by the electron of the $B^0_s \rightarrow e^+ \mu^-$ or an $h^+ h^-$ signal events in each BDT bin is determined by fitting the $h^+ h^-$ invariant mass distribution. The systematic uncertainty on the signal BDT probability distribution function is taken to be the maximum spread in the fractions of yields going into each bin, obtained by fitting the same $B^0_s \rightarrow h^+ h^-$ data set with different signal and background fit models. Corrections are applied to the BDT shape in order to take into account the effect of the different trigger requirements used for the signal and the $B^0_s \rightarrow h^+ h^-$ control sample.

The invariant mass line shape of the signal events is described by a Crystal Ball function [31] with two tails, left and right, defined by two parameters each. The values of the parameters depend on the momentum resolution, the momentum scale, and the amount of bremsstrahlung radiation recovered.

The signal shape parameters are obtained from simulation, but need to be reweighted to account for their dependency on the event multiplicity, which affects the amount of bremsstrahlung radiation recovered and differs between data and simulation. We use the number of hits in the scintillating pad detector ($N_{SPD}$) as a measure of the event multiplicity. The distribution of $N_{SPD}$ for $B^0_s \rightarrow e^+ \mu^-$ signal candidates is obtained from a $B^0 \rightarrow J/\psi(\mu^+ \mu^-) K^+$ data sample, which is selected with the same trigger conditions as the signal, ensuring a similar distribution of $N_{SPD}$. The signal mass shape parameters are determined by reweighting the $B^0_s \rightarrow e^+ \mu^-$ simulated events with the $N_{SPD}$ distribution measured in the $B^0 \rightarrow J/\psi(\mu^+ \mu^-) K^+$ sample.

This reweighting technique is used also for a $J/\psi \rightarrow e^+ e^-$ simulated sample and the reweighted parameters are then compared with those obtained with a $J/\psi \rightarrow e^+ e^-$ sample in data. The difference between the mean values of the $J/\psi \rightarrow e^+ e^-$ mass in data and simulation (+0.16%) is applied as a systematic shift to the peak values of the $B^0_s \rightarrow e^+ \mu^-$ and $B^0_s \rightarrow e^- \mu^+$ invariant mass in simulation. A systematic uncertainty is added to the $B^0_s \rightarrow e^+ \mu^-$ mass parameters when the differences in the values of the other mass parameters for the $J/\psi \rightarrow e^+ e^-$ sample in data and SPD-reweighted simulation are larger than their statistical uncertainties.

The signal region, defined by the invariant mass window [5.1, 5.5] GeV/c$^2$, retains $(85.0 \pm 0.1_{stat} \pm 5.0_{syst})\%$ and $(82.0 \pm 0.1_{stat} \pm 5.0_{syst})\%$ of the $B^0_s \rightarrow e^+ \mu^-$ and $B^0 \rightarrow e^- \mu^+$ signal decays, respectively. The systematic
uncertainties on these fractions are evaluated with pseudoexperiments that fluctuate each parameter of the mass line shape according to its uncertainty. The width of the corresponding fraction distribution is taken as the systematic uncertainty.

The $B_s^0 \to e^\pm \mu^\mp$ and $B^0 \to e^\pm \mu^\mp$ yields are translated into branching fractions according to

$$\mathcal{B}(B_s^0 \to e^\pm \mu^\mp) = \frac{\mathcal{B}_{\text{norm}} \mathcal{B}_{\text{norm}} f_{\text{d}} \times N_{\text{d}}}{N_{\text{norm}} \mathcal{B}_{\text{norm}} f_{\text{d}}} \times N_{\text{d}} \to e^\pm \mu^\mp,$$

where $N_{\text{norm}} = 10,120 \pm 920$ is the number of signal events in the normalziation channel and is determined from the number of events in the $B^0 \to h^+h^-$ channel and the fraction of $B^0 \to K^\pm \pi^\mp$ events in the inclusive sample. The systematic uncertainty is comparable to the statistical one and is dominated by the maximum spread in the yield obtained by fitting the same $B_s^0 \to h^+h^-$ data set with different fit models [29,30]. The branching fraction of the normalization channel is $\mathcal{B}_{\text{norm}} = (1.94 \pm 0.06) \times 10^{-5}$ [32] and $N_{\text{d}}$ is the number of observed signal events. The factors $f_{\text{d}}$ and $f_s$ indicate the probabilities that a $b$ quark fragments into a $B^0$ or $B_s^0$ meson, respectively. We use $f_{s}/f_{\text{d}} = 0.256 \pm 0.020$ measured in $pp$ collision data at $\sqrt{s} = 7$ TeV [33]. The measured dependence of $f_{s}/f_{\text{d}}$ on the $B$ meson $p_T$ [33] is found to be negligible for this analysis.

