Observation of the suppressed decay $\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-$

LHCb Collaboration; Bernet, R; Müller, K; Serra, N; Steinkamp, O; Straumann, U; Vollhardt, A; et al

Abstract: The suppressed decay $\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-$, excluding the $J/\psi$ and $\psi(2S) \rightarrow \mu^+\mu^-$ resonances, is observed for the first time with a significance of 5.5 standard deviations. The analysis is performed with proton-proton collision data corresponding to an integrated luminosity of $3\,fb^{-1}$ collected with the LHCb experiment. The $\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-$ branching fraction is measured relative to the $\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)p\pi^-$ branching fraction giving $\frac{B(\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-)}{B(\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)p\pi^-)} = 0.044 \pm 0.012 \pm 0.007$, where the first uncertainty is statistical and the second is systematic. This is the first observation of a $b \rightarrow d$ transition in a baryonic decay.

DOI: https://doi.org/10.1007/JHEP04(2017)029

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: https://doi.org/10.5167/uzh-146432

Journal Article

Published Version

The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:
LHCb Collaboration; Bernet, R; Müller, K; Serra, N; Steinkamp, O; Straumann, U; Vollhardt, A; et al (2017). Observation of the suppressed decay $\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-$. Journal of High Energy Physics, 04:029. DOI: https://doi.org/10.1007/JHEP04(2017)029
Observation of the suppressed decay $\Lambda_b^0 \rightarrow p\pi^- \mu^+\mu^-$

The LHCb collaboration

E-mail: smith@physik.rwth-aachen.de

ABSTRACT: The suppressed decay $\Lambda_b^0 \rightarrow p\pi^- \mu^+\mu^-$, excluding the $J/\psi$ and $\psi(2S) \rightarrow \mu^+\mu^-$ resonances, is observed for the first time with a significance of 5.5 standard deviations. The analysis is performed with proton-proton collision data corresponding to an integrated luminosity of 3 fb$^{-1}$ collected with the LHCb experiment. The $\Lambda_b^0 \rightarrow p\pi^- \mu^+\mu^-$ branching fraction is measured relative to the $\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)p\pi^-$ branching fraction giving

$$\frac{\mathcal{B}(\Lambda_b^0 \rightarrow p\pi^- \mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)p\pi^-)} = 0.044 \pm 0.012 \pm 0.007,$$

where the first uncertainty is statistical and the second is systematic. This is the first observation of a $b \rightarrow d$ transition in a baryonic decay.

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

ArXiv ePrint: 1701.08705
1 Introduction

The decay of the $\Lambda_b^0$ baryon into the $p\pi^-\mu^+\mu^-$ final state, where the muons do not originate from a hadronic resonance, is mediated by a $b \to d$ transition. Such decays are highly suppressed in the Standard Model (SM), as the leading order amplitudes are described by loop diagrams and are also suppressed by the relevant Cabibbo-Kobayshi-Maskawa (CKM) factors. This suppression is not necessarily present in extensions to the SM, and such decays are therefore sensitive to contributions from new particles. One of the lowest-order diagrams for the decay $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ is shown in figure 1.

The branching fraction of the decay\(^1\) $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ is expected to be of $\mathcal{O}(10^{-8})$. Together with the relevant form factors, a measurement of this branching fraction with respect to that of the analogous $b \to s$ transition $\Lambda_b^0 \to pK^-\mu^+\mu^-$, would allow the ratio of CKM elements $|V_{td}|/|V_{ts}|$ to be determined. Comparing the value of $|V_{td}|/|V_{ts}|$ from these processes with that measured via mixing processes would test the Minimal Flavour Violation hypothesis [1–3]. The decay, $\Lambda_b^0 \to pK^-\mu^+\mu^-$, has recently been observed for the first time by the LHCb collaboration [4].

At present, no form-factor calculations have been made for the $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ and $\Lambda_b^0 \to pK^-\mu^+\mu^-$ channels due to the complicated hadronic structure in the proton-meson systems. However, recent advances in lattice calculations [5] could make this possible in the future.

