B flavour tagging using charm decays at the LHCb experiment

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 JINST 10 P10005

(http://iopscience.iop.org/1748-0221/10/10/P10005)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 130.209.115.106
This content was downloaded on 15/08/2016 at 14:10

Please note that terms and conditions apply.
**B** flavour tagging using charm decays at the LHCb experiment

The LHCb collaboration

*E-mail: jwimberl@cern.ch*

**ABSTRACT:** An algorithm is described for tagging the flavour content at production of neutral *B* mesons in the LHCb experiment. The algorithm exploits the correlation of the flavour of a *B* meson with the charge of a reconstructed secondary charm hadron from the decay of the other *b* hadron produced in the proton-proton collision. Charm hadron candidates are identified in a number of fully or partially reconstructed Cabibbo-favoured decay modes. The algorithm is calibrated on the self-tagged decay modes $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow J/\psi K^{*0}$ using 3.0 fb$^{-1}$ of data collected by the LHCb experiment at *pp* centre-of-mass energies of 7 TeV and 8 TeV. Its tagging power on these samples of $B \rightarrow J/\psi X$ decays is $(0.30 \pm 0.01 \pm 0.01)$%.

**KEYWORDS:** Performance of High Energy Physics Detectors; Analysis and statistical methods

**ArXiv ePrint:** 1507.07892
1 Introduction

Measurements that involve mixing and time-dependent $CP$ asymmetries in decays of neutral $B$ mesons require the identification of their flavour content at production. This is achieved via various flavour tagging algorithms that exploit information from the rest of the $pp$ collision event. Same-side (SS) taggers look for particles produced in association with the signal $B$ meson during the hadronization of the $b$ quark [1]. The $d$ or $s$ partner of the light valence quark of the signal $B$ has a roughly 50% chance of hadronizing into a charged pion or kaon. Since $b$ quarks are mostly produced in $b\bar{b}$ pairs, the flavour content of the signal $B$ meson can also be deduced from available information on the opposite-side (OS) $b$ hadron, whose flavour is the opposite of the signal $B$ meson at the production time. OS muon and electron taggers look for leptons originating from semileptonic $b\to cW$ transitions of the $b$ hadron, and an OS kaon tagger looks for kaons coming from $b\to c\to s$ transitions. A vertex-charge tagger reconstructs the decay vertex of the OS $b$ hadron and predicts its charge by weighting the charges of its decay products according to their transverse momentum. The OS taggers employed by LHCb are described in ref. [2] and the SS taggers in refs. [3, 4]. This paper reports a new flavour tagging algorithm for the LHCb experiment that relies on reconstructed decays of charm hadrons produced in the OS $b$ hadron decay. For the development and evaluation of the tagging algorithm, signal $B$ meson and charm hadron candidates are reconstructed using data from 3 fb$^{-1}$ of integrated luminosity collected by LHCb at 7 TeV and 8 TeV centre-of-mass energies in 2011 and 2012, respectively.

The performance of a flavour tagging algorithm is defined by its tagging efficiency, $\varepsilon_{\text{tag}}$, mistag fraction, $\omega$, and dilution, $\mathcal{D} = 1 - 2\omega$. For a simple tagging algorithm with discrete decisions -- $B^0$, $\bar{B}^0$, or untagged -- these metrics are directly related to the numbers of rightly tagged ($R$), wrongly tagged ($W$), and untagged events ($U$) in a signal sample:

$$\varepsilon_{\text{tag}} = \frac{R + W}{R + W + U}, \quad \omega = \frac{W}{R + W}, \quad \mathcal{D} = \frac{R - W}{R + W}. \quad (1.1)$$
The performance of the flavour tagging algorithms is improved by assigning confidence weights to their tagging decisions. For each tagger, a multivariate classifier is trained using simulated data to distinguish between correct and incorrect decisions [2]. The inputs to the classifier are a selection of kinematic and geometric quantities describing the tagging track(s), the signal $B$ meson, and the event. This classifier then calculates a predicted mistag probability $\eta$ for each decision made. The predicted mistag probability is calibrated to data using an appropriate flavour self-tagged mode, such as $B^+ \rightarrow J/\psi K^+$, or a mode involving neutral $B$ oscillation, which self-tags its flavour at the decay-time, such as $B^0 \rightarrow J/\psi K^{*0}$ or $B^0 \rightarrow D^+ \pi^+$ [4, 5] (the use of charge-conjugate modes is implied throughout this paper). This calibration procedure provides a function $\omega(\eta)$, which relates the actual mistag probability $\omega$ to the predicted mistag probability $\eta$. Weighting each signal candidate by $1 - 2\omega(\eta)$ leads to an improved effective mistag fraction $\omega$ and associated dilution $D = 1 - 2\omega$. The statistical power of a $CP$ asymmetry measurement using a tagging algorithm is proportional to the effective tagging efficiency (or tagging power) $\varepsilon_{\text{eff}}$, defined as

