Search for the Rare Decays $B_s^0 \rightarrow e^+e^-$ and $B^0 \rightarrow e^+e^-$

LHCb Collaboration; Bernet, R; Müller, K; Owen, P; Serra, N; Steinkamp, O; et al

DOI: https://doi.org/10.1103/PhysRevLett.124.211802

Original published at:
LHCb Collaboration; Bernet, R; Müller, K; Owen, P; Serra, N; Steinkamp, O; et al (2020). Search for the Rare Decays $B_s^0 \rightarrow e^+e^-$ and $B^0 \rightarrow e^+e^-$. Physical Review Letters, 124(21):211802.
DOI: https://doi.org/10.1103/PhysRevLett.124.211802
Search for the Rare Decays $B^0_s \rightarrow e^+e^-$ and $B^0 \rightarrow e^+e^-$

R. Aaij et al.
(LHCb Collaboration)

A search for the decays $B^0_s \rightarrow e^+e^-$ and $B^0 \rightarrow e^+e^-$ is performed using data collected with the LHCb experiment in proton-proton collisions at center-of-mass energies of 7, 8, and 13 TeV, corresponding to integrated luminosities of 1, 2, and 2 fb$^{-1}$, respectively. No signal is observed. Assuming no contribution from $B^0 \rightarrow e^+e^-$ decays, an upper limit on the branching fraction $\mathcal{B}(B^0_s \rightarrow e^+e^-) < 9.4(11.2) \times 10^{-9}$ is obtained at 90(95)% confidence level. If no $B^0_s \rightarrow e^+e^-$ contribution is assumed, a limit of $\mathcal{B}(B^0 \rightarrow e^+e^-) < 2.5(3.0) \times 10^{-9}$ is determined at 90(95)% confidence level. These upper limits are more than one order of magnitude lower than the previous values.

DOI: 10.1103/PhysRevLett.124.211802

Searches for rare particle decays provide ideal probes for contributions from physics processes beyond the standard model (SM). Recent measurements of decays involving $b \rightarrow s \ell^+\ell^-$ transitions (the inclusion of charge-conjugated processes is implied throughout this Letter) hint at deviations from SM predictions in lepton-flavor universality tests [1–6] and thus motivate measurements of decay rates into final states involving leptons. Following the observation of the decay $B^0 \rightarrow \mu^+\mu^-$ [7,8], the search for $B^0_s \rightarrow e^+e^-$ and $B^0 \rightarrow e^+e^-$ decays provides an independent test of lepton-flavor universality. According to SM predictions (calculated from Ref. [9], neglecting QED corrections that are expected to be at the percent level), $B^0_s \rightarrow e^+e^-$ decays have branching fractions of $\mathcal{B}(B^0_s \rightarrow e^+e^-) = (8.60 \pm 0.36) \times 10^{-14}$ and $\mathcal{B}(B^0 \rightarrow e^+e^-) = (2.41 \pm 0.13) \times 10^{-15}$. With contributions beyond the SM, these branching fractions could be significantly larger, reaching values of $\mathcal{O}(10^{-8})$ for $\mathcal{B}(B^0_s \rightarrow e^+e^-)$ and $\mathcal{O}(10^{-10})$ for $\mathcal{B}(B^0 \rightarrow e^+e^-)$ [10]. These values are close to the current experimental bounds of $\mathcal{B}(B^0_s \rightarrow e^+e^-) < 2.8 \times 10^{-7}$ and $\mathcal{B}(B^0 \rightarrow e^+e^-) < 8.3 \times 10^{-8}$ at 90% confidence level (CL) [11], set by the CDF collaboration.

In this Letter, a search for $B^0_s \rightarrow e^+e^-$ and $B^0 \rightarrow e^+e^-$ decays is presented using data collected with the LHCb experiment in proton-proton collisions at center-of-mass energies of 7 TeV in 2011, 8 TeV in 2012 and 13 TeV in 2015 and 2016, corresponding to integrated luminosities of 1, 2, and 2 fb$^{-1}$, respectively. The signal yields are determined from a fit to the data and normalized to those of the $B^+ \rightarrow J/\psi K^+$ decay, where the $J/\psi$ meson decays to $e^+e^-$, which has a precisely measured branching fraction [12] and a similar dielectron signature in the detector.

