



# Observation of the semileptonic decay $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$

LHCb collaboration<sup>†</sup>

## Abstract

The Cabibbo-suppressed semileptonic decay  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$  is observed for the first time. The differential branching fraction is measured as a function of the  $p\bar{p}$  invariant mass, using the mode  $B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+$  for normalisation. The total branching fraction is also measured to be

$$\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu) = (5.27_{-0.24}^{+0.23} \pm 0.21 \pm 0.15) \times 10^{-6},$$

where the first uncertainty is statistical, the second systematic and the third is from the uncertainty on the normalisation branching fraction.

For submission to JHEP

© 2019 CERN for the benefit of the LHCb collaboration. CC-BY-4.0 licence.

---

<sup>†</sup>Authors are listed at the end of this paper.



# 1 Introduction

Studies of semileptonic  $B$  meson decays have recently generated interest due to a number of anomalous results. Measurements of the observables  $R(D)$  and  $R(D^*)$  [1–6] have shown hints of lepton non-universality with a combined significance of over  $3\sigma$  [7]. To probe the flavour structure of any possible new physics behind these results, it is desirable to make analogous measurements for decays involving different quark transitions, for example  $b \rightarrow u$ . To that end, the mode  $B^+ \rightarrow p\bar{p}l^+\nu_l$  is promising experimentally, particularly when performing the measurement at a hadron collider, as the two protons in the final state should significantly reduce combinatorial backgrounds. Semileptonic decays of  $B$  mesons to a final state containing multiple baryons are as yet unobserved, although the Belle collaboration has found evidence for  $B^+ \rightarrow p\bar{p}e^+\nu_e$  with  $3.2\sigma$  significance [8].

A theoretical model of  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$  has been constructed based on perturbative QCD (pQCD) [9]. This paper is based on studies of several similar fully hadronic  $B \rightarrow YY'X$  decays where  $Y$  represents a baryon and  $X$  one or more mesons. By fitting the angular distributions and decay rates of the hadronic modes [10, 11] the authors of this paper estimate the total branching fraction and the differential decay rate of  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ . However, the predicted branching fraction is two orders of magnitude larger than the Belle measurement.

The fully hadronic modes themselves raise some questions that merit further investigation. It is surprising that the branching fraction of  $B^0 \rightarrow p\bar{p}$  is two orders of magnitude smaller than that of the almost equivalent decay  $B^0 \rightarrow p\bar{p}\pi^+\pi^-$  [12, 13]. Furthermore, the invariant mass distributions of the di-baryon pair in  $B \rightarrow YY'X$  decays show a characteristic shape that peaks at low values [14–16]. Understanding the dynamics that leads to such features is difficult in fully hadronic decays, due to the interaction of the two baryons and the extra hadrons. It is therefore desirable to study semileptonic decays, such as  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ , where such final state interactions are absent.

In this paper, the first observation of the decay  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$  is presented. As the dynamics of the transition are not known, the branching fraction is measured in bins of  $p\bar{p}$  invariant mass. These bins are then summed to obtain a measurement of the total branching fraction. The measurement is made relative to the normalisation mode  $B^+ \rightarrow J/\psi K^+$ , with  $J/\psi \rightarrow \mu^+\mu^-$ , to reduce the systematic uncertainties. The branching fraction within a bin  $i$  is

$$\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_i = \frac{N(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_i}{N(B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+)} \times \frac{\epsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+)}{\epsilon(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_i} \times \mathcal{B}(B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+),$$

where  $N(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)_i$  is the yield of  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$  candidates in bin  $i$ ,  $N(B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+)$  is the total yield of  $B^+ \rightarrow J/\psi K^+$  candidates and  $\epsilon$  represents the reconstruction and selection efficiencies of the two modes. The branching fractions of  $B^+ \rightarrow J/\psi K^+$  and  $J/\psi \rightarrow \mu^+\mu^-$  are external inputs.

To establish the signal yield a variable called the corrected mass is defined [17]

$$m_{corr} = |p'_T| + \sqrt{|p'_T|^2 + m_{vis}^2},$$

where  $|p'_T|$  is defined as the magnitude of the reconstructed  $p\bar{p}\mu^+$  candidate momentum transverse to the flight direction and  $m_{vis}^2$  is the square of the invariant mass of the candidate.

41 The data for this study is that collected by the LHCb detector in proton-proton  
42 collisions in 2011, 2012 and 2016. This corresponds to integrated luminosities of  $1.0\text{ fb}^{-1}$ ,  
43  $2.0\text{ fb}^{-1}$  and  $1.9\text{ fb}^{-1}$  at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV respectively.  
44 The 2011 and 2012 datasets are treated together and collectively referred to by the  
45 sobriquet Run 1. Charge conjugate processes are implied throughout this document.

## 46 2 Detector and simulation

47 The LHCb detector [18, 19] is a single-arm forward spectrometer covering the  
48 pseudorapidity range  $2 < \eta < 5$ , designed for the study of particles containing  $b$  or  
49  $c$  quarks. The detector includes a high-precision tracking system consisting of a silicon-  
50 strip vertex detector surrounding the  $pp$  interaction region [20], a large-area silicon-strip  
51 detector located upstream of a dipole magnet with a bending power of about 4 Tm, and  
52 three stations of silicon-strip detectors and straw drift tubes [21, 22] placed downstream  
53 of the magnet. The tracking system provides a measurement of the momentum,  $p$ , of  
54 charged particles with a relative uncertainty that varies from 0.5% at low momentum  
55 to 1.0% at 200 GeV/ $c$ . The minimum distance of a track to a primary vertex (PV), the  
56 impact parameter (IP), is measured with a resolution of  $(15 + 29/p_T)\text{ }\mu\text{m}$ , where  $p_T$  is  
57 the component of the momentum transverse to the beam, in GeV/ $c$ . Different types of  
58 charged hadrons are distinguished using information from two ring-imaging Cherenkov  
59 (RICH) detectors [23]. Photons, electrons and hadrons are identified by a calorimeter  
60 system consisting of scintillating-pad and preshower detectors, an electromagnetic and a  
61 hadronic calorimeter. Muons are identified by a system composed of alternating layers of  
62 iron and multiwire proportional chambers [24].