This paper describes a search for the decay $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$, using proton-proton collision data corresponding to an integrated luminosity of $3\text{ fb}^{-1}$. The data were collected with

---

\(^1\)The inclusion of charge-conjugate processes is implied throughout this paper.
the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The branching fraction is determined relative to that of the tree-level decay, $\Lambda_b^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) p \pi^-$, denoted as $\Lambda_b^0 \rightarrow J/\psi p \pi^-$ hereafter, which has been measured with a precision of 15% \cite{6, 7}.

2 Detector and simulation

The LHCb detector \cite{8, 9} is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/$c$. The minimum distance of a track to a primary vertex (PV), the impact parameter, is measured with a resolution of $(15 + 29/p_T)$ μm, where $p_T$ is the component of the momentum transverse to the beam, in GeV/$c$. Different types of charged hadrons are distinguished using information from two Ring-Imaging Cherenkov (RICH) detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers. The online event selection is performed by a trigger \cite{10}, 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.

Simulated events are used to optimise selection criteria and calculate the relative efficiency between the signal and normalisation channels. In the simulation, $pp$ collisions are generated using PYTHIA \cite{11, 12} with a specific LHCb configuration \cite{13}. Decays of hadronic particles are described by EVTGEN \cite{14}, in which final-state radiation is generated using PHOTOS \cite{15}. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit \cite{16, 17}, as described in ref. \cite{18}.
3 Selection

The $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ signal candidates are first required to pass the hardware trigger, which selects events containing at least one muon with $p_T$ greater than 1.48 GeV/c in the 7 TeV data or $p_T > 1.76$ GeV/c in the 8 TeV data. In the subsequent software trigger, at least one of the final-state particles is required to have $p_T > 1.7$ GeV/c in the 7 TeV data or $p_T > 1.6$ GeV/c in the 8 TeV data. For muon candidates, a softer requirement of $p_T > 1.0$ GeV/c is applied. The final-state particles that satisfy these transverse momentum criteria are also required to have an impact parameter larger than 100 μm with respect to all PVs in the event. Finally, the tracks of two or more of the final-state particles are required to form a vertex that is significantly displaced from all PVs.

Signal candidates are reconstructed by combining two oppositely-charged muons with two additional tracks that are identified as a proton and a pion using particle identification (PID) information that comes primarily from the RICH detectors. All final-state particles are required to have a good-quality track fit and to be inconsistent with originating from a PV. The pion (proton) candidates are required to have $p_T > 0.4$ GeV/c and momentum greater than 2.0 (7.5) GeV/c. The four final-state particles are required to form a good-quality vertex, where the resulting $\Lambda_b^0$ candidate is consistent with originating from a PV. The vertex is also required to be significantly displaced from this PV. In order to reject the background from $\Lambda_b^0 \to J/\psi p\pi^-$ and $\Lambda_b^0 \to \psi(2S)p\pi^-$ decays, the regions $8.0 < q^2 < 11.0$ GeV$^2$/c$^4$ and $12.5 < q^2 < 15.0$ GeV$^2$/c$^4$ are excluded from the signal search, where $q^2$ refers to the invariant mass squared of the two muons. In addition, contributions from $\Lambda_b^0 \to \Lambda^0(\to p\pi^-)\mu^+\mu^-$ decays are removed by requiring $m_{p\pi^-} > 1.12$ GeV/c$^2$.