The LHCb detector [13,14] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons, and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The online event selection is performed by a trigger [15], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. At the hardware trigger stage, events are required to have a high-energy deposit in the calorimeters associated with a signal electron candidate, or a muon candidate with high transverse momentum $p_T$, or a photon, electron, or hadron candidate with high transverse energy from the decays of other particles from the $pp$ collision. The software trigger requires a two-track secondary vertex with a significant displacement from any primary $pp$ interaction vertex (PV). At least one charged particle must have high $p_T$ and be inconsistent with originating from a PV. A multivariate algorithm [16,17] is used in the trigger for the identification of secondary vertices consistent with the decay of a $b$
hadron. Simulated samples are used to optimize the candidate selection, estimate selection efficiencies and describe the expected invariant-mass shapes of the signal candidates and background decays. In the simulation, $pp$ collisions are generated using PYTHIA [18] with a specific LHCb configuration [19]. Decays of unstable particles are described by EVTGEN [20], in which final-state radiation is generated using PHOTOS [21]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [22] as described in Ref. [23]. The simulation is corrected for data-simulation differences in $B$-meson production kinematics, detector occupancy, and isolation criteria [24] using $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow J/\psi \phi$ decays, with $J/\psi \rightarrow e^+ e^-$ and $\phi \rightarrow K^+ K^-$. Particle identification variables are calibrated using data from $B^+ \rightarrow J/\psi K^+$ and $D^0 \rightarrow K^- \pi^+$ decays [25]. The calibration data are binned in momentum and pseudorapidity of the particle as well as detector occupancy to account for possible differences in kinematics between the investigated decay and the calibration data.

The $B^0_{(s)} \rightarrow e^+ e^-$ candidates are selected in events passing the trigger requirements by combining two tracks that are inconsistent with originating from any PV in the event and which form a good-quality secondary vertex. The tracks are also required to have a momentum larger than 3 GeV/$c$ and $p_T$ greater than 500 MeV/$c$, and must be identified as electrons using information from the Cherenkov detectors and calorimeters. The dielectron candidate’s momentum must be aligned with the vector pointing from a PV (the associated PV) to the two-track vertex and have a considerable transverse component. The candidate must also have an invariant mass in the range [4166, 6566] MeV/$c^2$.

The measured electron momenta are corrected for losses due to bremsstrahlung radiation by adding the momentum of photons consistent with being emitted upstream of the magnet [26]. Candidates in data and simulation are separated into three categories with either zero, one, or both electrons having a bremsstrahlung correction applied. To avoid experimenters’ bias, the narrowest dielectron invariant-mass region containing 90% of simulated $B^0_{(s)} \rightarrow e^+ e^-$ decays, corresponding to a range of [4689, 5588] MeV/$c^2$, was removed from the data set until the analysis procedure was finalized.

Candidates for the normalization mode, $B^+ \rightarrow J/\psi K^+$, are constructed similarly, but require an additional track consistent with being a kaon and originating from the same vertex as the dielectron candidate. The dielectron candidate must have an invariant mass in the range [2450, 3176] MeV/$c^2$, consistent with arising from a $J/\psi$ meson decay. In addition, the reconstructed $B^+$-candidate mass, when the dielectron candidate is constrained to the known $J/\psi$ mass [12], must be above 5175 MeV/$c^2$, suppressing partially reconstructed decays.

A boosted decision tree (BDT) algorithm [27–29] is used to separate $B^0_{(s)} \rightarrow e^+ e^-$ signal from random combinations of two electrons (combinatorial background). The BDT is trained separately for data taking periods 2011–2012 (Run 1) and 2015–2016 (Run 2) on simulated $B^0_{(s)} \rightarrow e^+ e^-$ decays as signal proxy and dielectron candidates from data with a mass above 5588 MeV/$c^2$ as background proxy. The split between the data taking periods is done to account for changes in the center-of-mass energies and trigger strategies, which significantly impact the data distributions and improve the BDT and the particle identification algorithms in Run 2. It is checked that the data behave consistently across the data-taking periods. The BDT input variables comprise of the following: kinematic information on the electron tracks and $B$ candidate, information on the displacement of the electrons and $B$ candidate from the associated PV, and isolation variables that quantify the compatibility of other tracks in the event with originating from the same decay as the $B$ candidate [24,30]. Candidates with a BDT response compatible with that of the background are discarded, with the threshold chosen by maximizing the figure of merit $c_{\text{signal}}(\sqrt{N_{\text{background}}} + 3/2)$ [31], where $c_{\text{signal}}$ is the signal efficiency and the expected background yield in the signal region is $N_{\text{background}}$.