63 The online event selection is performed by a trigger [25], which consists of a hardware  
64 stage that performs some basic selection, followed by a software stage, which applies a  
65 full event reconstruction. At the first level, a track consistent with being a muon with  
66 significant  $p_T$  is required to be present in the event. Subsequently in the software stage,  
67 two tracks are required to form a secondary vertex with significant displacement from a  
68  $pp$  interaction vertex. A multivariate algorithm [26] is used to identify vertices that are  
69 consistent with the decay of a  $b$  hadron.

70 Simulation is used to determine the efficiency of the signal mode and estimate the shapes  
71 of the signal and several background modes in the fit. In the simulation,  $pp$  collisions are  
72 generated using PYTHIA [27] with a specific LHCb configuration [28]. Decays of unstable  
73 particles are described by EVTGEN [29], in which final-state radiation is generated using  
74 PHOTOS [30]. The interaction of the generated particles with the detector, and its response,  
75 are implemented using the GEANT4 toolkit [31], as described in Ref. [32]. The generated  
76  $B$  meson  $p$  and  $p_T$  spectra are corrected to better match the data. A boosted decision  
77 tree (BDT) reweighter [33] is trained on samples of  $B^+ \rightarrow J/\psi K^+$  data and simulation,  
78 independent of those used for the normalisation of the branching fraction. This is then  
79 used to correct all of the simulation used in the analysis.

## 80 3 Selection

81 Signal candidates are constructed from three tracks which are each required to be of  
82 good quality and have a large IP with respect to any PV. The tracks must also have

83 particle identification criteria consistent with their particle hypothesis. The requirement  
84 for positive proton identification in turn enforces a minimum  $p$  requirement on the protons  
85 of 15 GeV such that they are above the threshold for radiating in the RICH. Similarly,  
86 the muons must have  $p$  above 3 GeV to propagate through the muon stations. All tracks  
87 must also have large  $p_T$ . Finally, the three candidate tracks must form a good quality  
88 vertex that is significantly displaced from the PV with which the total momentum vector  
89 of the candidate has the smallest IP. The muon of the candidate must be that which fired  
90 the hardware trigger and then the reconstructed candidate itself must be consistent with  
91 the object that fired the software trigger. Potential decays of  $\eta_c$ ,  $J/\psi$  and  $\psi(2S)$  to  $p\bar{p}$   
92 are vetoed with cuts on the  $p\bar{p}$  invariant mass of  $\pm 50$  MeV around the respective nominal  
93 masses.

94 The selection of the  $B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+$  normalisation mode is aligned with  
95 that of the signal to reduce systematic uncertainties. The cuts on the signal protons are  
96 applied to the  $K^+$  and the muon of opposite sign ( $\mu^-$ ), with the exception of the particle  
97 identification criteria. The remaining muon has the signal muon selection applied to it.

98 Further selection is used to reduce several sources of backgrounds relative to the  
99 signal. The largest contribution comes from a melange of partially reconstructed decays  
100 producing two protons and a muon in the final state. It is expected that the largest among  
101 these would originate in  $b \rightarrow c$  transitions. The most pernicious is  $B \rightarrow \bar{A}_c^- p \mu^+ \nu_\mu X$ ,  
102 where  $X$  represents any number of charged or neutral pions (including none), where the  
103  $\bar{A}_c^-$  decays to a final state including one proton. The other major background arises  
104 from  $B \rightarrow p\bar{p}DX$ , where the  $D$  meson may be of any variety ( $D^0$ ,  $D^+$ ,  $D^{*+}$  etc.) but  
105 ultimately decays to a final state with a muon. A similar decay of  $B \rightarrow p\bar{A}_c^- X$  with the  $\bar{A}_c^-$   
106 decaying semileptonically is comparatively small, as the semileptonic branching fraction is  
107 dominated by  $\bar{A}_c^- \rightarrow \Lambda l^- \bar{\nu}_l$  decays and the  $\Lambda$  flies a sufficient distance within the detector  
108 before decaying such that the resulting proton is not associated with the  $B$  decay vertex.  
109 Another source of partially reconstructed background is formed of  $B \rightarrow p\bar{p}\mu^+\nu_\mu X$  decays,  
110 where  $X$  denotes one or more charged or neutral pions. These decays may proceed with  
111 intermediate  $N^*$  or  $\Delta$  resonances and could naively be expected to have similar branching  
112 fractions to the signal.

