IL NUOVO CIMENTO **39 C** (2016) 335
DOI 10.1393/ncc/i2016-16335-5

Colloquia: La Thuile 2016

b-flavour tagging in pp collisions at LHCb

V. Battista on behalf of the LHCb Collaboration

École Polytechnique Fédérale de Lausanne (EPFL) - Lausanne, Switzerland

received 26 July 2016

Summary. — Measurements of CP violation and flavour oscillations of neutral B mesons require the knowledge of the meson flavour at the production time. Flavour-tagging algorithms in the LHCb experiment allow to perform such measurements with very high precision. Recent examples include the determination of the CKM angles 2β and $2\beta_s$. The details of these flavour-tagging algorithms are presented, together with their performances.

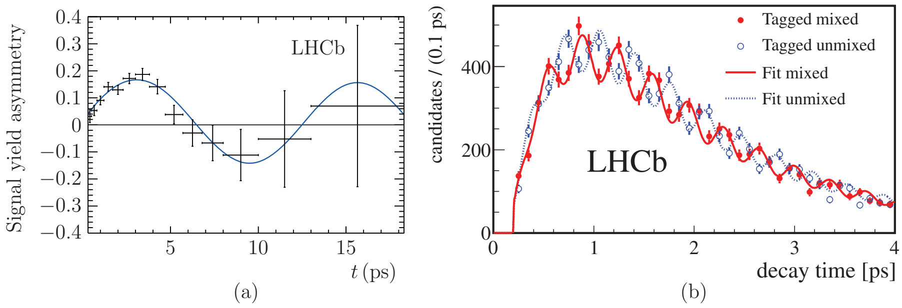
1. – Introduction

The LHCb experiment, a forward spectrometer optimized for b- and c- hadron physics, allows to perform time-dependent analyses with very high precision thanks to excellent resolutions of the decay time, tracks impact parameter and momentum, and the good particle identification [1]. The measurement of time-dependent asymmetries and decay rates of B and $\bar{B}$ mesons relies on the knowledge of the meson flavour at the production time. Examples of these measurements are shown in fig. 1. Flavour tagging algorithms, by exploiting correlations between the B meson flavour and features of the global event, tag the candidate as B or $\bar{B}$ with some efficiency and mistag probability.

A sketch of the LHCb flavour tagging algorithm is presented in fig. 2. Same-side (SS) algorithms rely on the correlation between the flavour of the B candidate and the charge of a particle (proton, kaon or pion) produced in the same hadronisation process of the B candidate. Opposite-side (OS) algorithms exploit the correlation between the flavour of the B candidate and the charge of a particle (pion, kaon, lepton, c-hadron) or the reconstructed secondary vertex produced from the other b-hadron in the event.

2. – Relevant flavour tagging parameters

The performance of a flavour tagging algorithm is quantified by means of the tagging efficiency, the mistag fraction and the tagging power. The tagging efficiency $\epsilon_{\rm tag}$ is the



(a) (b)

Fig. 1. – (a) Time-dependent asymmetry in $B^0 \to J/\psi$ decays [2]. (b) Decay time distributions of mixed and unmixed $B_s^0 \to D_s^- \pi^+$ decays [3].

fraction of tagged events:

$$\epsilon_{\text{tag}} = \frac{N_{\text{tag}}}{N_{\text{tag}} + N_{\text{untag}}}. \tag{1}$$

The tagging efficiency depends on the transverse momentum p_T spectrum of the B meson, and improves for higher p_T. The mistag fraction ω is the fraction of events with a wrong tag decision:

$$\omega = \frac{N_{\text{wrong}}}{N_{\text{wrong}} + N_{\text{right}}}. \tag{2}$$

A non-zero mistag induces a dilution of the time-dependent asymmetry, as shown in fig. 3. A tagging algorithm predicts a mistag probability η which needs to be calibrated via a function $\omega(\eta)$ to provide an unbiased estimate of ω. The tagging power or effective tagging effciency ϵ_{eff} is defined as

$$\epsilon_{\text{eff}} = \epsilon_{\text{tag}} D^2 = \epsilon_{\text{tag}} \langle (1 - 2\omega(\eta))^2 \rangle\,. \tag{3}$$