The final selected data set is separated by data-taking period and by category of bremsstrahlung correction. The branching fraction $B(B^0_{(s)} \rightarrow e^+ e^-)$ is measured relative to that of the normalization channel via

$$B(B^0_{(s)} \rightarrow e^+ e^-) = N(B^0_{(s)} \rightarrow e^+ e^-) \times \alpha \times B(B^+ \rightarrow J/\psi K^+) \times \left( \frac{f_{d(s)}}{f_u} \right)^{-1},$$

where

$$\alpha \equiv \frac{e(B^+ \rightarrow J/\psi K^+)}{e(B^0_{(s)} \rightarrow e^+ e^-)} \times \frac{1}{N(B^+ \rightarrow J/\psi K^+)},$$

and $e(B^0_{(s)} \rightarrow e^+ e^-)$ and $e(B^+ \rightarrow J/\psi K^+)$ denote the efficiencies of the signal and normalization modes, and $N(B^0_{(s)} \rightarrow e^+ e^-)$ and $N(B^+ \rightarrow J/\psi K^+)$ their yields. The normalization mode branching fraction (including that for the decay $J/\psi \rightarrow e^+ e^-$) is $B(B^+ \rightarrow J/\psi K^+) = (6.03 \pm 0.17) \times 10^{-5}$, taken from Ref. [12]. The $b$-hadron fragmentation fraction ratio $f_{d(s)}/f_u$ is assumed to be unity, while $f_s/f_u = 0.259 \pm 0.015$ [32] is used for the Run 1 data and is scaled by 1.068 ± 0.016 for the Run 2 data, according to Ref. [33], to account for center-of-mass energy differences. A measurement of $f_s/f_u$ from Run 2 yields a consistent, but less precise, result [34].

The yield of the normalization mode is determined using an unbinned maximum-likelihood fit to the $K^+ e^+ e^-$ invariant mass separately for each year of data.
taking and bremsstrahlung category. The fit model comprises a Gaussian function with power-law tails [35] for the signal component, where the tail parameters are fixed from simulation, and an exponential function to describe combinatorial background. Summed over the bremsstrahlung categories, the yield of the normalization mode is 20480 ± 140 in the Run 1 data and 33080 ± 180 in the Run 2 data.

The selection efficiencies \( \epsilon(B^0_{sl} \rightarrow e^+e^-) \) and \( \epsilon(B^+ \rightarrow J/\psi K^+) \) are determined separately for each year of data taking and bremsstrahlung category using simulated decays that are weighted to better represent the data. Calibration data are used to evaluate particle-identification efficiencies [25]. Trigger efficiencies are also estimated from data, using the technique described in Ref. [36]. For simulated \( B^0_{sl} \rightarrow e^+e^- \) decays, the mean \( B^0_{sl} \) lifetime [37] is assumed. The selection efficiency is assumed to be the same for both \( B^0 \rightarrow e^+e^- \) and \( B^0 \rightarrow e^+e^- \) decays, which is consistent with results from simulation. The normalization factors, \( \alpha \), are combined across the data-taking periods and, given in Table I, split by bremsstrahlung category (for the selection efficiency ratio between normalization and signal mode, see the Supplemental Material [38]).

In addition to the combinatorial background, backgrounds due to misidentification and partial reconstruction are present in the data. These backgrounds differ significantly between the categories of bremsstrahlung correction. Their invariant-mass shapes and relative contributions are evaluated using simulation. In the lower mass region, partially reconstructed backgrounds of the types \( B \rightarrow Xe^+e^- \) and \( B^+ \rightarrow \bar{D}^0\rightarrow Y^+e^-\bar{\nu}_e e^+\nu_e \) dominate, where \( X \) and \( Y \) represent hadronic systems. The main source of background in the \( B \)-mass region, however, stems from misidentified particles in the decays \( B^0 \rightarrow \pi^-e^+\nu_e \) and \( B \rightarrow h^+h^- \), where \( h \) and \( h' \) are hadrons. The latter has a peaking structure in the \( B \)-mass region. Backgrounds involving misidentified particles contribute mostly to categories in which at most one of the electrons has a bremsstrahlung correction applied. The contribution from combinatorial background is evaluated from same-sign lepton pairs in data and found to be small. The yields of the backgrounds are Gaussian constrained to their expected values, estimated from simulation using their known branching fractions [12].