113 If any of these partially reconstructed decay modes produces a surplus charged track  
114 they can be efficiently dealt with by an isolation technique. Once a signal candidate has  
115 been constructed, the other tracks in the event that are close to the  $B$  decay-vertex are  
116 then examined and a BDT is used to give a probability that these nearby tracks can be  
117 associated with the signal candidate decay-vertex. If the candidate is truly signal, there  
118 should be few other tracks that can be associated with it and the BDT should classify  
119 them with a low probability. On the other hand, the extra track(s) from a partially  
120 reconstructed decay will give a high probability of association if such tracks are found.  
121 The algorithm returns the BDT probabilities for the four tracks most likely to have come  
122 from the  $B$  vertex. These four probabilities are themselves combined into a single BDT  
123 classifier, known as the charged isolation BDT. This BDT is trained on simulation to  
124 discriminate signal from  $B^+ \rightarrow \bar{A}_c^- p \mu^+ \nu_\mu$ , which is expected to be the largest decay mode  
125 with extra charged tracks. The efficacy of this BDT in reducing such background is shown  
126 in Fig. 1.

127 For those partially reconstructed final states with only extraneous neutral particles,  
128 additional suppression is achieved by considering the kinematics of the decays. A further  
129 BDT, the so called part-reco BDT, considers 11 variables: the impact parameter  $\chi^2$ ,

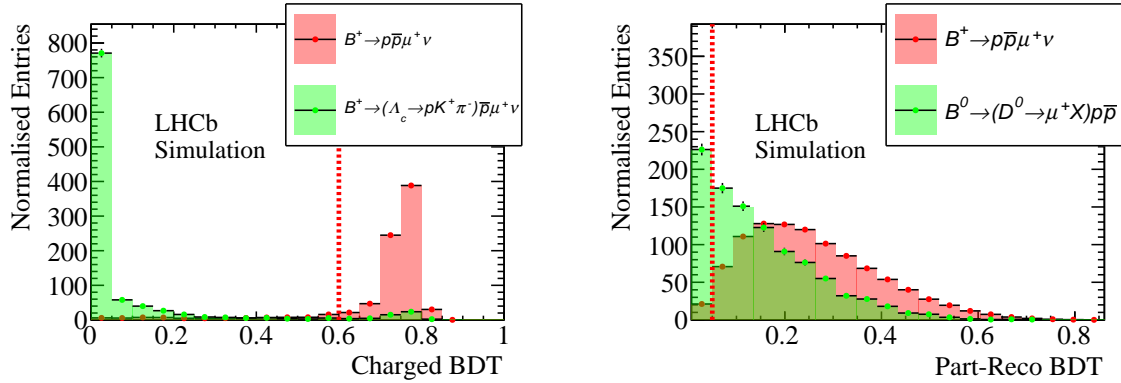


Figure 1: The result of training the (left) charged isolation BDT and the (right) part-reco BDT on the right. The chosen cut values are indicated by the red line. For some events there are no additional tracks near the  $B$ -decay vertex; these events are accepted and do not appear in the charged BDT output.

130 defined as the change in  $\chi^2$  of the vertex fit when removing the candidate, of the three  
 131 daughter tracks, the  $p\bar{p}$  pair and the  $B^+$  candidate with respect to the PV; the impact  
 132 parameters of the tracks with respect to the fitted  $B^+$  decay vertex; the  $\chi^2$  of the  $B^+$   
 133 vertex fit; the angle between the  $B^+$  candidate momentum and flight distance vectors; and  
 134 the difference between the  $p$  and  $\bar{p}$  momenta. The part-reco BDT is trained on simulation  
 135 in order to discriminate signal from a mixture of all the considered background modes.  
 136 The result of this training is shown in Fig. 1

137 An additional background arises from particles that are mis-identified as protons (mis-  
 138 ID). The particle identification requirements on the proton tracks are therefore further  
 139 tightened relative to those from earlier. The background due to the muon being another  
 140 mis-identified particle has been considered but is reduced to a negligible amount with  
 141 only a loose particle identification cut.

142 In addition to the two BDTs and proton identification criteria, one further quantity is  
 143 considered: the uncertainty on the corrected mass of the candidate. This is calculated from  
 144 the estimated uncertainties on the positions of the  $B^+$  primary and secondary vertices,  
 145 and the momenta of the tracks. Selecting lower values of the corrected-mass uncertainty  
 146 produces a sharper peak for the signal mode in the corrected mass variable, which will  
 147 aid the discrimination of the signal from background in the fit to determine the yield.  
 148 Therefore in total the selection uses four quantities (two BDTs, the proton PID and the  
 149 corrected-mass error). In order to ascertain the optimal cut point, a four dimensional grid  
 150 search is performed using pseudo-experiments. Datasets are generated from the simulation  
 151 samples with the expected proportions of each background. The expected signal amount is  
 152 taken from the central value of the  $B^+ \rightarrow p\bar{p}e^+\nu_e$  branching fraction reported by Belle [8].  
 153 For the backgrounds, the current averages for the branching fractions are used if they  
 154 have been measured. For those backgrounds that have not been measured their branching  
 155 fractions are estimated relative to that expected for the signal, accounting for different  
 156 CKM factors and the available phase-space. For each trialed set of cuts the expected  
 157 relative uncertainty on the signal yield is found by a fit to the simulated data. These  
 158 fits are not binned in  $m(p\bar{p})$  but take the entire statistics together. The set of cuts that  
 159 produces the smallest relative uncertainty on the signal yield is chosen.