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2.$$  \hspace{1cm} (1.2)

The typical combined tagging power of the current set of OS tagging algorithms used by LHCb is approximately 2.5% [3, 6–8]. Any augmentation to this tagging power increases the statistical precision achievable in $CP$ measurements at LHCb.

2 Detector and simulation

The LHCb detector [9, 10] is a single-arm forward spectrometer covering the pseudorapidity range between 2 and 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 [11], 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 [12] 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 $(15 + 29/p_T) \mu$m, where $p_T$ is the component of the momentum transverse to the beam, in GeV/$c$. Different types of charged hadrons in the momentum range 2–100 GeV/$c$ are distinguished using information from two ring-imaging Cherenkov detectors [13]. 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 [14]. 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.

In the simulation, $pp$ collisions are generated using PYTHIA [16, 17] with a specific LHCb configuration [18]. Decays of hadronic particles are described by EvtGen [19], in which final-state radiation is generated using PHOTOS [20]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [21] as described in ref. [22].
3 Tagging potential of OS charm hadrons

In events containing a signal $B$ decay, the opposite-side $D^+, D^0,$ and $\Lambda_c^+$ charm hadrons are primarily produced through the quark-level $b \to c$ transition. The charge of the $D^+$ or $\Lambda_c^+$ determines the flavour of the $b$ hadron parent. For $D^0$ decays through the dominant Cabibbo-favoured process $D^0 \to K^- X$, the kaon charge determines the flavour of the charm hadron, and thereby that of the parent $B$ hadron (the effect of $D^0$ mixing is negligible). The OS charm tagging algorithm uses charm hadron candidates reconstructed in a number of decay modes, chosen for their relatively large branching fractions, listed in Table 1. These include fully reconstructed hadronic modes with a single charged kaon in the final state, partially reconstructed hadronic modes with an unobserved neutral pion, and partially reconstructed semileptonic modes. Table 1 also reports the breakdown of the charm tagger’s performance by decay mode. The relative rate and relative power of each mode are the amounts that it contributes to the algorithm’s total tagging rate $\varepsilon_{\text{tag}}$ and tagging power $\varepsilon_{\text{eff}}$, which are presented in section 6 and Table 3. The algorithm predicts the flavour of the signal $B$ meson using the charge of the kaon in the same manner as the OS kaon tagger; however, the selection based on the reconstruction of $c$ hadrons (rather than the selection of kaons based on their individual kinematic properties) results in a different set of selected kaons and provides a complementary source of tagging information.

Several effects contribute an irreducible component to the mistag probability for the OS charm tagging algorithm. The dominant impact comes from $B^0 \to \bar{B}^0$ oscillation and from the contributions of “wrong sign” charm hadrons produced in $b \to c\bar{q}g$ transitions. The impact of Cabibbo-suppressed $D^0 \to K^+ X$ decays is negligible, as these typically produce additional kaons and do not mimic modes used by the tagging algorithm, and doubly Cabibbo-suppressed decays such as $D^0 \to K^+ \pi^-$ have a negligibly small branching fraction. Accounting for relative production rates of $b$ hadrons, neutral $B$ oscillation, and branching fractions of the decay modes used in the tagger, the irreducible mistag probabilities for $D^0$, $D^+$ and $\Lambda_c^+$ modes are estimated to be 23%, 19%, and 6%, respectively.