TABLE I. Normalization factors \( \alpha \) for \( B^0_{sl} \rightarrow e^+e^- \). The bremsstrahlung category denotes whether zero, one or both electrons are corrected for bremsstrahlung losses. The uncertainties include statistical uncertainties and uncertainties due to limited size of the simulated samples.

| Bremsstrahlung category | 2011–2012 [10^{-5}] | 2015–2016 [10^{-5}] |
|--------------------------|----------------------|----------------------|
| No correction            | 2.85 ± 0.24          | 1.84 ± 0.08          |
| One electron corrected   | 1.13 ± 0.08          | 0.70 ± 0.03          |
| Both electrons corrected | 1.73 ± 0.20          | 1.04 ± 0.06          |

The shape of the invariant mass of the \( B^0_{sl} \rightarrow e^+e^- \) and \( B^0 \rightarrow e^+e^- \) components is modeled using a Gaussian function with power-law tails, where the parameters are obtained from simulation and differ between each bremsstrahlung category and year of data taking. The peak values and the widths of the functions are corrected for data-simulation differences by a factor determined from the normalization mode. The parameters of the \( B^0_{sl} \rightarrow e^+e^- \) and \( B^0 \rightarrow e^+e^- \) line shapes are fixed to the same values with the exception of the peak value, which is shifted according to the known \( B^0_{sl}-B^0 \) mass difference [12]. Due to the limited mass resolution, arising from imperfect bremsstrahlung recovery, the line shapes from \( B^0_{sl} \rightarrow e^+e^- \) and \( B^0 \rightarrow e^+e^- \) are highly overlapping. Therefore the branching fraction of \( B^0_{sl} \rightarrow e^+e^- \) is obtained by performing a simultaneous fit to the dielectron invariant-mass distribution of all six data sets while neglecting the contribution from \( B^0 \rightarrow e^+e^- \), and vice versa. In these fits, the only shared parameters between categories are the branching fractions \( B(B^0_{sl} \rightarrow e^+e^-) \) and \( B(B^+ \rightarrow J/\psi K^+) \), and the ratio of the fragmentation fractions \( f_s/f_o \).

Systematic uncertainties are estimated separately for each data set. Dominant sources of systematic uncertainties in the normalization arise from the uncertainty on the fragmentation fraction ratio, the technique used to evaluate the trigger efficiencies, and the determination of particle-identification efficiencies; the systematic uncertainties from these sources extend to 5.8%, 5.3%, and 5.3% on the branching fractions, respectively. The uncertainty on \( B(B^+ \rightarrow J/\psi K^+) \) of 2.8% [12] is taken into account. A difference of up to 4.1% is found between the efficiency of the BDT selection on simulated \( B^+ \rightarrow J/\psi K^+ \) decays and \( B^+ \rightarrow J/\psi K^+ \) decays in data, which is assigned as a systematic uncertainty. The fraction of candidates in each bremsstrahlung-correction category of the signal modes is taken from simulation. The difference between simulation and data is investigated using \( B^+ \rightarrow J/\psi K^+ \) decays and its effect on the normalization, up to 4.0%, is taken as a systematic uncertainty. Systematic uncertainties on the invariant-mass resolution corrections are determined by repeating the correction procedure with pseudoexperiments obtained with the bootstrapping method [39], yielding up to 1.1%. A difference between the total selection efficiencies in the \( B^0_{sl} \rightarrow e^+e^- \) and \( B^0 \rightarrow e^+e^- \) channels of up to 2.5% is assigned as a systematic uncertainty on the \( B^0 \rightarrow e^+e^- \) normalization factor. Due to the presence of an additional kaon in the final state of the normalization mode, the track-reconstruction efficiency is different between the signal and normalization modes. An uncertainty of 1.1% is assigned to the branching fraction as a systematic uncertainty on the kaon reconstruction efficiency arising from the limited knowledge of the interactions in the detector material [40]. Finally, an uncertainty of 1.0% is assigned to account for small differences in detector occupancy between the signal
and normalization mode arising from the trigger selection. The dominant sources of systematic uncertainties on the background composition are due to the imprecise knowledge of the branching fractions of the background components. The largest uncertainty of this type on the expected background yield in the $B$-mass region is 14%, determined from refitting the mass sidebands while varying the background components according to their uncertainties. Taking all correlations into account, overall single event sensitivities are shown in Fig. 1, with the result of the fit. The measured branching fraction is obtained.