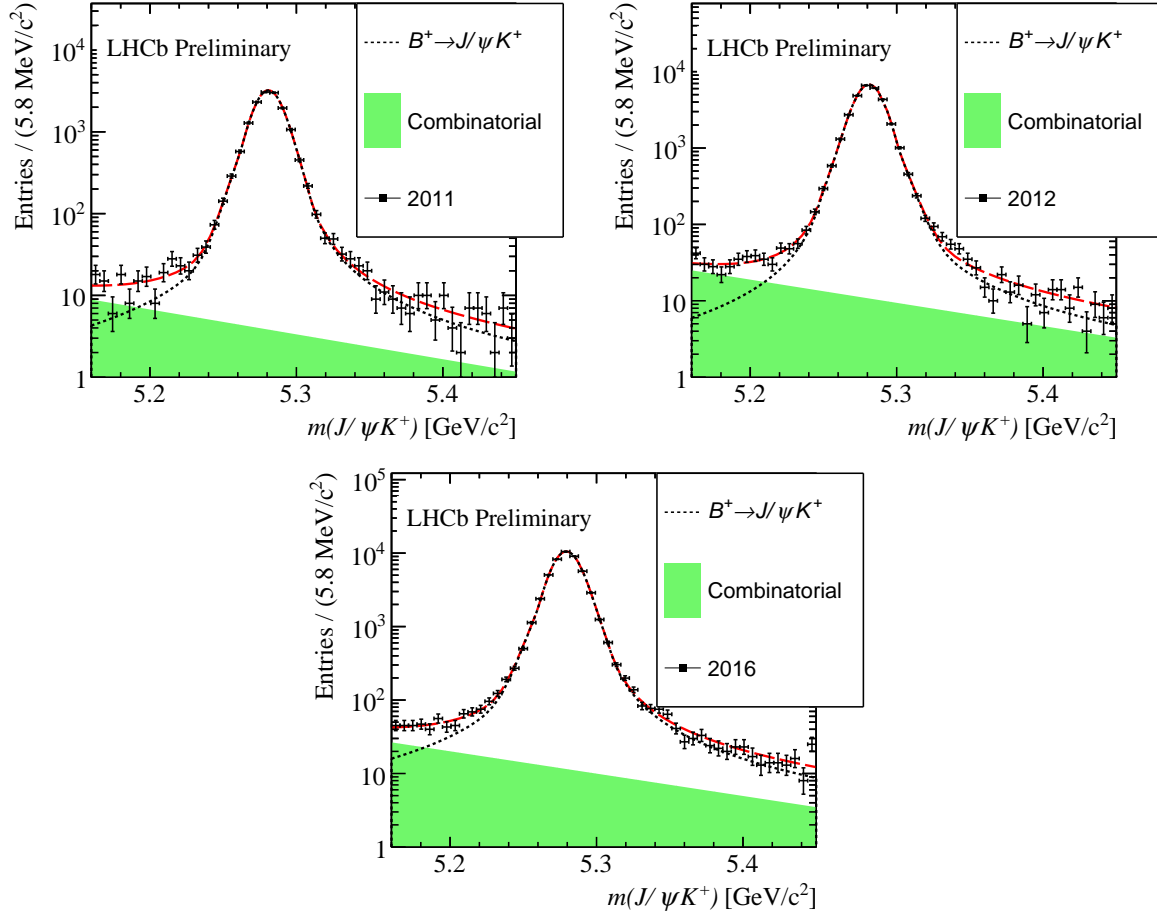


Figure 2: The  $m(J/\psi K^+)$  distribution with the fit result shown for the (top left) 2011 data, (top right) 2012 and (bottom) 2016.

## 4 Signal and normalisation yields

160

161 The yields of signal and normalisation are ascertained with unbinned, extended, maximum-  
 162 likelihood fits. In the case of the normalisation mode, the invariant mass distribution of  
 163 the  $J/\psi K^+$  candidates is fitted. The 2011, 2012 and 2016 datasets are fitted separately  
 164 and then the yields combined. The fits are shown in Fig. 2.

165 For the signal mode, the corrected mass is fitted. The distribution of this variable  
 166 peaks at the true  $B^+$  rest mass for candidates where one massless particle has not been  
 167 reconstructed. On the other hand, candidates from partially reconstructed decays that are  
 168 missing one or more massive particles in addition to the neutrino have wide distributions  
 169 concentrated at lower values of corrected mass. To help the fits converge, the run 1 and  
 170 2016 data are combined and fitted together.

171 For the signal component and contributions from partially-reconstructed decays the  
 172 shapes are fixed from the simulation. The shape of the proton mis-ID comes from a  
 173 dedicated data sample in which the particle identification requirements on one of the  
 174 protons have been removed. Using this data sample the expected number of the proton  
 175 mis-ID candidates can be estimated and a Gaussian constraint is applied to the yield in  
 176 the fit. A final background due to random combinations of protons and muons, referred

177 to as the combinatorial background, is included in the fit. A sample of data for which the  
 178  $B^+$  decay vertex quality selection has been reversed is used to estimate the shape of this  
 179 background. The form of the signal probability density function (PDF) is parameterised by  
 180 the sum of four bifurcated Gaussian functions with a shared mean. All of the background  
 181 PDFs are accounted for with kernel density estimation [34].

182 The yields of the signal, proton mis-ID, combinatorial and total partially-reconstructed  
 183 decays are determined by the fit, as are the relative fractions of each partially reconstructed  
 184 mode. All of the fit parameters are free with the exception of the Gaussian constrained  
 185 proton mis-ID yield.

186 The fit in each  $m(p\bar{p})$  bin is performed independently. The  $m_{corr}$  distributions in  
 187 each bin and the resulting fits are shown in Fig. 3. In each bin the fits are validated  
 188 with pseudo-experiments. An ensemble of  $10^5$  datasets are generated and fitted with the  
 189 component yields taken from the fits to data. Some small biases on the signal yield are  
 190 found and these are considered as a source of systematic uncertainty.