The tagging power quantifies the effective statistical reduction of the data sample due to the mistag probability and the tagging efficiency. In fact, the statistical uncertainty

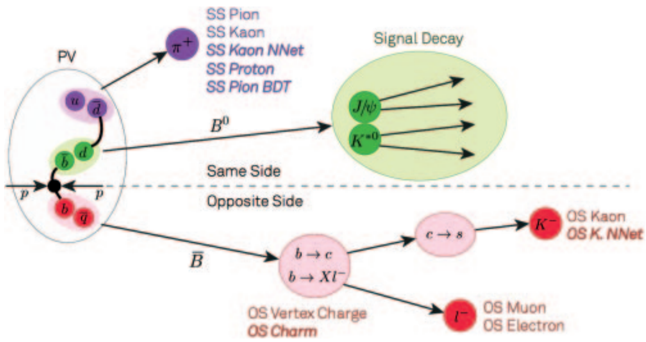


Fig. 2. – Flavour tagging algorithms in LHCb.

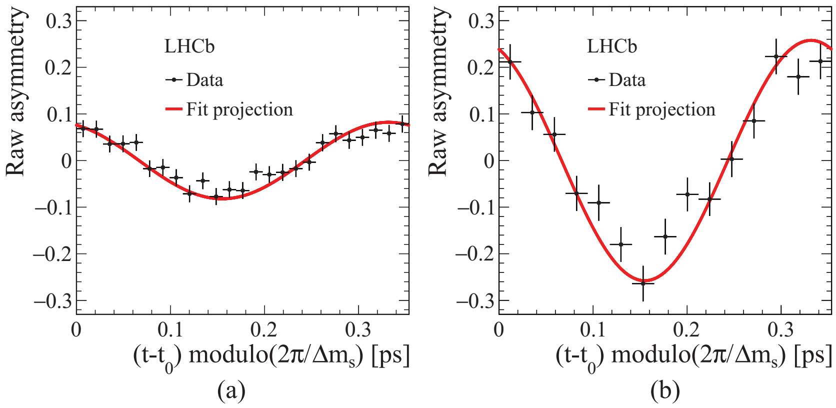


Fig. 3. – Time dependent asymmetry in $B_s \to D_s\pi$ decays for (a) all tagged B candidates and (b) B candidates with $\omega < 0.35$ [4].

on a time-dependent symmetry measured on a sample of size N depends on $\epsilon_{\rm eff}$ as $\sigma \propto 1/\sqrt{\epsilon_{\rm eff}N}$.

3. – Flavour tagging calibration

The predicted mistag probability η is calibrated with data via a linear function $\omega(\eta)$ with parameters p_0, p_1 and $\langle\eta\rangle$, the latter being the average predicted mistag probability. Differences between B and $\bar{B}$ are taken into account with additional parameters Δp_0 and Δp_1:

$$\omega = p_0 + p_1(\eta - \langle\eta\rangle), \tag{4}$$

$$\omega(B) - \omega(\bar{B}) = \Delta\omega = \Delta p_0 + \Delta p_1(\eta - \langle\eta\rangle). \tag{5}$$

Examples of calibration curves are shown in fig. 4. Different decay modes can be used for the calibration. Self-tagged charged B decays ($B^+ \to J/\psi K^+$, $B^+ \to D^0\pi^+$) are used for OS taggers calibration and ensure high statistics and low systematic uncertainties; the charge of the B (true flavour) is compared with the tagger prediction. Neutral B decays ($B^0 \to J/\psi K^*$, $B^0 \to D^{*-}\mu^+\nu_\mu$) require a measure of the B-$\bar{B}$ oscillation amplitude to