The dielectron invariant-mass spectrum, summed over bremsstrahlung categories, is shown in Fig. 1, with the result of the $B^0 \rightarrow e^+e^-$ fit. The individual categories are shown in the Supplemental Material [38], as well as the distributions with the result of the $B^0 \rightarrow e^+e^-$ fit. The measured branching fractions are $B(B^0 \rightarrow e^+e^-) = (2.4 \pm 4.4) \times 10^{-9}$ and $B(B^0 \rightarrow e^+e^-) = (0.30 \pm 1.29) \times 10^{-9}$, where the uncertainties include both statistical and systematic components. The results are in agreement with the background-only hypothesis.

Upper limits on the branching fractions are set using the CL$_s$ method [41], as implemented in the GAMMACOMBO framework [42,43] with a one-sided profile likelihood ratio [44] as test statistic. The likelihoods are computed from fits to the invariant-mass distributions. In the fits, the normalization factor, normalization mode branching fraction, fragmentation fraction ratio, and background yields are Gaussian constrained to their expected values within statistical and systematic uncertainties. Pseudoexperiments, in which the nuisance parameters are set to their fitted values from data, are used for the evaluation of the test statistic.

The expected and observed CL$_s$ distributions are shown in Fig. 2. The upper limits are $B(B^0 \rightarrow e^+e^-) < 9.4(11.2) \times 10^{-9}$ and $B(B^0 \rightarrow e^+e^-) < 2.5(3.0) \times 10^{-9}$ at 90(95)% confidence level. These are consistent with the expected upper limits of $B(B^0 \rightarrow e^+e^-) < 7.0(8.6) \times 10^{-9}$ and $B(B^0 \rightarrow e^+e^-) < 2.0(2.5) \times 10^{-9}$ at 90(95)% confidence level, obtained as the median of limits determined on background-only pseudoexperiments.

In conclusion, a search for the rare decays $B^0 \rightarrow e^+e^-$ is performed using data from proton-proton collisions

![Figure 1](image1.png)  
**FIG. 1.** Simultaneous fit to the dielectron invariant-mass distribution, with $B(B^0 \rightarrow e^+e^-)$ fixed to zero. The sum of bremsstrahlung categories is shown for (left) Run 1 and (right) Run 2. The relative proportions of background contributions change between Run 1 and Run 2 due to different performances of the particle identification algorithms and BDT selections.

![Figure 2](image2.png)  
**FIG. 2.** CL$_s$ values as a function of the branching fractions of the decays (left) $B^0_s \rightarrow e^+e^-$ and (right) $B^0 \rightarrow e^+e^-$. The red solid line (black solid line with data points) corresponds to the distribution of the expected (observed) upper limits, and the light blue (dark blue) band contains the 1σ (2σ) uncertainties on the expected upper limits. Thresholds corresponding to 90% and 95% confidence level are indicated with dashed lines. The observed values are plotted for branching fractions greater than the measured branching fraction in the data; the test statistic is defined to be nonzero only in that region.
recorded with the LHCb experiment, corresponding to a total integrated luminosity of 5 fb$^{-1}$. No excess of events is observed over the background. The resulting limits on the branching fractions are $\mathcal{B}(B^0_\ell \rightarrow e^+e^-) < 9.4(11.2) \times 10^{-9}$ and $\mathcal{B}(B^0 \rightarrow e^+e^-) < 2.5(3.0) \times 10^{-9}$ at 90(95)% confidence level, when neglecting the contribution from the other decay. The mean $B^0_\ell$ lifetime is assumed for $B^0_\ell \rightarrow e^+e^-$ decays. Assuming SM-like CP-odd (CP-even) $B^0_\ell \rightarrow e^+e^-$ decays, an increase (decrease) of 2.4% with respect to the quoted limit is found. The results improve the limits on these branching fractions [11] by more than one order of magnitude and constrain contributions beyond the SM, for example from scalar and pseudoscalar currents [10].

