Flavour Tagging Algorithms and Performances in LHCb

Marta Calvi[1], Olivier Leroy[2], Marco Musy[1]

[1] *University of Milano-Bicocca and INFN, Milano, Italy*
[2] *CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*

Abstract

In this note we describe the general characteristics of the LHCb flavour tagging algorithms and summarize the tagging performances on the Monte Carlo samples generated for the Data Challenge 2004 in different decay channels. We also discuss some systematics effects and possible methods to extract the mistag fraction in real data.

Contents

1   Introduction

The identification of the initial flavour of reconstructed B^0 and B_s^0 mesons (flavour tagging) is necessary for most of the measurements of CP asymmetries and flavour oscillations. The statistical uncertainty on the measured CP asymmetries is directly related to the effective tagging efficiency $\varepsilon_{\mathrm{eff}}$, defined as

$$\varepsilon_{\mathrm{eff}} = \varepsilon_{\mathrm{tag}} D^2 = \varepsilon_{\mathrm{tag}}(1 - 2\omega)^2 \,, \tag{1}$$

$$\varepsilon_{\mathrm{tag}} = \frac{N_R + N_W}{N_R + N_W + N_U} \qquad \omega = \frac{N_W}{N_R + N_W} \,, \tag{2}$$

where $\varepsilon_{\mathrm{tag}}$ is the tagging efficiency (fraction of events in which the tagging procedure gives an answer), D is the dilution factor, and ω is the wrong tag fraction (probability for the answer to be incorrect when a tag is present). N_R, N_W, N_U are the number of correctly tagged, incorrectly tagged, and untagged events, respectively.

This note is meant as an update of the results obtained in [1] and a detailed description of the improvements on the LHCb flavour tagging algorithms.

In Section 2, we present the simulation framework. In Section 3, we describe the different algorithms designed to tag the B-meson flavour. The combination of taggers is explained in Section 4 and the resulting performances are given in Section 5. Systematics effects are discussed in Section 6.

2   Simulation framework

The physics generator used is Pythia 6.226 [2]. This model includes the description of multiple parton-parton interaction with varying impact parameter. The track multiplicity has been tuned to reproduce CDF low energy data. The Monte Carlo also includes the effect of multiple interactions in the same bunch crossing (pileup).

The response of the detector is simulated in a realistic way including noise and "spillover" effects (event acquisition can be affected by the previous or subsequent acquisition). Reconstruction and selection algorithms do not make use of the true Monte Carlo information at any stages. This means that track reconstruction, particle identification with RICH, calorimetry and muon systems are fully realistic. The truth information has been used only to determine whether a tag is correct or not, and for evaluating efficiencies.

This particular study has been carried out in the analysis framework of the LHCb C++ object oriented software and the standard tools for the physics analysis [3].

Event selection in the various decay channels is provided by individual off-line selection algorithms, which are described in notes for individual channel selection. In the present note, results refer to event samples passing Level-0, Level-1 and HLT (High Level Trigger) trigger selections, unless otherwise specified.