Several fully reconstructed decays with at least one misidentified particle can form backgrounds that peak in the distribution of the $p\pi^-\mu^+\mu^-$ mass, $m_{p\pi^-\mu^+\mu^-}$. Specific vetoes are used to reject such backgrounds. The vetoes require that if the invariant mass of the candidate is consistent with a particular hypothesis, then a more restrictive PID requirement is applied. For example, if the proton candidate is assigned the kaon mass and falls within the mass range $5246 < m_{K^+\pi^-\mu^+\mu^-} < 5330$ MeV/c$^2$, the PID cut is significantly tightened to reduce $K \to p$ misidentification from $B^0 \to K^+\pi^-\mu^+\mu^-$ decays. Other possible sources of specific backgrounds are the decays $\Lambda_b^0 \to pK^-\mu^+\mu^-$, $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$, $B_s^0 \to K^+K^-\mu^+\mu^-$ and crossfeed from the normalisation channel $\Lambda_b^0 \to J/\psi p\pi^-$. After the vetoes have been applied, the only significant residual background contribution for the signal (normalisation) channel comes from the decay $\Lambda_b^0 \to pK^-\mu^+\mu^-$ ($\Lambda_b^0 \to J/\psi(\to \mu^+\mu^-)pK^-$). This contamination is treated as a systematic uncertainty in the signal channel and is considered explicitly when extracting the yield of the normalisation channel. Partially reconstructed $\Lambda_b^0 \to (\Lambda_b^0 \to p\pi^-\mu^+\nu)\mu^-\nu$ decays, which contribute at $m_{p\pi^-\mu^+\mu^-}$ below the nominal $\Lambda_b^0$ mass, are also explicitly considered when determining the signal yield.

A boosted decision tree (BDT) [19], with the AdaBoost algorithm [20] and a five-fold cross-validation method [21], is used to reduce combinatorial background. The BDT is trained and optimised on data. Candidates with $m_{p\pi^-\mu^+\mu^-} > 6000$ MeV/c$^2$ are used as a sample representative of the background, and $\Lambda_b^0 \to J/\psi(\to \mu^+\mu^-)pK^-$ candidates
selected from the data are used as a proxy for the signal sample. The BDT uses kinematic, geometric and PID variables associated with the proton to discriminate between the signal and background candidates. The two most discriminating input variables are the vertex quality of the $\Lambda_b^0$ candidate and its consistency with originating from a PV. In order to reject background containing additional tracks in close proximity to the $\Lambda_b^0$ vertex, an isolation parameter [22] is also used as an input variable. As the presence of a proton from a displaced vertex is a distinctive signature, PID information on the proton candidate is used in the BDT in order to improve the rejection of background. Other, less discriminating variables used in the BDT include the minimum impact parameter with respect to any PV and the momenta of the final-state particles. The requirement on the BDT response is optimised by maximising the figure of merit [23] defined as

$$P = \frac{\varepsilon_{sel}}{3/2 + \sqrt{B}},$$

where $\varepsilon_{sel}$ is the selection efficiency for the signal and $B$ is the background expected within 40 MeV/$c^2$ of the $\Lambda_b^0$ mass. After candidates have been reconstructed and the above selection criteria have been applied, the requirement on the BDT output retains 65% of signal events and rejects 99% of the background.

4 Normalisation

The branching fraction of $\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-$ can be determined from

$$B(\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-) = \frac{B(\Lambda_b^0 \rightarrow J/\psi p\pi^-)B(J/\psi \rightarrow \mu^+\mu^-)}{N(\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-)} \frac{\varepsilon(\Lambda_b^0 \rightarrow J/\psi p\pi^-)}{\varepsilon(\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-)},$$

(4.1)

where $N(X)$ is the yield of the final state $X$ and $\varepsilon(X)$ is the efficiency to select that final state. The efficiencies are obtained from simulated events and specific control samples in the data. Since the normalisation channel $\Lambda_b^0 \rightarrow J/\psi p\pi^-$ has the same final state and similar kinematics as the signal decay, many systematic uncertainties cancel in the efficiency ratio.