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); MOST and NSFC (China); CRNS/IN2P3 (France); BMBF, DFG, and MPG (Germany); INFN (Italy); NWO (Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MSHE (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); DOE NP and NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany); EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union); ANR, Labex P2IO and OCEVU, and Région Auvergne-Rhône-Alpes (France); Key Research Program of Frontier Sciences of CAS, CAS PIFI, and the Thousand Talents Program (China); RFBR, RSF, and Yandex LLC (Russia); GVA, XuntaGal, and GENCAT (Spain); the Royal Society and the Leverhulme Trust (United Kingdom).

[1] R. Aaij et al. (LHCb Collaboration), Test of lepton universality with $B^0 \rightarrow K^{(*)}e^+\ell^-$ decays, J. High Energy Phys. 08 (2017) 055.
[2] R. Aaij et al. (LHCb Collaboration), Search for Lepton-Universality Violation in $B^+ \rightarrow K^+\ell^+\ell^-$ Decays, Phys. Rev. Lett. 122, 191801 (2019).
[3] R. Aaij et al. (LHCb Collaboration), Test of lepton universality using $A_0^0 \rightarrow pK^{*}e^+\ell^-$ decays, J. High Energy Phys. 05 (2020) 040.
[23] M. Clemencic, G. Corti, S. Easo, C. R. Jones, S. Miglioranz, M. Pappalagallo, and P. Robbe, The LHCb simulation application, Gaussion: design, evolution and experience, J. Phys. Conf. Ser. 331, 032023 (2011).

[24] R. Aaij et al. (LHCb Collaboration), Search for the rare decays $B^0 \to \mu^+ \mu^-$ and $B^0 \to \mu^+ \mu^-$, Phys. Lett. B 708, 55 (2012).

[25] L. Anderlini et al., The PIDCalib package, Report No. LHCb-PUB-2016-021, 2016.

[26] R. Aaij et al. (LHCb Collaboration), Measurement of the $B^0 \to K^{\pm} e^+ e^-$ branching fraction at low dilepton mass, J. High Energy Phys. 05 (2013) 159.

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

[28] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees (Wadsworth International Group, Belmont, California, USA, 1984).

[29] F. Pedregosa et al., Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12, 2825 (2011), and online at http://scikit-learn.org/stable/.

[30] L. Gavardi, Search for lepton flavour violation in $\tau$ decays at the LHCb experiment, Master's thesis, Universit degli studi di Milano-Bicocca, 2013, CERN-THESIS-2013-259.

[31] G. Punzi, Sensitivity of searches for new signals and its optimization, eConf C030908, MODT002 (2003).

[32] R. Aaij et al. (LHCb Collaboration), Measurement of the fragmentation fraction ratio $f_2/f_4$ and its dependence on $B$ meson kinematics, J. High Energy Phys. 04 (2013) 001, $f_2/f_4$ value updated in LHCb-CONF-2013-011.

[33] R. Aaij et al. (LHCb Collaboration), Measurement of $f_2/f_4$ Variation with Proton-Proton Collision Energy and $B$-Meson Kinematics, Phys. Rev. Lett. 124, 120002 (2020).

[34] R. Aaij et al. (LHCb Collaboration), Measurement of $b$-hadron fractions in 13 TeV $pp$ collisions, Phys. Rev. D 100, 031102(R) (2019).

[35] T. Skwarnicki, A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances, Ph.D. thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[36] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, Data driven trigger efficiency determination at LHCb, Report No. LHCb-PUB-2014-039, 2014.

[37] K. A. Olive et al. (Particle Data Group), Review of particle physics, Chin. Phys. C 38, 090001 (2014).

[38] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.211802 for the selection efficiency ratio of normalization and signal mode, the individual categories of the $B^0 \to e^+ e^-$ fit and the distributions with the result of the $B^0 \to e^+ e^-$ fit.

[39] B. Efron and R. J. Tibshirani, An Introduction to the Bootstrap, Mono. Stat. Appl. Probab. (Chapman and Hall, London, 1993).

[40] R. Aaij et al. (LHCb Collaboration), Measurement of the track reconstruction efficiency at LHCb, J. Instrum. 10, P02007 (2015).

[41] A. L. Read, Presentation of search results: The $CL_s$ technique, J. Phys. G 28, 2693 (2002).

[42] R. Aaij et al. (LHCb Collaboration), Measurement of the CKM angle $\gamma$ from a combination of LHCb results, J. High Energy Phys. 12 (2016) 087.