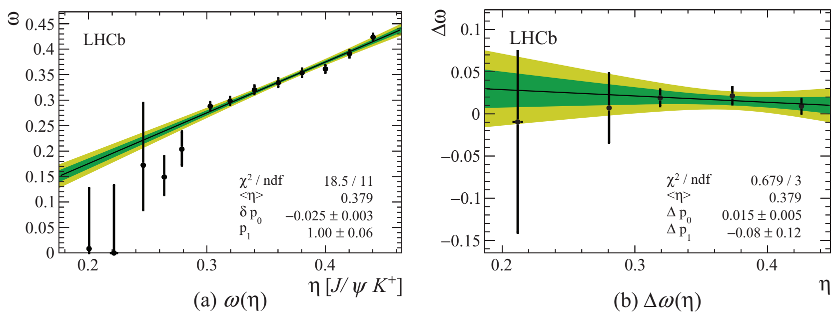


Fig. 4. – Calibration of the OSCharm tagger [5].

TABLE I. – *Performances of the OS taggers.*

Taggers	$\varepsilon_{\rm tag}[\%]$	$\omega[\%]$	$\varepsilon_{\rm tag}(1-2\omega)^2[\%]$
μ	4.8 ± 0.1	29.9 ± 0.7	0.77 ± 0.07
e	2.2 ± 0.1	33.2 ± 1.1	0.25 ± 0.04
K	11.6 ± 0.1	38.3 ± 0.5	0.63 ± 0.06
$Q_{\rm vtx}$	15.1 ± 0.1	40.0 ± 0.4	0.60 ± 0.06
OS average ($\eta_c < 0.42$)	17.8 ± 0.1	34.6 ± 0.4	1.69 ± 0.10
OS sum of η_c bins	27.3 ± 0.2	36.2 ± 0.5	2.07 ± 0.11

infer the mistag ω, which are affected by higher systematics. Finally, $B_s^0 \to D_s^- \pi^+$ and $B_s^{**} \to B^+ K^-$ decays, which suffer a lower statistics, are used for analyses involving B_s^0 mesons.

4. – Opposite-Side taggers

A standard combination of OS taggers (OSComb) is used in LHCb [6], which includes electron (OSe), muon (OSμ), kaon (OSK) and vertex charge (OSVtx) taggers. Electrons, muons and kaons are required to have large impact parameter (IP) and p_T and to match particle identification (PID) criteria. For the OSVtx taggers, two pion tracks compatible with a B decay vertex are combined, and additional tracks are added afterwards. The predicted mistag for each OS tagger is obtained from neural networks (NN) trained on simulated $B^+ \to J/\psi K^+$ events. The NN combines both global information (*e.g.* number of tagging particles and pile-up vertices) and tagging particle properties (*e.g.* kinematics). The mistag probability is then calibrated using $B^+ \to J/\psi K^+$ data samples. The tagging decision and the mistag probability for each tagger are finally combined in a single tagging decision and a single mistag probability. The tagging performances are reported in table I, while the calibrated mistag distribution for each tagger are shown in fig. 5. Thanks to improvements in the selection and the usage of $B^+ \to J/\psi K^+$ data in the training of the NN, a relative increase of $\sim 15\%$ was obtained since 2011.

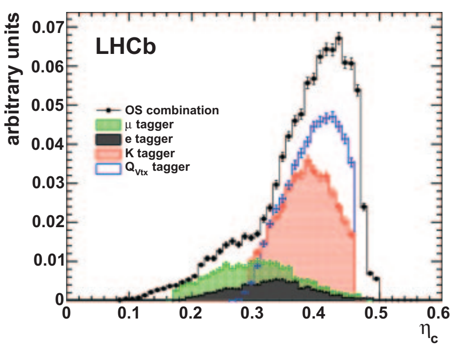


Fig. 5. – Calibrated mistag probabilities for the OS taggers.