In addition to the irreducible mistag probability arising from physics effects, the charm hadron candidates are contaminated by combinatorial and partially reconstructed $b$ and $c$ hadron background that can lead to an incorrect flavour tag result. For each mode, the charm tagger uses a multivariate algorithm that combines geometric and kinematic quantities and properties of the $c$ hadron candidate and its daughters. The resulting discriminating variable is used both to suppress the combinatorial background and to predict the corresponding mistag probability for the surviving candidate.

4 Selection of charm candidates

Charm decay candidates are formed by combining kaon, pion, and proton candidates that satisfy particle identification criteria. These particles are required to have momentum $p > 1000\text{MeV}/c$, transverse momentum with respect to the beam axis $p_T > 100\text{MeV}/c$, and to be significantly displaced from any PV. For the candidates in the partially reconstructed modes and the decay $D^0 \to K^- \pi^+ \pi^+ \pi^-$, which contain large combinatorial backgrounds, more stringent requirements are imposed on the displacement of the final-state particles from the PV. In addition, particles are required to have $p_T > 150\text{MeV}/c$ for candidates in the mode $D^0 \to K^- \pi^+ \pi^+ \pi^-$. 

Table 1. Decay modes used in the OS charm tagger. The symbol $H_c$ stands for any $c$ hadron. The definition of the two right-most columns is given in the text.

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>Relative rate</th>
<th>Relative power</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^0 \rightarrow K^- \pi^+$</td>
<td>10.0%</td>
<td>24.0%</td>
</tr>
<tr>
<td>$D^0 \rightarrow K^- \pi^+ \pi^-$</td>
<td>5.9%</td>
<td>8.4%</td>
</tr>
<tr>
<td>$D^+ \rightarrow K^- \pi^+ \pi^+$</td>
<td>10.3%</td>
<td>2.6%</td>
</tr>
<tr>
<td>$H_c \rightarrow K^- \pi^+ X$</td>
<td>69.7%</td>
<td>61.5%</td>
</tr>
<tr>
<td>$H_c \rightarrow K^- e^+ X$</td>
<td>0.5%</td>
<td>0.2%</td>
</tr>
<tr>
<td>$H_c \rightarrow K^- \mu^+ X$</td>
<td>3.4%</td>
<td>0.3%</td>
</tr>
<tr>
<td>$\Lambda_c^+ \rightarrow p^+ K^- \pi^+$</td>
<td>0.2%</td>
<td>2.4%</td>
</tr>
</tbody>
</table>

Charm hadron candidates are required to pass a number of selection requirements. These include a maximum distance of closest approach between each pair of daughter tracks and a minimum quality of the decay vertex fit. Each candidate is required to be well separated from any PV and to have a trajectory that leads back to the best PV, chosen to be the PV for which the impact parameter significance of the charm hadron is smallest. The invariant mass of the charm hadron candidate is required to be consistent with the known mass of the corresponding charm hadron, within $100\text{MeV}/c^2$ for the $\Lambda_c^+$ channel and $50\text{MeV}/c^2$ for all other fully reconstructed $D$ decay modes. For the partially reconstructed $D \rightarrow K^- \pi^+ X$ modes, the $K^- \pi^+$ mass is required to be in a $[-400\text{MeV}/c^2, +0\text{MeV}/c^2]$ window around the known $D^0$ mass or in a window of $\pm 50\text{MeV}/c^2$ around the $K^*(892)^0$ resonance. The former is favoured by the invariant mass distribution of $K^- \pi^+$ pairs from the quasi-two body decay $D^0 \rightarrow K^- \rho^+$, and the latter selects $D \rightarrow K^*(892)^0 X$ decays. Charm candidates surviving these criteria still contain significant background contamination, which must be further reduced in order to lower the mistag probability of the algorithm.

For each mode, an adaptive-boosted decision tree (BDT) [23, 24] is used both to suppress background candidates and to estimate mistag probabilities. The inputs to the BDT are variables describing the decay kinematics, decay vertex and displacement, and particle identification information on the decay products. A variable related to the decay-time is calculated from the distance between the $c$ hadron’s decay vertex and the corresponding best PV; this approximates the sum of the decay-times of the $c$ hadron and its parent $b$ hadron. The BDT algorithms are trained using Monte Carlo (MC) simulations of $b\bar{b}$ events containing $B^+ \rightarrow J/\psi K^+, B^0 \rightarrow J/\psi K^{*0}$, and $B_s^0 \rightarrow J/\psi \phi$ decays on the signal side and inclusive decays of the $b$ hadron on the opposite-side. These $B$ decays are used to model the various sources and amounts of background when reconstructing OS $c$ hadrons recoiling against signal $B$ decays.