## 191 5 Efficiency

192 The efficiencies for the signal and normalisation modes to be reconstructed and selected  
 193 are both assessed with simulation. Further corrections are applied to account for known  
 194 differences between data and simulation in the track reconstruction efficiency [35] and  
 195 the efficiency of the hardware trigger [36]. The efficiency of the particle identification  
 196 requirements on each track is evaluated with a data driven method [37] and applied to  
 197 the simulation.

198 The binning in  $m(p\bar{p})$  reduces the dependence on the model of the  $B^+$  decay when  
 199 calculating the efficiency of the signal mode. However, as the selection cuts on kinematic  
 200 quantities of the candidates there is still some residual dependence on the physics of the  
 201 decay. The simulation is therefore re-weighted to represent the pQCD model of Ref. [9]  
 202 as the current best estimate of how the decay proceeds. This weighting corrects the  
 203 distributions of  $m(p\bar{p})$  and the invariant mass of the di-lepton system ( $q^2$ ). The variation  
 204 of the parameters of this model is considered as a source of systematic uncertainty.

205 The ratio of selection efficiencies between the signal and normalisation in each bin of  
 206  $m(p\bar{p})$  is shown in Table 1. These efficiencies are presented separately for the Run 1 and  
 207 2016. They are combined to form an overall efficiency ratio, accounting for the difference  
 208 in sample sizes between Run 1 and 2016.

Table 1: Relative efficiencies for the run 1 and 2016 and the weighted combination of both.

$\frac{\epsilon(B \rightarrow p\bar{p}\mu\nu)_{m(p\bar{p})}}{\epsilon(B \rightarrow J/\psi K)}$	Relative efficiency in bins of $m(p\bar{p})$ [ GeV ]				
	1.87 – 2.0	2.0 – 2.2	2.2 – 2.4	2.4 – 2.6	2.6 – 5.0
Run 1	0.37	0.37	0.36	0.36	0.35
2016	0.56	0.51	0.50	0.52	0.49
Run 1&2016	0.48	0.45	0.44	0.45	0.43



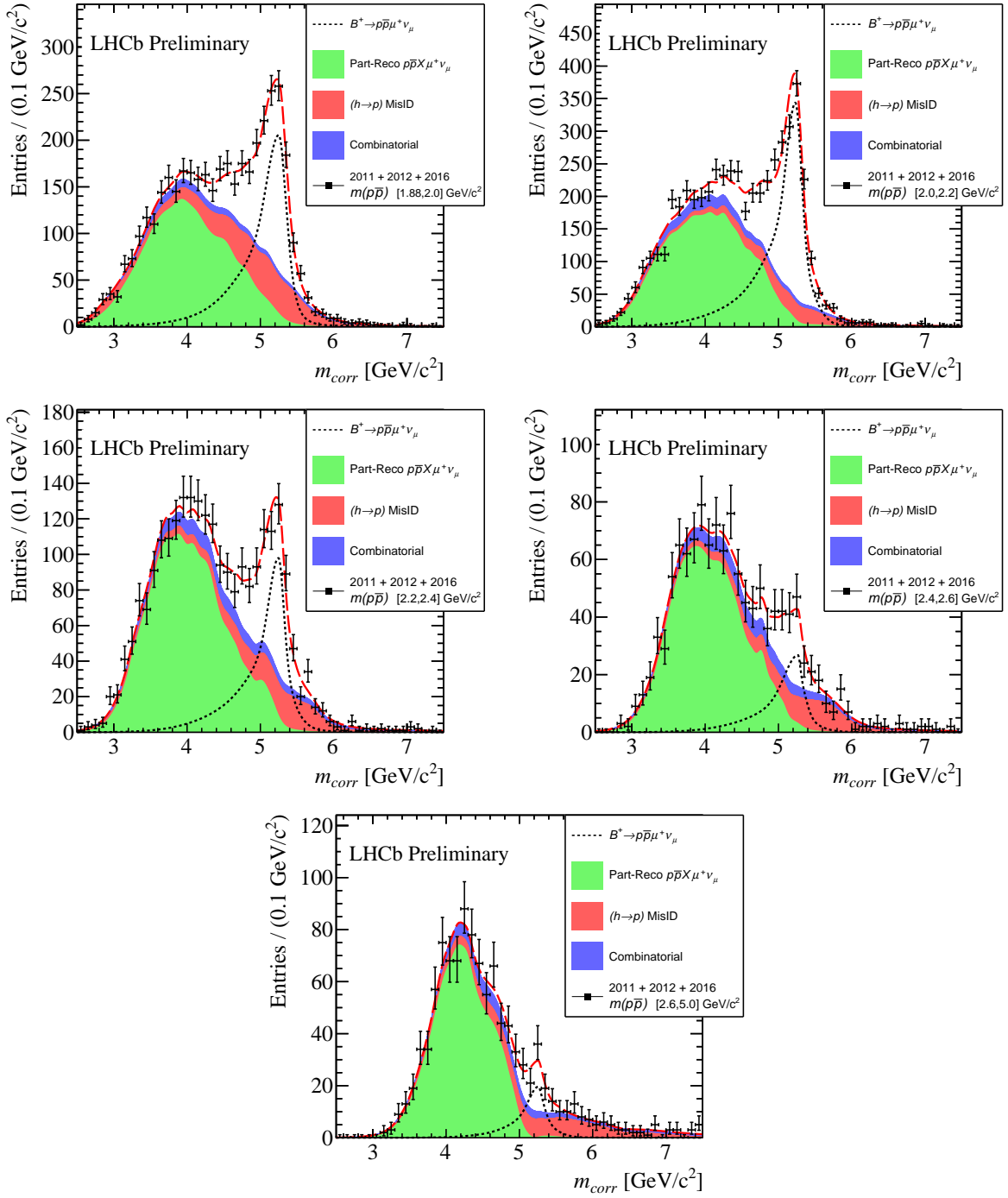


Figure 3: The distributions of  $m_{corr}$  in each  $m(p\bar{p})$  bin with the fit results shown.