Control channels selected from the data are used to account for several effects that are mismodelled in the simulation. For example, the PID efficiencies are obtained from data samples with decays where the final-state particles can be identified by kinematic constraints alone [24]. Further corrections are derived by comparing the data and simulation distributions of the $\Lambda_b^0$ momentum, transverse momentum, decay time and the track multiplicity for the normalisation channel. The relative efficiency of the BDT is calculated using both $\Lambda_b^0 \rightarrow J/\psi pK^-$ and $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$ candidates selected from the data; the resulting efficiencies are consistent with each other. The most important difference in the efficiency between the signal and normalisation modes is due to the $q^2$ selection for the signal decay, which removes 30% of the signal candidates. For the full selection, including the dimuon mass vetoes, the total relative efficiency is found to be $\varepsilon(\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-)/\varepsilon(\Lambda_b^0 \rightarrow J/\psi p\pi^-) = 0.487 \pm 0.022$. 

- 4 -
For the normalisation channel, candidates are required to have a dimuon mass within 60 MeV/c² of the known J/ψ mass. The yield of the normalisation channel is obtained by performing an extended unbinned maximum likelihood fit to the Λ_b^0 → J/ψ pπ⁻ mass distribution, as shown in figure 2. The shape of the Λ_b^0 → J/ψ pπ⁻ mass distribution is described by the sum of two Gaussian functions with power law tails and a shared mean, where the Gaussian parameters are allowed to vary in the fit and the tail parameters are obtained from the simulation. Combinatorial background is parameterised with an exponential function with a decay constant that is allowed to vary in the fit. Finally, there is a small contribution from the decay Λ_b^0 → J/ψ pK⁺⁻, the shape of which is determined from the simulation and included in the fit to the data. In total, 1017 ± 41 Λ_b^0 → J/ψ pπ⁻ candidates are observed. This yield is significantly lower than in refs. [7, 25], owing to the tighter selection employed to search for the Λ_b^0 → pπ⁻μ⁺μ⁻ decay.

5 Results

The fit to the invariant mass distribution of Λ_b^0 → pπ⁻μ⁺μ⁻ candidates, excluding the J/ψ and ψ(2S) regions, is shown in figure 3. The signal shape is determined from the fit to the normalisation decay mode in data, with corrections for the differences between the signal and normalisation modes obtained from the simulation. The combinatorial background is parameterised as in the fit for the normalisation mode. The shape of the partially reconstructed background is obtained from a fit to the Λ_b^0 → pK⁻μ⁺μ⁻ mass spectrum and the yield is allowed to vary in the fit to the Λ_b^0 → pπ⁻μ⁺μ⁻ mass distribution.

A signal contribution is clearly visible and Wilks’ theorem [26] gives a significance of 5.5 standard deviations. The systematic uncertainties described in section 6 are mainly associated with the normalisation. Only the systematic uncertainty arising from the shape assumed for the partially reconstructed background has any appreciable impact on the significance. Releasing the constraints on the relevant parameters, the significance increases to 5.7 standard deviations. Pseudoexperiments indicate that, on-average, the significance would be expected to decrease by 0.3 standard deviations. Given the statistical variation, the observed increase is perfectly compatible with the expectation. This analysis therefore
Figure 3. Mass distribution of $\Lambda_b^0 \to p\pi^- \mu^+ \mu^-$ candidates compared to the result of the fit. The fit parameterisation is described in the text.

constitutes the first observation of the decay $\Lambda_b^0 \to p\pi^- \mu^+ \mu^-$. The number of signal candidates is found to be $22 \pm 6$, which is converted to relative and absolute branching fractions of

$$\frac{B(\Lambda_b^0 \to p\pi^- \mu^+ \mu^-)}{B(\Lambda_b^0 \to J/\psi (\to \mu^+ \mu^-) p\pi^-)} = 0.044 \pm 0.012 \pm 0.007$$

and

$$B(\Lambda_b^0 \to p\pi^- \mu^+ \mu^-) = (6.9 \pm 1.9 \pm 1.1^{+1.3}_{-1.0}) \times 10^{-8}$$

using eq. (4.1). In both cases, the first uncertainty given is statistical and the second is the systematic uncertainty, which is discussed in the next section. The third uncertainty on $B(\Lambda_b^0 \to p\pi^- \mu^+ \mu^-)$ arises from the limited knowledge of the $\Lambda_b^0 \to J/\psi p\pi^- [6, 7]$ and $J/\psi \to \mu^+ \mu^-$ [27] branching fractions.