The output of the BDT, along with the simulation record of candidate identification, is used to compute the predicted mistag probability $\eta$ for each $c$ hadron candidate. Candidates with $\eta < 45\%$ are used in the flavour tagging decision. Removing candidates that fail this criterion significantly reduces the computing time of the algorithm at little cost to tagging performance. In cases where multiple charm candidates are present, the candidate with the lowest predicted mistag probability is retained. The combined efficiency of these requirements for retaining tagged events is $(59.00 \pm 0.07)\%$ and $(53.4 \pm 0.3)\%$ in simulation and data, respectively.
5 Calibration

While simulated data are used to develop and optimize the charm tagging algorithm, its performance is calibrated with collision data by comparing the algorithm’s predictions to the known flavours of signal \( B \) candidates, according to the procedure detailed in ref. [2]. The calibration parameters \( p_0, p_1, \Delta p_0, \) and \( \Delta p_1 \) are defined by

\[
\omega = p_0 + p_1 \left( \langle \eta \rangle - \langle \eta \rangle \right) \\
\Delta \omega = \Delta p_0 + \Delta p_1 \left( \langle \eta \rangle - \langle \eta \rangle \right)
\]

where \( \langle \eta \rangle \) is the average predicted mistag probability, \( \omega \) is the actual mistag probability averaged over \( B^+ \) and \( B^- \) signal mesons, and \( \Delta \omega \) is the excess mistag probability for \( B^+ \) mesons with respect to \( B^- \) mesons; equivalent definitions hold for \( B^0/\bar{B}^0 \) signal. In the ideal case, the offset parameter \( p_0 \) should equal \( \langle \eta \rangle \), and so the related parameter \( \delta p_0 = p_0 - \langle \eta \rangle \) is often more convenient.

A calibration of the algorithm has been performed using the flavour self-tagged mode \( B^+ \rightarrow J/\psi K^+ \). The signal candidates are selected by combining pairs of oppositely charged muons, with invariant mass consistent with the known \( J/\psi \) mass, with charged kaons, and are required to pass a set of cuts to obtain a good signal to background ratio [2]. When multiple candidates are present for a single event, that with the best decay vertex fit is kept. A fit to the reconstructed \( B^+ \) mass distribution is used to separate signal and background via the \( sPlot \) procedure, which computes signal and background weights for each candidate [25]. The empirical model for the signal is a sum of two Crystal Ball functions [26], while background is modeled by an exponential distribution. A total of \( 1.1 \times 10^6 \) signal candidates in this channel are found in the full dataset. The parameters \( p_0 \) and \( p_1 \) are determined by splitting the data into 13 bins of \( \eta \) between 0.19 and 0.45, calculating \( \omega_i \) and \( \bar{\eta}_i \) (the average \( \eta \) in each bin, and performing a linear fit to the set of values \( \omega_i, \bar{\eta}_i \)). The calibration parameters \( \Delta p_0 \) and \( \Delta p_1 \) are obtained from fits to the \( B^+ \) and \( B^- \) data each split into 5 bins of \( \eta \). The quantities \( \Delta \omega_i \) and \( \bar{\eta}_i \) are calculated in each of the 5 bins, and a linear fit is performed to the set of values \( \langle \Delta \omega_i, \bar{\eta}_i \rangle \). These fits are shown in Figs. 1 and 2. The resulting calibration parameters are given in table 2.