## 6 Systematic uncertainties

The systematic uncertainties can be split into two categories: those that affect the calculation of the ratio of efficiencies of the signal and normalisation modes, and those that may change the determination of the signal yield in the fit. For the former, each of the corrections to the simulation contributes a source of uncertainty both from the limited sizes of the samples used to derive the corrections and from the method of deriving them. The uncertainty due to the limited simulation sample sizes used to calculate the efficiencies is also included. A further uncertainty is due to the physics model that the simulation is weighted to represent. The model affects the kinematic distributions of the daughter tracks which feeds into the efficiency calculation as these distributions are cut on. Due to the unproven nature of the model, a conservative uncertainty is taken. New sets of weights for the simulation are created that sample extreme variations of the model parameters ( $\pm 5\sigma$ ) and for each variation the efficiency is recalculated. Despite this extreme test the systematic due to the physics model is not dominant, which reflects the flat selection efficiency over the kinematic ranges in which the daughter tracks lie within each bin of  $m(p\bar{p})$ .

In the corrected mass fit uncertainties arise from potential variations in the shapes of the components. This variation is assessed with pseudo-experiments. Datasets are generated with the nominal fit model and then fitted with the nominal model and an alternative. The width of the distribution of differences between the nominal and alternative fits gives the uncertainty. For those components that rely on kernel density estimators, a systematic is assessed for the choice of smoothing parameter by varying it. The uncertainty due to the choice of model for the signal shape is found by replacing the nominal PDF with one constructed with kernel density estimators. The uncertainty due to the limited statistics of the simulation samples is determined by generating new simulation from the nominal fit PDFs with the same statistics and making alternative PDFs with those samples. For the shape of the combinatorial background component an alternative data sample is trialled which requires the two protons to be of the same charge.

A summary of the systematic uncertainties is presented in Table 2. They are given as relative uncertainties on the branching fraction with the combination accounting for the correlation of the uncertainties between the two datasets.

## 7 Results

The fitted yields for the signal mode are presented in Table 3. The extracted yields of the normalisation channel are  $(1.49 \pm 0.03) \times 10^4$  for 2011,  $(3.14 \pm 0.02) \times 10^4$  for 2012 and  $(4.927 \pm 0.025) \times 10^4$  for 2016. Combining these with the efficiency ratios from Section 5, the differential branching fraction in each  $m(p\bar{p})$  bin is calculated, relative to the branching fraction of the normalisation mode. The results are presented in the third column of Table 3. The relative differential branching fractions are then summed over the bins, with the correlation of the systematic uncertainties between the bins accounted for, to give the total relative branching fraction of

$$\frac{\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)}{\mathcal{B}(B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+)} = (8.75 \pm 0.39 \pm 0.35) \times 10^{-2},$$

Table 2: A summary of the systematic uncertainties on the differential branching fractions. The particle identification and tracking efficiency systematics are assumed to be 100% correlated between run 1 and run 2.

Source	Range of relative uncertainties on BF [%]
Kinematic reweighting	0.39 – 0.56
Simulation statistics	2.92 – 3.48
Tracking efficiency	2.74– 2.74
Physics model	0.01 – 0.05
Particle identification	0.68 – 1.67
Data-simulation agreement	0.40
Run 1 and run 2 combination	1.63 – 2.06
Kernel smoothing	0.02 – 7.90
Signal model	0.59 – 9.87
MisID model	0.13 – 13.46
Simulation statistics	0.02 – 5.16
Combinatorial model	0.90 – 8.51
Fit bias	0.13 – 7.80
Total systematic uncertainty	5.20 – 20.37
Total statistical uncertainty	5.54 – 29.64

249 where the first uncertainty is statistical and the second systematic. Multiplying this by  
 250 the current average of the normalisation branching fraction leads to

$$\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu) = (5.27_{-0.24}^{+0.23} \pm 0.21 \pm 0.15) \times 10^{-6},$$

251 where the third uncertainty is from the normalisation branching fraction. Finally, the  
 252 absolute differential branching fraction as a function of  $m(p\bar{p})$  is shown in Fig. 4, where  
 253 the indicated uncertainties include statistical, systematic and normalisation uncertainty  
 254 contributions. As expected from the theory model and the analogous hadronic decays,  
 255 the differential distribution peaks at a very low value and falls off sharply. The measured  
 256 total branching fraction agrees with the previous measurement from Belle and represents  
 257 the first observation of the  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$  decay mode.

Table 3: Table of the number of observed  $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$  candidates in each bin of  $p\bar{p}$  mass with the associated statistical uncertainty.