6 Systematic uncertainties

The systematic uncertainties are summarised in table 1. The total systematic uncertainty is 16.1%, which is comparable to but smaller than the statistical uncertainty.

The largest systematic uncertainty originates from the decay model used to simulate the signal. There are two components to this uncertainty. The first originates from the unknown $q^2$ distribution for the signal decay. As no model for the $\Lambda_b^0 \to p\pi^- \mu^+ \mu^-$ decay currently exists, the model for the decay $\Lambda_b^0 \to \Lambda^0(\to p\pi^-) \mu^+ \mu^-$ from ref. [28] is used to derive the $q^2$ distribution. To assess the systematic uncertainty from this assumption, the decay $\Lambda_b^0 \to pK^- \mu^+ \mu^-$ is instead assumed to describe the signal $q^2$ distribution and the difference in relative efficiency is assigned as a systematic uncertainty. The $q^2$ distribution for the $\Lambda_b^0 \to pK^- \mu^+ \mu^-$ decay is obtained from data weighted using the sPlot technique [29]. An uncertainty of 7.9% is found. The second component of the systematic uncertainty due to the decay model is the distribution of the $p\pi^-$ invariant mass. In this case, the distribution in the simulation is weighted to match the data for the $\Lambda_b^0 \to J/\psi p\pi^-$ decay.
Table 1. Summary of the systematic uncertainties on the $\Lambda_{b}^{0} \rightarrow J/\psi p \pi^{-} \mu^{+}\mu^{-}$ branching fraction.

| Source                                      | Uncertainty (%) |
|---------------------------------------------|-----------------|
| Modelling of the $q^{2}$ distribution       | 7.9             |
| Modelling of the $p\pi^{-}$ mass distribution | 7.7             |
| Effect of the partially reconstructed background fit shape | 6.9             |
| Choice of BDT efficiency proxy              | 5.6             |
| Finite size of the simulated sample         | 4.4             |
| Statistical uncertainty on $\Lambda_{b}^{0} \rightarrow J/\psi p \pi^{-} \mu^{+}\mu^{-}$ yield | 4.0             |
| Trigger efficiency                          | 3.4             |
| Fit bias                                    | 2.2             |
| $\Lambda_{b}^{0} \rightarrow pK^{-}\mu^{+}\mu^{-}$ contamination | 1.6             |
| Simulation corrections                      | 1.3             |
| PID efficiency                              | 1.0             |
| Total                                       | 16.1            |

and the efficiency is reevaluated. The difference of 7.7% in relative efficiency between these two cases is taken as a systematic uncertainty.

Another important source of systematic uncertainty is related to the assumption that the partially reconstructed background for the signal has the same shape as the partially reconstructed background in $\Lambda_{b}^{0} \rightarrow pK^{-}\mu^{+}\mu^{-}$ decays. The effect of this assumption is estimated by allowing the shape parameters for the partially reconstructed background component to vary in the fit, and then calculating the resulting bias in the background estimation using pseudoexperiments. This results in a 6.9% uncertainty on the signal yield. As noted above, this is the only systematic uncertainty that has an appreciable effect on the significance for the observation of the decay $\Lambda_{b}^{0} \rightarrow p\pi^{-} \mu^{+}\mu^{-}$.

Other, smaller uncertainties are assigned to the calculation of the efficiency: the calibration of the BDT efficiency using data (5.6%); the finite size of the simulation samples used (4.4%) and possible mismodelling of the trigger (3.4%). The statistical uncertainty on the $\Lambda_{b}^{0} \rightarrow J/\psi p \pi^{-}$ yield gives rise to a systematic uncertainty of 4.0%. Due to the low number of signal candidates, a small bias in the signal yield is observed. The size of this bias is calculated using pseudoexperiments and results in a 2.2% systematic uncertainty. No $\Lambda_{b}^{0} \rightarrow pK^{-}\mu^{+}\mu^{-}$ contribution is considered for the $\Lambda_{b}^{0} \rightarrow p\pi^{-} \mu^{+}\mu^{-}$ fit, due to the low expected yield. The $\Lambda_{b}^{0} \rightarrow J/\psi pK^{-}$ decay is used to assess the resulting systematic uncertainty, which is 1.6%. The corrections applied to the simulation give rise to a small systematic uncertainty (1.3%), as does the calibration of the PID efficiency using data (1.0%).