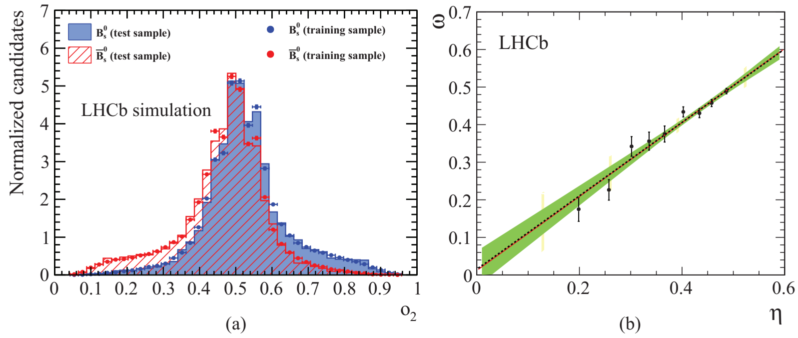


Fig. 6. – (a) Distribution of the output of NN2 for B_s^0 and $\bar{B}_s^0$ candidates. The training and testing samples are superimposed. (b) Calibration curve from the fit of the $B_s^0 \to D_s^- \pi^+$ decay time distribution (red line). The green and yellow bands are the 1σ and 2σ intervals, respectively. The black points correspond to the average ω in bins of η. The black, dashed line is a linear fit to these points.

5. – Neural Network-based Same-Side Kaon tagger

A new NN-based SS kaon algorithm (SSKaonNNet) was recently developed to improve an existing SSKaon algorithm used in LHCb [4]. Two NNs, both trained on simulated $B_s^0 \to D_s^- \pi^+$ events, are implemented to discriminate fragmentation kaons from background tracks (NN1) and to determine tagging decision and mistag probability (NN2). The output of NN1 is used as input feature for NN2. The distribution of NN2 output is shown in fig. 6(a). The calibration is performed with $B_s^0 \to D_s^- \pi^+$ data samples by means of an unbinned maximux likelihood fit of the decay time distribution. The fit is done simultaneously in the untagged, mixed and unmixed samples. In the mixed (unmixed) sample, the B flavour at decay is opposite (equal) to the B flavour at the production time; in the untagged sample, no tag is provided by the tagging algorithms. The predicted mistag η is treated as a per-event observable of the probability density function (PDF), p_0 and p_1 are fitted and $\langle \eta \rangle$ is fixed to 0.4377. The resulting calibration curve is presented in fig. 6(b). This calibration is combined with the calibration obtained from self-tagged, hadronic $B_{s2}^*(5840)^0 \to B^+ K^-$ decays, shown in fig. 7. The portability of the calibration from the calibration sample to different decays used in analyses is

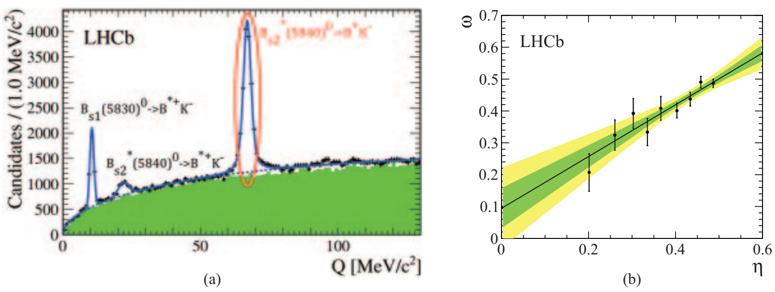


Fig. 7. – (a) $m_{B^+K^-} - M_{B^+} - M_{K^-}$ distribution. (b) Calibration curve obtained from $B_{s2}^*(5840)^0 \to B^+ K^-$ decays.