A cross-check of the calibration has been carried out using a \( B^0 \rightarrow J/\psi K^{*0} \) control sample. For this calibration, \( B^0/\bar{B}^0 \) oscillation must be taken into account. The Hypatia function [27] is used to model the signal’s mass distribution, while the background is modeled with a sum of two exponential functions. A set of simultaneous fits to the \( B^0 \) lifetime distribution in bins of \( \eta \) is performed, in which \( p_0, p_1, \Delta p_0, \) and \( \Delta p_1 \) are parameters of the fit model. In each bin, the raw \( B^0/\bar{B}^0 \) mixing asymmetry is defined as

\[
A_{\text{mixing}} = \frac{N(D = P) - N(D \neq P)}{N(D = P) + N(D \neq P)}, \tag{5.1}
\]

where \( D \) is the \( B \) meson flavour at decay-time and \( P \) is the production flavour predicted by the charm tagger. The amplitude of this asymmetry is governed by the actual mistag fraction \( \omega_i \) in the bin, while the bin’s average predicted mistag probability is \( \bar{\eta}_i \). The fit attempts to match the calibrated value \( \omega_i(\bar{\eta}_i) \) to \( \omega_i \) in each bin by adjusting the calibration parameters. A projection of the fitted model to the mixing asymmetry is shown in Fig. 3. The values of the calibration parameters obtained from the fit are given in table 2. The parameters are compatible with those obtained in the \( B^+ \rightarrow J/\psi K^+ \) mode, with the total \( \chi^2 \) per degree of freedom equal to 0.65.
Figure 1. Mistag probability $\omega$ as a function of the predicted mistag probability $\eta$ for the $B^+ \rightarrow J/\psi K^+$ data sample. A straight line fit to extract the parameters $p_0$ and $p_1$ is superimposed. The dark (green) and light (yellow) bands are the regions within 1σ and 2σ of the fitted value, respectively.

Figure 2. Excess mistag probability $\Delta \omega$ as a function of the predicted mistag probability $\eta$ for the $B^+ \rightarrow J/\psi K^+$ data sample. A straight line fit to extract the parameters $\Delta p_0$ and $\Delta p_1$ is superimposed. The dark (green) and light (yellow) bands are the regions within 1σ and 2σ of the fitted value, respectively.
Figure 3. Raw $B^0 - B^0$ mixing asymmetry (defined in eq. 5.1) vs. decay-time for the $B^0 \to J/\psi K^{*0}$ data sample. The amplitude of the asymmetry is diluted due to mistagging by the charm tagger. The mixing asymmetry from the fit is superimposed.

Table 2. Calibration parameters as determined from the $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ control samples. For both calibration modes, the average predicted mistag probability $\langle \eta \rangle$ is 0.379. The first uncertainties are statistical and the second are systematic. The systematic uncertainties are evaluated using simulation.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\delta p_0 (10^{-3})$</th>
<th>$p_1$</th>
<th>$\Delta p_0 (10^{-3})$</th>
<th>$\Delta p_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \to J/\psi K^+$</td>
<td>$-25 \pm 3 \pm 3$</td>
<td>$1.00 \pm 0.06 \pm 0.02$</td>
<td>$15 \pm 5 \pm 4$</td>
<td>$-0.08 \pm 0.12 \pm 0.04$</td>
</tr>
<tr>
<td>$B^0 \to J/\psi K^{*0}$</td>
<td>$-18 \pm 8 \pm 3$</td>
<td>$1.16 \pm 0.17 \pm 0.02$</td>
<td>$23 \pm 11 \pm 4$</td>
<td>$0.21 \pm 0.25 \pm 0.04$</td>
</tr>
</tbody>
</table>

The relatively small yield of the decay $B^0_s \to D_s^- \pi^+$ precludes performing a data-driven calibration on a $B^0_s$ mode. Therefore, in order to ensure that the algorithm performs similarly for $B^0_s$ channels as well as $B^+$ and $B^0$ channels, separate calibrations to simulated $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{*0}$, and $B^0_s \to J/\psi \phi$ events are performed. Where statistically significant differences between the calibration parameters in the three channels are found, a systematic uncertainty, corresponding to half of the maximum difference, is assigned to the parameter. These systematic uncertainties are roughly the size of the statistical uncertainties for the parameters $p_0$ and $\Delta p_0$, but are negligible for $p_1$ and $\Delta p_1$. The propagation of these uncertainties results in a 0.011% absolute systematic uncertainty on the tagging power, comparable to its statistical uncertainty.

Other sources of systematic uncertainty on calibration parameters have been investigated and found to have negligible effect. These include the potential effect of the chosen model of the invariant $B$ mass distribution for the channel $B^+ \to J/\psi K^+$. Two alternative models of the mass distribution were used and gave nearly identical results.