The efficiency $\mathcal{E}_{\text{norm}}$ (signal) 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 ratios of acceptance, reconstruction, and selection efficiencies are computed with simulation. A systematic uncertainty is assigned to these ratios, to take into account the difference between the tracking efficiencies measured in data and predicted in simulation. Reweighting techniques are used to correct distributions in the simulation that do not match those from data, in particular for those variables that depend on $N_{\text{SPD}}$. The trigger efficiency of L0 and HLIT1 on signal decays is evaluated using data, while the HLIT2 efficiency is evaluated in simulation after validation with control samples. The electron and muon identification efficiencies are evaluated from data using the $B^+ \to J/\psi (B^{+}\mu^{-})K^+$ and $B^+ \to J/\psi (B^{+}\mu^{-})K^+$ control samples. The two normalization factors $\alpha_{B^0}$ and $\alpha_{B_s^0}$ are determined to be $\alpha_{B^0} = (1.1 \pm 0.2) \times 10^{-5}$ and $\alpha_{B_s^0} = (2.8 \pm 0.5) \times 10^{-10}$.

The BDT range is divided into eight bins with boundaries at 0.0, 0.25, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. The number of expected combinatorial background events in each BDT bin and the invariant mass signal region is determined from data by fitting to an exponential function the events in the mass sidebands, defined by $[4.9, 5.0]$ GeV/$c^2$ and $[5.5, 5.9]$ GeV/$c^2$. The invariant mass distributions of the selected candidates in BDT bins and the binned BDT distributions for the signals and the combinatorial background samples are available in the Supplemental Material [34].

In the exponential function both the slope and the normalization are 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. As a cross-check, two other models, the sum of two exponential functions and a single exponential fitted to the right sideband only, have been used and provide consistent background estimates inside the signal region.

The low-mass sideband and the signal region are potentially polluted by exclusive backgrounds. The background from $B_s^0 \to J/\psi (B^{+}\mu^{-})e^+\nu_e$ and $B_s^0 \to J/\psi (B^{+}\mu^{-})e^+\nu_e$ decays is evaluated assuming the branching fraction value from Ref. [35]. The decays $B^0 \to \pi^- l^- l^+$, $B^0_s \to h^+h^-$, $B_s^0 \to K^- l^+ l^-$, and $B^+ \to \pi^+ l^- l^+$ branching fractions are taken from Ref. [32]. The $B^+ \to \pi^+ l^- l^+$ branching fraction is taken from Ref. [36]. The theoretical estimates of the $\Lambda_b^0 \to p l^- l^+$ and $B_s^0 \to K^- l^+ l^-$ branching fractions are taken from Refs. [37,38], respectively. We use the $\Lambda_b^0$ fragmentation fraction $f_{\Lambda_b^0}$ measured by LHCb [39] and account for its $p_T$ dependence.

The mass and BDT distributions of these background modes are evaluated from simulated samples, using the probabilities of misidentifying a kaon, pion, and proton as a muon or electron as functions of momenta and transverse momenta, which are determined from $D^{*+} \to D^{0}(\to K^- \pi^+)^+ \pi^+$ and $\Lambda \to p \pi^- \pi^+$ data samples. The yield of the $B_s^0 \to h^+h^- \to e^+\mu^-$ peaking background in each BDT bin is obtained by multiplying the $B_s^0 \to h^+h^-$ yields obtained by fitting the invariant mass distribution of an inclusive $B_s^0 \to h^+h^-$ sample in BDT bins [29,30] with the probabilities of misidentifying a kaon, pion, and proton as a muon or electron as functions of momenta and transverse momenta, as determined from control samples. The mass line shape of the $B_s^0 \to h^+h^- \to e^+\mu^-$ peaking background is obtained from a simulated sample of doubly misidentified $B_s^0 \to h^+h^-$ events. Apart from $B^0 \to h^+h^-$, all background modes are normalized relative to the $B^+ \to J/\psi (B^{+}\mu^{-})K^+$ decay. We assume $f_{\text{d}} = f_{\text{d}}$ where $f_{\text{u}}$ is the $B^+$ fragmentation fraction.

The $\Lambda_b^0 \to p l^- l^+$ and the $B_s^0 \to J/\psi (B^{+}\mu^{-})e^+\nu_e$ and $B_s^0 \to J/\psi (B^{+}\mu^{-})e^+\nu_e$ modes are the dominant exclusive modes in the range BDT > 0.5, where the combinatorial background is reduced by a factor $\sim 500$ according to simulation. These decay modes have an invariant mass

\[ J=\{f_{\Lambda_b^0} \to p l^- l^+ \} \]
distribution that is compatible with an exponential in the region [4.9, 5.9] GeV/c², and hence are taken into account by the exponential fit to the mass sidebands.

In the entire BDT and mass range ([4.9, 5.9] GeV/c²), 4.5 ± 0.7 doubly misidentified $B_{s}^{0} \rightarrow h^{+} h^{-}$ decays are expected, with (87.9% ± 0.1%) lying in the signal mass interval of [5.1, 5.5] GeV/c².

For each BDT bin we count the number of candidates observed in the signal region, and compare to the expected number of signal and background events.

The systematic uncertainties in the background and signal predictions in each bin are computed by varying the normalization factor, and the mass and BDT shapes within their Gaussian uncertainties.