7 Conclusions

A search for the rare decay $\Lambda_{b}^{0} \rightarrow p\pi^{-} \mu^{+}\mu^{-}$ has been performed with proton-proton collision data collected with the LHCb experiment corresponding to 3fb$^{-1}$ of integrated luminosity. The search is made excluding the $J/\psi$ and $\psi(2S) \rightarrow \mu^{+}\mu^{-}$ resonances. A signal is observed
with a significance of 5.5 standard deviations, which constitutes the first observation of a $b \to d$ transition in a baryonic decay. The relative and absolute branching fractions are measured to be

$$\frac{\mathcal{B}(\Lambda_b^0 \to p\pi^- \mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to J/\psi (\to \mu^+\mu^-)p\pi^-)} = 0.044 \pm 0.012 \pm 0.007$$

and

$$\mathcal{B}(\Lambda_b^0 \to p\pi^- \mu^+\mu^-) = (6.9 \pm 1.9 \pm 1.1_{-1.0}^{+1.3}) \times 10^{-8},$$

where the first uncertainties are statistical and the second are systematic. The third uncertainty on $\mathcal{B}(\Lambda_b^0 \to p\pi^- \mu^+\mu^-)$ arises from the limited knowledge of the $\Lambda_b^0 \to J/\psi p\pi^-$ [7] and $J/\psi \to \mu^+\mu^-$ [27] branching fractions. With further advances in lattice QCD combined with a $\Lambda_b^0 \to pK^- \mu^+\mu^-$ branching fraction measurement, this result will allow $|V_{td}|/|V_{ts}|$ to be measured, enabling a test of the Minimal Flavour Violation hypothesis.

Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.A.). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (U.S.A.). We are indebted to the communities behind the multiple open source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Sklodowska-Curie Actions and ERC (European Union), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

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

References

[1] A.J. Buras, P. Gambino, M. Gorbahn, S. Jager and L. Silvestrini, *Universal unitarity triangle and physics beyond the standard model*, Phys. Lett. B 500 (2001) 161 [hep-ph/0007085] [hep-ph/0007085].
[2] T. Feldmann and T. Mannel, Minimal flavour violation and beyond, JHEP 02 (2007) 067 [hep-ph/0611095] [INSPIRE].

[3] LHCb collaboration, First measurement of the differential branching fraction and CP asymmetry of the $B^\pm \to \pi^\pm \mu^+\mu^-$ decay, JHEP 10 (2015) 034 [arXiv:1509.00414] [INSPIRE].

[4] LHCb collaboration, Observation of the decay $B^0 \to p\bar{K}^-\mu^+\mu^-$ and a search for CP violation, arXiv:1703.00256 [INSPIRE].

[5] D.J. Wilson, R.A. Briceno, J.J. Dudek, R.G. Edwards and C.E. Thomas, Coupled $\pi\pi, K\bar{K}$ scattering in $P$-wave and the $\rho$ resonance from lattice QCD, Phys. Rev. D 92 (2015) 094502 [arXiv:1507.02599] [INSPIRE].

[6] LHCb collaboration, Study of the production of $A^0_b$ and $B^0$ hadrons in pp collisions and first measurement of the $A^0_b \to J/\psi p\bar{K}^-$ branching fraction, Chin. Phys. C 40 (2016) 011001 [arXiv:1509.00292] [INSPIRE].