3   Flavour tagging of B mesons

Different sources of information can be used to asses the initial flavour of a B-meson candidate (signal B), as illustrated on figure 1. In the following Sections, two types of
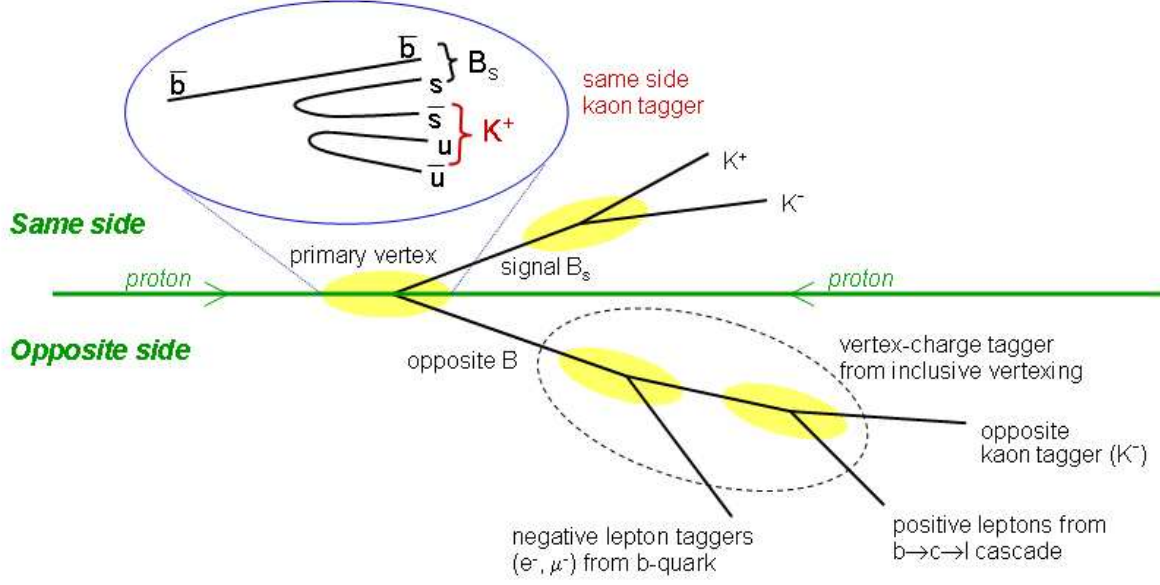
Figure 1: Schematic representation of the different sources of information available to tag the initial flavour of a signal B candidate, here $B_s^0 \to K^+K^-$. The *same* side and *opposite* side (which can be any b-hadron) are shown. Since LHCb is a forward spectrometer, *same* and *opposite* tagging particles can be close in phase-space.

tagger algorithms are described: opposite-side taggers, based on muons, electrons, kaons and inclusive secondary vertex, which are used to tag all types of B-mesons and same-side taggers based on kaons and pions which are used to tag B_s^0 and B^0 or B^+ mesons respectively. The selection cuts used to define the taggers are summarized in Table 1.

If more than one primary vertex is reconstructed in the event, all single particle taggers are required to have an impact parameter significance IP/σ_{IP} in excess of 3 with respect to any primary vertex which was not chosen as the B signal production vertex. When more than one track is selected, the one with highest p_T is chosen. The polar angle of the track with respect to the beam pipe axis is required to be larger than 12 mrad to reject badly reconstructed tracks.