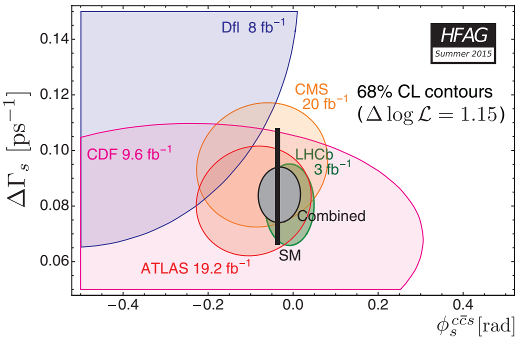


Fig. 8. – World averages for $\Delta\Gamma_s$ and ϕ_s parameters.

checked on $B_s^0 \to J/\psi\phi$, $B_s^0 \to D_s^+ D_s^-$ and $B_s^0 \to \phi\phi$ data samples. The performance is evaluated on $B_s^0 \to D_s^-\pi^+$ data samples. The measured tagging efficiency and tagging power are $\epsilon_{\text{tag}} = (60.38 \pm 0.16)\%$ and $\epsilon_{\text{eff}} = (1.80 \pm 0.19(\text{stat.}) \pm 0.18(\text{syst.}))\%$. The relative improvement of ϵ_{eff} with respect to the previous implementation of the SSKaon tagger is $\sim 50\%$.

The SSKaonNNet tagger, together with OSComb, was used in the measurement of the weak phase ϕ_s via time-dependent analyses of $B_s^0 \to J/\psi K^+K^-$, $B_s^0 \to J/\psi\pi^+\pi^-$ [7] and $B_s^0 \to D_s^+D_s^-$ [8] decays. The tagging power obtained in the $B_s^0 \to J/\psi K^+K^-$ analysis is $\epsilon_{\text{eff}} = (3.73 \pm 0.15)\%$, which represents an absolute improvement of $+0.60\%$ compared to the same analysis performed with a previous version of the SSK tagger [9]. The $B_s^0 \to D_s^+D_s^-$ analysis was the first measurement of ϕ_s in this decay mode; the tagging power obtained is $\epsilon_{\text{eff}} = (5.33 \pm 0.18(\text{stat.}) \pm 0.17(\text{syst.}))\%$. The current world average for ϕ_s [10], which is driven by LHCb measurements, is shown in fig. 8. A summary of the inclusive tagging power for OSComb and SSKaonNNet in the $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to D_s^+D_s^-$ analyses is reported in table II. Other LHCb analyses using SSKaonNNet are $B_s^0 \to J/\psi\pi^+\pi^-$ [11], $B_s^0 \to \phi\phi$ [12] and $B_s^0 \to D_s^-K^+$ [13].

6. – Opposite-Side Charm tagger

A new OS tagger (OSCharm) was implemented recently. [5]. The OSCharm tagger exploits the correlation between the B meson flavour and the flavour of charmed hadrons produced in the decay of the other b-hadron in the event. The charmed hadrons are reconstructed exclusively and partially in several final states, *e.g.* $D^0 \to K^-\pi^+$ or $D^+ \to K^-\pi^+\pi^+$. A boosted decision tree (BDT) is used to both discriminate signal charmed

TABLE II. – *Inclusive tagging power for OSComb and SSKaonNNet algorithms in* $B_s^0 \to J/\psi K^+K^-$ *and* $B_s^0 \to D_s^+D_s^-$ *analyses.*

Tagger	$B_s^0 \to J/\psi K^+K^-$	$B_s^0 \to D_s^+D_s^-$
OSComb	$(2.55 \pm 0.14)\%$	$(3.49 \pm 0.10 \pm 0.17)\%$
SSKaonNNet	$(1.26 \pm 0.17)\%$	$(2.37 \pm 0.23 \pm 0.18)\%$

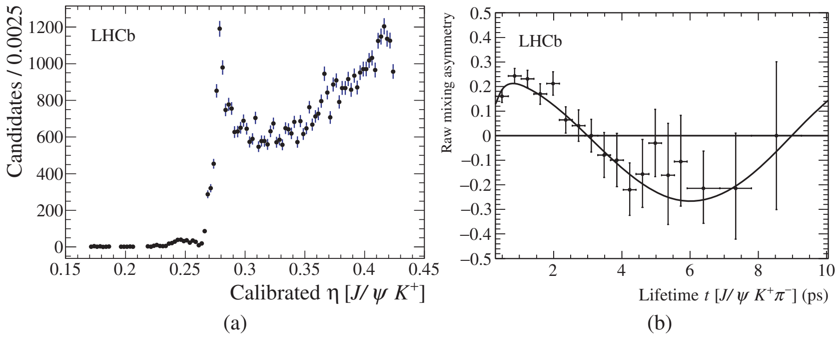


Fig. 9. – (a) Calibrated mistag distribution for OSCharm. (b) Mixing asymmetry in $B^0 \to J/\psi K^{0*}$ decays, used to evaluate OSCharm performance.