$p\bar{p}$ Mass range [GeV/c <sup>2</sup> ]	Signal Yield	$d\mathcal{B}(B^+ \rightarrow p\bar{p}\mu^+\nu_\mu)/dm(p\bar{p}) [\times 10^{-6}]$
1.87 – 2.0	1208 <sup>+111</sup> <sub>-110</sub>	12.86 <sup>+1.18</sup> <sub>-1.17</sub> $\pm$ 0.68 $\pm$ 0.38
2.0 – 2.2	1826 <sup>+97</sup> <sub>-105</sub>	12.88 <sup>+0.69</sup> <sub>-0.74</sub> $\pm$ 0.67 $\pm$ 0.38
2.2 – 2.4	526 <sup>+67</sup> <sub>-64</sub>	3.78 <sup>+0.48</sup> <sub>-0.46</sub> $\pm$ 0.24 $\pm$ 0.11
2.4 – 2.6	148 <sup>+40</sup> <sub>-35</sub>	1.038 <sup>+0.279</sup> <sub>-0.248</sub> $\pm$ 0.162 $\pm$ 0.030
2.6 – 5.0	88 <sup>+26</sup> <sub>-26</sub>	0.0542 <sup>+0.0163</sup> <sub>-0.0160</sub> $\pm$ 0.0113 $\pm$ 0.0016

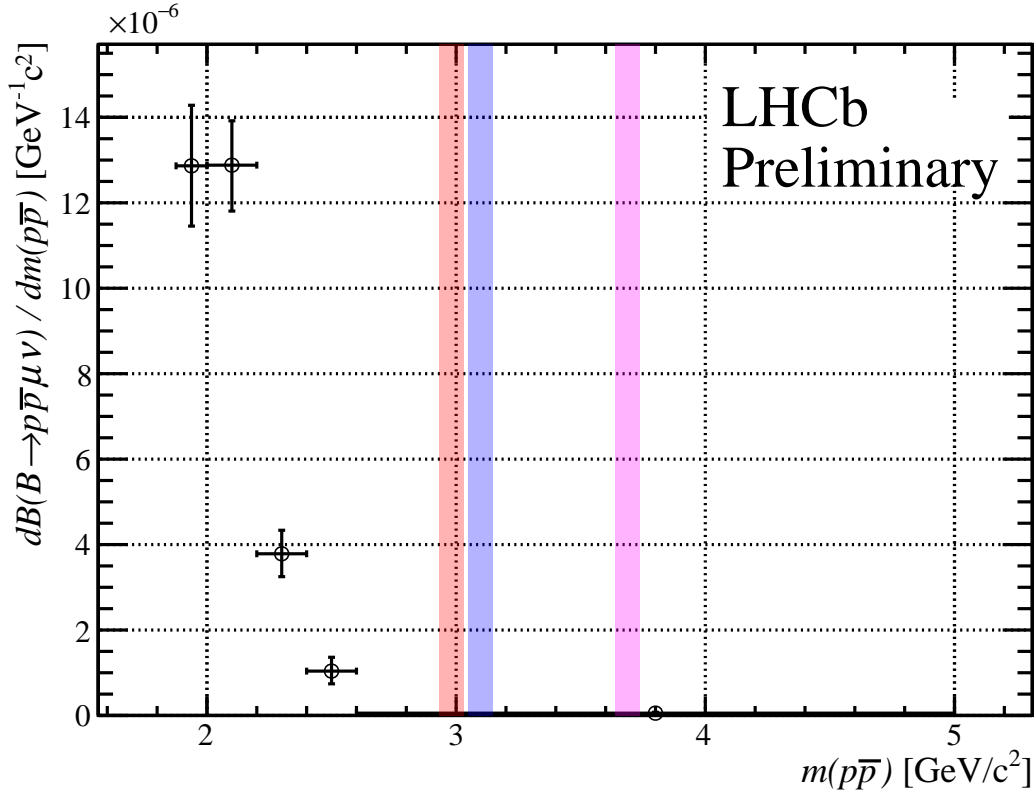


Figure 4: Differential branching fraction as a function of  $p\bar{p}$  mass. The (left) red band indicates the  $\eta_c \rightarrow p\bar{p}$  veto and (middle) blue band indicates the  $J/\psi \rightarrow p\bar{p}$  veto. The  $\psi(2S)$  veto is the (right) pink band.

## Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at 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); 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).

277 **8 Supplementary material for LHCb-PAPER-20XX-**  
278 **YYY**

279 This appendix contains supplementary material that will posted on the public CDS record  
280 but will not appear in the paper.