The results for the $B_{s}^{0} \rightarrow e^{±} \mu^{±}$ and $B^{0} \rightarrow e^{±} \mu^{±}$ decays are summarized in Table 1. In the high BDT range, the observed number of candidates is in agreement with the expected, with (90% and 95% C.L. are shown in Table II. Note that since the same events are used to set limits for both $B_{s}^{0}$ and $B^{0}$ decays, the results are strongly correlated. The inclusion of systematic uncertainties increases the expected upper limits by ~20%.

The systematic uncertainties are dominated by the uncertainty in the interpolation of the background yields inside the signal region. The observed limits are ~1σ below the expectation due to the lower than expected numbers of observed events in the fourth and last BDT bins.

In the framework of the Pati-Salam model, the relation linking the $B_{s}^{0} \rightarrow e^{±} \mu^{±}$ branching fraction and the leptoquark mass ($M_{LQ}$) [10] is

$$B(B_{s}^{0} \rightarrow e^{±} \mu^{±}) = \frac{\alpha_{3}(M_{LQ})^{2}}{M_{LQ}} \cdot \frac{F_{B_{s}^{0}}}{m_{B}} \cdot \frac{m_{B}}{m_{B}^{0}} \frac{\tau_{B_{s}^{0}}}{\hbar},$$

where

$$R = \frac{m_{B_{s}^{0}}}{m_{B}} \left( \frac{\alpha_{3}(M_{LQ})}{\alpha_{3}(m_{B})} \right)^{-4/7} \left( \frac{\alpha_{3}(m_{B})}{\alpha_{3}(m_{B})} \right)^{-12/23}.$$

The $B^{0}$ and $B_{s}^{0}$ masses $m_{B^{0}}$ and $m_{B_{s}^{0}}$ and the average lifetimes $\tau_{B^{0}}$ and $\tau_{B_{s}^{0}}$ are taken from Ref. [32]. The factors $F_{B_{s}^{0}} = 0.190 ± 0.004$ GeV and $F_{B^{0}} = 0.227 ± 0.004$ GeV are the decay constants of the $B^{0}$ and $B_{s}^{0}$ mesons [41].
TABLE II. Expected (background only) and observed limits on the $B^0 \rightarrow e^\pm \mu^\mp$ branching fractions.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Limit</th>
<th>90% C.L.</th>
<th>95% C.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow e^\pm \mu^\mp$</td>
<td>Expected</td>
<td>$1.5 \times 10^{-8}$</td>
<td>$1.8 \times 10^{-8}$</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>$1.1 \times 10^{-8}$</td>
<td>$1.4 \times 10^{-8}$</td>
</tr>
<tr>
<td>$B^0 \rightarrow e^\pm \mu^\mp$</td>
<td>Expected</td>
<td>$3.8 \times 10^{-9}$</td>
<td>$4.8 \times 10^{-9}$</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>$2.8 \times 10^{-9}$</td>
<td>$3.7 \times 10^{-9}$</td>
</tr>
</tbody>
</table>

and $m_b$ and $m_t$ are the bottom and top quark masses [32], respectively, computed in the $\overline{\text{MS}}$ scheme [42]. The value of $\alpha_s$ at an arbitrary scale $M_{LQ}$ is determined using the software package RUNDEC [43].

Using the limits on the branching fractions shown in Table II, we find the following lower bounds for the leptoquark masses if the leptoquark links the $\tau$ lepton to the first and second quark generation, $M_{LQ}(B_s^0 \rightarrow e^\pm \mu^\mp) > 107(101)$ TeV/$c^2$ and $M_{LQ}(B^0 \rightarrow e^\pm \mu^\mp) > 135(126)$ TeV/$c^2$ at 90% (95%) C.L., respectively. When the parameters entering in Eq. (2) are fluctuated within $\pm 1 \sigma$, the limits on the leptoquark masses change by $\sim \pm 1$ TeV.

In summary, a search for the lepton-flavor-violating decays $B_s^0 \rightarrow e^\pm \mu^\mp$ and $B^0 \rightarrow e^\pm \mu^\mp$ has been performed on a data sample, corresponding to an integrated luminosity of 1.0 fb$^{-1}$, collected in $pp$ collisions at $\sqrt{s} = 7$ TeV. The data are consistent with the background-only hypothesis. Upper limits are set on the branching fractions, $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 1.1(1.4) \times 10^{-8}$ and $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 2.8(3.7) \times 10^{-9}$ at 90% (95%) C.L., that are the most restrictive to date. These limits translate into lower bounds on the leptoquark masses in the Pati-Salam model [10] of $M_{LQ}(B_s^0 \rightarrow e^\pm \mu^\mp) > 107(101)$ TeV/$c^2$ and $M_{LQ}(B^0 \rightarrow e^\pm \mu^\mp) > 135(126)$ TeV/$c^2$ at 90% (95%) C.L., respectively. These are a factor of 2 higher than the previous bounds.

We thank Diego Guadagnoli for the theory inputs. 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); MEN/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 (United Kingdom). 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.

O. Yushchenko,1 M. Zangoli,14 M. Zavertyaev,10,5 F. Zhang,3 L. Zhang,58 W. C. Zhang,12 Y. Zhang,3 A. Zhelezov,11 A. Zhokhov,30 L. Zhong,3 and A. Zvyagin37

(LHCb Collaboration)

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