[43] M. Kenzie et al., GammaCombo: A statistical analysis framework for combining measurements, fitting datasets and producing confidence intervals, https://doi.org/10.5281/zenodo.3371421.

[44] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C 71, 1554 (2011); Erratum, Eur. Phys. J. C 73, 2501 (2013).
Z. Yang, 65 Y. Yao, 67 L. E. Yeomans, 59 H. Yin, 7 J. Yu, 7 X. Yuan, 67 O. Yushchenko, 43 K. A. Zarebski, 52 M. Zavertyaev, 15, v M. Zdybal, 33 M. Zeng, 3 D. Zhang, 6 L. Zhang, 6 Y. Zhang, 3, 6 W. C. Zhang, 43 Y. Zheng, 5 X. Zhou, 7 Y. Zhou, 5 X. Zhu, 3 V. Zhukov, 13, 39 J. B. Zonneveld, 57 and S. Zucchelli 19, c

(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 School of Physics State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
5 University of Chinese Academy of Sciences, Beijing, China
6 Institute Of High Energy Physics (IHEP), Beijing, China
7 Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China
8 Université Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IN2P3-LAPP, Annecy, France
9 Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France
10 Aix Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
11 Université Paris-Saclay, CNRS/IN2P3, IJCLab, Orsay, France
12 LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France
13 I. Physikalisches Institut, RWTH Aachen University, Aachen, Germany
14 Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
15 Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
16 Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
17 School of Physics, University College Dublin, Dublin, Ireland
18 INFN Sezione di Bari, Bari, Italy
19 INFN Sezione di Bologna, Bologna, Italy
20 INFN Sezione di Ferrara, Ferrara, Italy
21 INFN Sezione di Firenze, Firenze, Italy
22 INFN Laboratori Nazionali di Frascati, Frascati, Italy
23 INFN Sezione di Genova, Genova, Italy
24 INFN Sezione di Milano-Bicocca, Milano, Italy
25 INFN Sezione di Milano, Milano, Italy
26 INFN Sezione di Cagliari, Monsummaro, Italy
27 INFN Sezione di Padova, Padova, Italy
28 INFN Sezione di Pisa, Pisa, Italy
29 INFN Sezione di Roma Tor Vergata, Roma, Italy
30 INFN Sezione di Roma La Sapienza, Roma, Italy
31 Nikhef National Institute for Subatomic Physics, Amsterdam, Netherlands
32 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, Netherlands
33 Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
34 AGH—University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
35 National Center for Nuclear Research (NCBJ), Warsaw, Poland
36 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
37 Petersburg Nuclear Physics Institute NRC Kurchatov Institute (PNPI NRC KI), Gatchina, Russia
38 Institute of Theoretical and Experimental Physics NRC Kurchatov Institute (ITEP NRC KI), Moscow, Russia, Moscow, Russia
39 Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
40 Institute for Research of the Russian Academy of Sciences (INR RAS), Moscow, Russia
41 Yandex School of Data Analysis, Moscow, Russia
42 Budker Institute of Nuclear Physics (SB RAS), Novosibirsk, Russia
43 Institute for High Energy Physics NRC Kurchatov Institute (IHEP NRC KI), Protvino, Russia, Protvino, Russia
44 ICUB, Univeritat de Barcelona, Barcelona, Spain
45 Instituto Galego de Física de Altas Enxeñarias (IGFAE), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
46 Instituto de Física Corpuscular, Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain
47 European Organization for Nuclear Research (CERN), Geneva, Switzerland
48 Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
49 Physik-Institut, Universität Zürich, Zürich, Switzerland
50 NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
51 Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
52 University of Birmingham, Birmingham, United Kingdom
Also at Università di Siena, Siena, Italy.
8 Also at Università di Padova, Padova, Italy.
9 Also at Scuola Normale Superiore, Pisa, Italy.
10 Also at MSU - Iligan Institute of Technology (MSU-IIT), Iligan, Philippines.
11 Also at Hanoi University of Science, Hanoi, Vietnam.
12 Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
13 Also at Università di Roma Tor Vergata, Roma, Italy.
14 Also at Università della Basilicata, Potenza, Italy.
15 Also at Università di Urbino, Urbino, Italy.
16 Also at School of Physics and Information Technology, Shaanxi Normal University (SNNU), Xi’an, China.