There are additional systematic uncertainties related to flavour tagging that must be considered in a $CP$ asymmetry analysis. These include differences between the signal channel sample and calibration channel sample in phase space distribution, event multiplicity, number of primary vertices,
Figure 4. Distribution of the calibrated predicted mistag probability \( \omega(\eta) \) for the \( B^{+} \to J/\psi K^+ \) data sample.

Table 3. Tagging efficiencies (\( \epsilon_{\text{tag}} \)), effective mistag fractions (\( \omega \)), and tagging powers (\( \epsilon_{\text{eff}} \)) in the various data samples studied. The first uncertainties are statistical and the second are systematic. The sample labeled Simulation is the training sample of simulated \( B^{+} \to J/\psi K^+ \), \( B^0 \to J/\psi K^0 \), and \( B^0_s \to J/\psi \phi \) decays, which has negligible statistical uncertainties.

<table>
<thead>
<tr>
<th>Sample</th>
<th>( \epsilon_{\text{tag}} )</th>
<th>( \omega )</th>
<th>( \epsilon_{\text{eff}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation</td>
<td>4.88%</td>
<td>37.0%</td>
<td>0.33%</td>
</tr>
<tr>
<td>( B^{+} \to J/\psi K^+ )</td>
<td>(3.11 ( \pm ) 0.02)%</td>
<td>(34.6 ( \pm ) 0.3 ( \pm ) 0.3)%</td>
<td>(0.30 ( \pm ) 0.01 ( \pm ) 0.01)%</td>
</tr>
<tr>
<td>( B^0 \to J/\psi K^0 )</td>
<td>(3.32 ( \pm ) 0.04)%</td>
<td>(35.0 ( \pm ) 0.8 ( \pm ) 0.3)%</td>
<td>(0.30 ( \pm ) 0.03 ( \pm ) 0.01)%</td>
</tr>
<tr>
<td>( B^0 \to D^- \pi^+ )</td>
<td>(4.11 ( \pm ) 0.03)%</td>
<td>(34.4 ( \pm ) 0.4 ( \pm ) 0.3)%</td>
<td>(0.40 ( \pm ) 0.02 ( \pm ) 0.01)%</td>
</tr>
<tr>
<td>( B^0_s \to D_s^- \pi^+ )</td>
<td>(3.99 ( \pm ) 0.07)%</td>
<td>(34.4 ( \pm ) 0.6 ( \pm ) 0.3)%</td>
<td>(0.39 ( \pm ) 0.03 ( \pm ) 0.01)%</td>
</tr>
</tbody>
</table>

or other variables. These differences would require corrections and would introduce tagging-related systematic uncertainties. Such effects are dependent on the signal channel and selection, and must be determined separately for each analysis.

6 Performance

The distribution of \( \eta \) after calibration for the \( B^{+} \to J/\psi K^+ \) control sample is shown in Fig. 4. The tagging efficiency, mistag fraction, and the tagging power of the charm tagger are reported in table 3 for the training sample of simulated \( B \to J/\psi X \) decays and for both calibration channels. The propagated statistical uncertainty of the calibration parameters dominates the statistical uncertainty of the tagging power. The overall tagging power is slightly higher in simulation than in data, due to differences in the distributions of input variables. The tagging powers in the two \( B \to J/\psi X \) calibration channels are consistent.

Table 3 also reports the tagging metrics for the decays \( B^0 \to D^- \pi^+ \) and \( B^0_s \to D^- \pi^+ \). Fits to the mass distributions of the signal candidates are performed to separate signal from background.
In each fit, the signal is modeled by a sum of two Crystal Ball functions and the combinatorial background is described by an exponential function. Several fully and partially reconstructed backgrounds are also modeled in the fit to the $B^0_s \rightarrow D^- \pi^+$ sample. The tagging efficiency for these samples is found to be higher than for the samples of $B \rightarrow J/\psi X$ decays, due to correlations between the kinematics of the signal $B$ and the opposite-side charm hadrons. The effective mistag fraction for these samples is consistent with that on the $B \rightarrow J/\psi X$ samples. The net effect is an increased tagging power for these $B \rightarrow DX$ decays, similar to that observed for other opposite-side tagging algorithms [7, 28].