hadrons from background and to estimate the mistag probability. The BDT, trained on simulated events, includes different features of the charmed hadron like kinematic quantities, vertex quality and flight distance. The calibration is performed on $B^+ \to J/\psi K^+$ data samples; the calibrated mistag is shown in fig. 9(a). The performance is evaluated on different data samples ($B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{0*}$, $B^0 \to D^-\pi^+$, $B^0_s \to D^-_s$); the tagging efficiency spans the interval between 3.1% and 4.1%, while the tagging power is comprised between 0.3% and 0.4%. The mixing asymmetry in $B^0 \to J/\psi K^{0*}$ decays is presented in fig. 9(b). A test of the combination between OSComb and OSCharm on $B^+ \to J/\psi K^+$ data samples is performed as well; the resulting tagging power had an absolute increase of $\sim 0.11\%$ compared to the performance of OSComb only ($\epsilon_{\mathrm{eff}} \sim 2.5\%$).

7. – Future developments

New BDT-based SS taggers, SSPionBDT and SSProtonBDT, are currently under study. For both taggers, a BDT is trained on $B^0 \to D^\mp\pi^\pm$ data samples in order to discriminate signal pions and protons from background tracks, and to evaluate the mistag probability η. In this training, the decay time of the B is required to be smaller than 0.2 ps to suppress B oscillations. The calibration is performed by evaluating the average mistag in bins of the BDT output. The tagging power measured on $B^0 \to D^\mp\pi^\pm$ is $\sim 0.5\%$ for SSProtonBDT and $\sim 1.6\%$ for SSPionBDT (the latter represents a relative improvement of $\sim 20\%$ with respect to a previous non BDT-based implementation of SSPion).

A new inclusive tagger is under development as well. This tagger relies on a BDT which includes features related to the signal B meson and reconstructed tracks/vertices from the entire event, and does not make any distinction between SS and OS.

REFERENCES

[1] LHCb Collaboration (Aaij R. *et al.*), *Int. J. Mod. Phys. A*, **30** (2015) 1530022.
[2] LHCb Collaboration (Aaij R. *et al.*), *Phys. Rev. Lett.*, **115** (2015) 031601.
[3] LHCb Collaboration (Aaij R. *et al.*), *New J. Phys.*, **15** (2013) 053021.
[4] LHCb Collaboration (Aaij R. *et al.*), *JINST*, **11** (2016) P05010.

[5] LHCb Collaboration (Aaij R. *et al.*), *JINST*, **10** (2015) P10005.
[6] LHCb Collaboration (Aaij R. *et al.*), *Eur. Phys. J. C*, **72** (2012) 2022.
[7] LHCb Collaboration (Aaij R. *et al.*), *Phys. Rev. Lett.*, **114** (2015) 041801.
[8] LHCb Collaboration (Aaij R. *et al.*), *Phys. Rev. Lett.*, **113** (2014) 211801.
[9] LHCb Collaboration (Aaij R. *et al.*), *Phys. Rev. D*, **87** (2013) 112010.
[10] Heavy Flavor Averaging Group (Amhis Y. *et al.*), arXiv:1412.7515
[11] LHCb Collaboration (Aaij R. *et al.*), *Phys. Lett. B*, **736** (2014) 186.
[12] LHCb Collaboration (Aaij R. *et al.*), *Phys. Rev. D*, **90** (2014) 052011.
[13] LHCb Collaboration (Aaij R. *et al.*), *JHEP*, **11** (2014) 060.