[7] LHCb collaboration, Observation of the $\Lambda^0_b \to J/\psi p\pi^-$ decay, JHEP 07 (2014) 103 [arXiv:1406.0755] [INSPIRE].

[8] LHCb collaboration, The LHCb detector at the LHC, 2008 JINST 3 S08005 [INSPIRE].

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

[10] R. Aaij et al., The LHCb trigger and its performance in 2011, 2013 JINST 8 P04022 [arXiv:1211.3056] [INSPIRE].

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

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

[13] I. Belyaev et al., Handling of the generation of primary events in Gauss, the LHCb simulation framework, J. Phys. Conf. Ser. 331 (2011) 032047 [INSPIRE].

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

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

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

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

[18] M. Clemencic et al., The LHCb simulation application, Gauss: design, evolution and experience, J. Phys. Conf. Ser. 331 (2011) 032023 [INSPIRE].

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

[20] Y. Freund and R. E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, J. Comput. Syst. Sci. 55 (1997) 119.
[21] M. Stone, Cross-validatory choice and assessment of statistical predictions, J. R. Stat. Soc. B 36 (1974) 111.

[22] LHCb collaboration, Measurement of the ratio of branching fractions \( B(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/B(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu) \), Phys. Rev. Lett. 115 (2015) 111803 [arXiv:1506.08614] [inSPIRE].

[23] G. Punzi, Sensitivity of searches for new signals and its optimization, eConf C 030908 (2003) MODT002 [physics/0308063] [inSPIRE].

[24] LHCb RICH GROUP collaboration, M. Adinolfi et al., Performance of the LHCb RICH detector at the LHC, Eur. Phys. J. C 73 (2013) 2431 [arXiv:1211.6759] [inSPIRE].

[25] LHCb collaboration, Evidence for exotic hadron contributions to \( \Lambda^0_b \to J/\psi \pi^- \) decays, Phys. Rev. Lett. 117 (2016) 082003 [arXiv:1606.06999] [inSPIRE].

[26] S.S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, Ann. Math. Statist. 9 (1938) 60.

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

[28] W. Detmold, C.J.D. Lin, S. Meinel and M. Wingate, \( \Lambda_b \to \Lambda\ell^+\ell^- \) form factors and differential branching fraction from lattice QCD, Phys. Rev. D 87 (2013) 074502 [arXiv:1212.4827] [inSPIRE].

[29] M. Pivk and F.R. Le Diberder, SPlot: a statistical tool to unfold data distributions, Nucl. Instrum. Meth. A 555 (2005) 356 [physics/0402083] [inSPIRE].
University of Chinese Academy of Sciences, Beijing, China, associated to 3
School of Physics and Technology, Wuhan University, Wuhan, China, associated to 3
Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to 3
Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia, associated to 6
Institut für Physik, Universität Rostock, Rostock, Germany, associated to 12
National Research Centre Kurchatov Institute, Moscow, Russia, associated to 72
Instituto de Física Corpuscular, Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain, associated to 38
Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to 53

a Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
b Laboratoire Leprince-Ringuet, Palaiseau, France
c P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
d Università di Bari, Bari, Italy
e Università di Bologna, Bologna, Italy
f Università di Cagliari, Cagliari, Italy
g Università di Ferrara, Ferrara, Italy
h Università di Genova, Genova, Italy
i Università di Milano Bicocca, Milano, Italy
j Università di Roma Tor Vergata, Roma, Italy
k Università di Roma La Sapienza, Roma, Italy
l AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
m LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
n Hanoi University of Science, Hanoi, Viet Nam
o Università di Padova, Padova, Italy
p Università di Pisa, Pisa, Italy
q Università degli Studi di Milano, Milano, Italy
r Università di Urbino, Urbino, Italy
s Università della Basilicata, Potenza, Italy
t Scuola Normale Superiore, Pisa, Italy
u Università di Modena e Reggio Emilia, Modena, Italy
v Iligan Institute of Technology (IIT), Iligan, Philippines
w Novosibirsk State University, Novosibirsk, Russia 1 Deceased