To use the charm tagger in a physics analysis, the flavour tagging information from the charm tagger can be combined with information from other tagging algorithms. Assessing the actual gain in performance depends on the method of combination and calibration, as well as on the set of tagging algorithms being combined. Due to correlations with other tagging algorithms, in particular the OS kaon and vertex-charge taggers, the maximum possible increase in tagging power after the addition of the charm tagging algorithm is less than its individual tagging power. The performance of the combination of the current OS tagging algorithms with and without the addition of the charm tagger has been measured on the $B^+ \rightarrow J/\psi K^+$ data sample. The absolute net gain in tagging power using the current combination algorithm is found to be around 0.11%, compared to the current total OS tagging power of about 2.5% [3, 6–8].

7 Conclusion

An algorithm has been developed that determines the flavour of a signal $b$ hadron at production time by reconstructing opposite-side charm hadrons from a number of decay channels. The flavour tagger uses boosted decision tree algorithms trained on simulated data, and has been calibrated and evaluated on data using the self-tagged decay $B^+ \rightarrow J/\psi K^+$. Its tagging power for data in this channel is found to be $(0.30 \pm 0.01 \text{ (stat)} \pm 0.01 \text{ (syst)})\%$. The calibration has been cross-checked using the decay $B^0 \rightarrow J/\psi K^{*0}$, giving consistent results. The tagging power is found to be higher for the decays $B^0 \rightarrow D^- \pi^+$ and $B^0_s \rightarrow D^- \pi^+$, at $(0.40 \pm 0.02 \text{ (stat)} \pm 0.01 \text{ (syst)})\%$ and $(0.39 \pm 0.03 \text{ (stat)} \pm 0.01 \text{ (syst)})\%$, respectively.

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); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG, HGF and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). The Tier1 computing centres are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom). We are indebted to the communities behind the multiple open source software packages on which we depend. We are also thankful for the computing resources and the access to software R&D tools provided by Yandex LLC (Russia). Individual groups
or members have received support from EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), Conseil général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR (Russia), XuntaGal and GENCAT (Spain), Royal Society and Royal Commission for the Exhibition of 1851 (United Kingdom).

References


The LHCb collaboration

European Organization for Nuclear Research (CERN), Geneva, Switzerland

Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Physik-Institut, Universität Zürich, Zürich, Switzerland

Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands

NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine

Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine

University of Birmingham, Birmingham, United Kingdom

H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom

Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom

Department of Physics, University of Warwick, Coventry, United Kingdom

STFC Rutherford Appleton Laboratory, Didcot, United Kingdom

School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom

School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom

Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom

Imperial College London, London, United Kingdom

College of Science and Technology, University of Manchester, Manchester, United Kingdom

Department of Physics, University of Oxford, Oxford, United Kingdom

Massachusetts Institute of Technology, Cambridge, MA, United States

University of Cincinnati, Cincinnati, OH, United States

University of Maryland, College Park, MD, United States

Syracuse University, Syracuse, NY, United States

Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to

Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to

Departamento de Física, Universidad Nacional de Colombia, Bogota, Colombia, associated to

Institut für Physik, Universität Rostock, Rostock, Germany, associated to

National Research Centre Kurchatov Institute, Moscow, Russia, associated to

Yandex School of Data Analysis, Moscow, Russia, associated to

Instituto de Física Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain, associated to

Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to

Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil

P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia

Università di Bari, Bari, Italy

Università di Bologna, Bologna, Italy

Università di Cagliari, Cagliari, Italy

Università di Ferrara, Ferrara, Italy

Università di Urbino, Urbino, Italy

Università di Modena e Reggio Emilia, Modena, Italy

Università di Genova, Genova, Italy

Università di Milano Bicocca, Milano, Italy

Università di Roma Tor Vergata, Roma, Italy

Università di Roma La Sapienza, Roma, Italy

Università della Basilicata, Potenza, Italy

AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland

LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain

Hanoi University of Science, Hanoi, Viet Nam

Università di Padova, Padova, Italy

Università di Pisa, Pisa, Italy

Scuola Normale Superiore, Pisa, Italy

Università degli Studi di Milano, Milano, Italy

† Deceased