## References

- 281
- 282 [1] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the  $\mathcal{B}(B^0 \rightarrow D^{*-} \tau^+ \nu_\tau)$*   
283 *and  $\mathcal{B}(B^0 \rightarrow D^{*-} \mu^+ \nu_\mu)$  branching fractions using three-prong  $\tau$ -lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, [arXiv:1708.08856](#).  
284
- 285 [2] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching frac-*  
286 *tions  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)$* , Phys. Rev. Lett. **115** (2015) 111803,  
287 *Publisher's Note ibid.* **115** (2015) 159901, [arXiv:1506.08614](#).
- 288 [3] Belle, A. Abdesselam *et al.*, *Measurement of  $\mathcal{R}(D)$  and  $\mathcal{R}(D^*)$  with a semileptonic*  
289 *tagging method*, [arXiv:1904.08794](#).
- 290 [4] Belle, S. Hirose *et al.*, *Measurement of the  $\tau$  lepton polarization and  $R(D^*)$  in the*  
291 *decay  $\bar{B} \rightarrow D^* \tau^- \bar{\nu}_\tau$* , Phys. Rev. Lett. **118** (2017) 211801, [arXiv:1612.00529](#).
- 292 [5] Belle, M. Huschle *et al.*, *Measurement of the branching ratio of  $\bar{B} \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$*   
293 *relative to  $\bar{B} \rightarrow D^{(*)} \ell^- \bar{\nu}_\ell$  decays with hadronic tagging at Belle*, Phys. Rev. **D92**  
294 (2015) 072014, [arXiv:1507.03233](#).
- 295 [6] BaBar, J. P. Lees *et al.*, *Evidence for an excess of  $\bar{B} \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$  decays*, Phys. Rev.  
296 Lett. **109** (2012) 101802, [arXiv:1205.5442](#).
- 297 [7] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of  $b$ -hadron,  $c$ -*  
298 *hadron, and  $\tau$ -lepton properties as of summer 2016*, Eur. Phys. J. **C77**  
299 (2017) 895, [arXiv:1612.07233](#), updated results and plots available at  
300 <https://hflav.web.cern.ch>.
- 301 [8] Belle, K.-J. Tien *et al.*, *Evidence for semileptonic  $B^- \rightarrow p \bar{p} \ell^- \bar{\nu}_\ell$  decays*, Phys. Rev.  
302 **D89** (2014) 011101, [arXiv:1306.3353](#).
- 303 [9] C. Q. Geng and Y. K. Hsiao, *Semileptonic  $B^- \rightarrow p \bar{p} \ell^- \bar{\nu}_\ell$  decays*, Phys. Lett. **B704**  
304 (2011) 495, [arXiv:1107.0801](#).
- 305 [10] C. Q. Geng and Y. K. Hsiao, *Angular distributions in three-body baryonic  $b$  decays*,  
306 Phys. Rev. D **74** (2006) 094023.
- 307 [11] C.-H. Chen, H.-Y. Cheng, C. Q. Geng, and Y. K. Hsiao, *Charmful three-body baryonic*  
308  *$b$  decays*, Phys. Rev. D **78** (2008) 054016.
- 309 [12] LHCb collaboration, R. Aaij *et al.*, *First observation of the rare purely baryonic decay*  
310  *$B^0 \rightarrow p \bar{p}$* , Phys. Rev. Lett. **119** (2017) 232001, [arXiv:1709.01156](#).
- 311 [13] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev.  
312 **D98** (2018) 030001.
- 313 [14] BaBar, B. Aubert *et al.*, *Evidence for the  $B^0 \rightarrow p \bar{p} K^{*0}$  and  $B^+ \rightarrow p \bar{p} \eta(c)$*   
314  *$K^{*+}$  decays and Study of the Decay Dynamics of  $B$  Meson Decays into  $p$  anti- $p$   $h$*   
315 *final states*, Phys. Rev. **D76** (2007) 092004, [arXiv:0707.1648](#).

- 316 [15] BaBar, B. Aubert *et al.*, *Measurements of the Decays  $B^0 \rightarrow \bar{D}^0 p \bar{p}$ ,  $B^0 \rightarrow \bar{D}^* 0 p \bar{p}$ ,  $B^0 \rightarrow D^- p \bar{p} \pi^+$ , and  $B^0 \rightarrow D^{*-} p \bar{p} \pi^+$* , Phys. Rev. **D74** (2006) 051101,  
317 arXiv:hep-ex/0607039.  
318
- 319 [16] Belle, M. Z. Wang *et al.*, *Observation of  $B^+ \rightarrow p \text{ anti-}p \pi^+$ ,  $B^0 \rightarrow p \text{ anti-}p K^0$ , and  
320  $B^+ \rightarrow p \text{ anti-}p K^{*+}$* , Phys. Rev. Lett. **92** (2004) 131801, arXiv:hep-ex/0310018.
- 321 [17] SLD, K. Abe *et al.*, *A Measurement of  $R(b)$  using a vertex mass tag*, Phys. Rev. Lett.  
322 **80** (1998) 660, arXiv:hep-ex/9708015.
- 323 [18] LHCb, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008)  
324 S08005.
- 325 [19] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015)  
326 1530022, arXiv:1412.6352.
- 327 [20] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007,  
328 arXiv:1405.7808.
- 329 [21] LHCb Outer Tracker Group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*,  
330 JINST **9** (2014) P01002, arXiv:1311.3893.
- 331 [22] LHCb Outer Tracker Group, P. d'Argent *et al.*, *Improved performance of the LHCb  
332 Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, arXiv:1708.00819.
- 333 [23] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at  
334 the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.
- 335 [24] A. A. Alves, Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013)  
336 P02022, arXiv:1211.1346.
- 337 [25] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013)  
338 P04022, arXiv:1211.3055.
- 339 [26] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using  
340 a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.
- 341 [27] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP  
342 **05** (2006) 026, arXiv:hep-ph/0603175; T. Sjöstrand, S. Mrenna, and P. Skands,  
343 *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852,  
344 arXiv:0710.3820.
- 345 [28] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb  
346 simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- 347 [29] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.  
348 **A462** (2001) 152.
- 349 [30] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections  
350 in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.

- 351 [31] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE  
352 Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4:*  
353 *A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- 354 [32] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and*  
355 *experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- 356 [33] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762**  
357 (2016) , [arXiv:1608.05806](https://arxiv.org/abs/1608.05806), [https://github.com/arogozhnikov/hep\\_ml](https://github.com/arogozhnikov/hep_ml).
- 358 [34] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun.  
359 **136** (2001) 198, [arXiv:hep-ex/0011057](https://arxiv.org/abs/hep-ex/0011057).
- 360 [35] T. L. collaboration, *Measurement of the track reconstruction efficiency at LHCb*,  
361 Journal of Instrumentation **10** (2015) P02007.
- 362 [36] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency*  
363 *determination at LHCb*, LHCb-PUB-2014-039, 2014.
- 364 [37] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.



## LHCb collaboration

365 A. N. Other<sup>1</sup>.

366 <sup>1</sup>*University of nowhere*