3.1 Particle identification

A reliable Particle IDentification (PID) is of great importance for both event selection and flavour tagging. The combined information of the RICH, Electromagnetic Calorimeter and Muon system have been used to tune the PID cuts to obtain the best performance in the flavour tagging. Examples of Log-Likelihood distributions for PID determination are shown in figure 2. The arrows indicate the cuts applied that maximize the final effective tagging efficiencies. All these distributions correspond to the tracks which are used as tagger candidates. For kaons and pions, not only *long* tracks have been considered as tagger candidates, but also *upstream* tracks with a tighter requirement on the χ^2 from
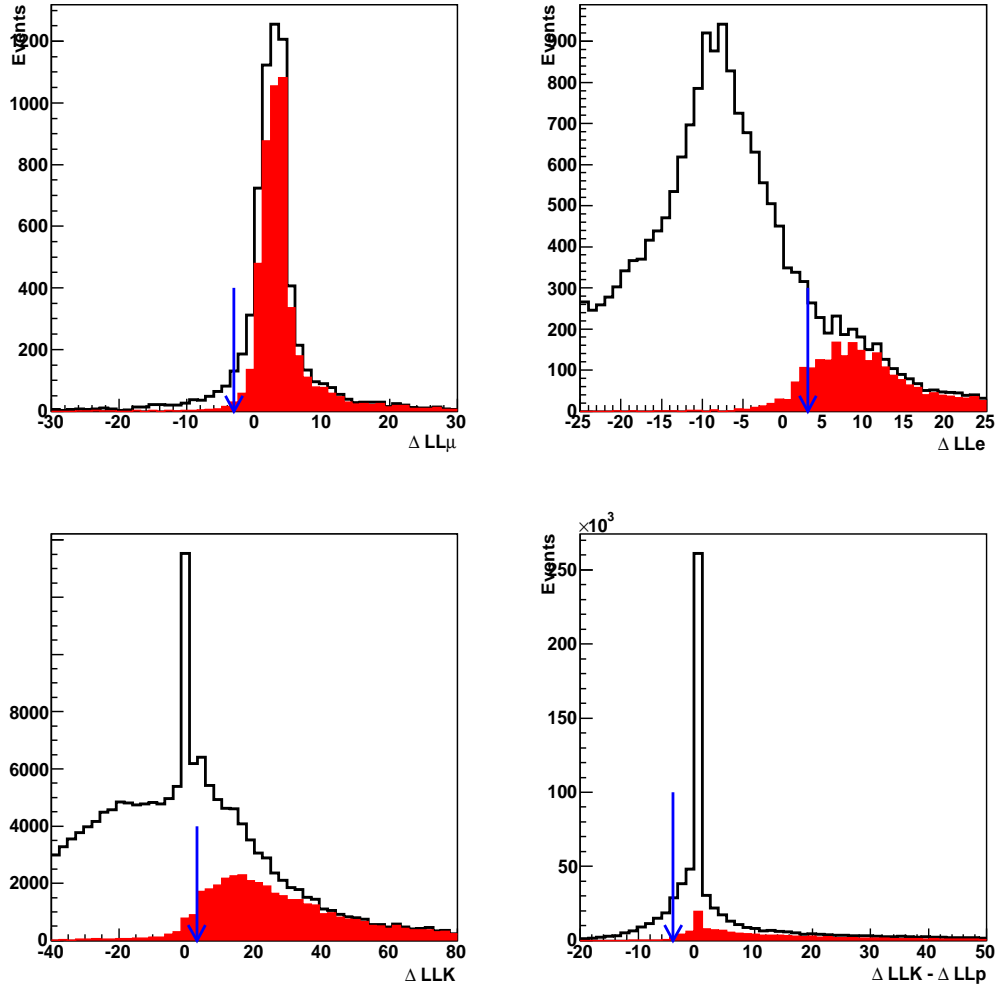
Figure 2: Log-Likelihood distributions for tagging candidates (top: muons and electrons, bottom: opposite kaons) in $B_s^0 \to D_s^- \pi^+$ events. Shaded (red) histograms correspond to the distributions for true muons, electrons and kaons compared to pion hypothesis, and for kaons compared to proton hypothesis (bottom right plot). Open histogram corresponds to the likelihood distributions for tracks which are not of that given type.

3

the tracking system[1]. Long tracks can be reconstructed by two types of algorithms called *forward* or *matched*. A tighter χ^2 cut is applied to the matched tracks. Additional information on how these likelihood functions are built can be found in [4].

With these requirements, the *ghost* rate for muons, electrons, kaon-opposite and kaon-same side taggers is 2.9%, 3.2% 5.9% and 1.0% respectively. *Ghost* rate refers to tracks which have been reconstructed, but are not associated to any track at Monte Carlo truth level.

PID purities, defined as the fraction of true-type particles among the identified ones, are of the order of 76%, 85% and 74% for muons electrons and kaons respectively while cut efficiencies are 87%, 75% and 80% respectively when considering only *long* tracks of momentum $p > 5\,\mathrm{GeV}/c$ in $\mathrm{B}_\mathrm{s}^0 \rightarrow \mathrm{D}_\mathrm{s}^- \pi^+$ events. A lower ghost rate for Same-Side-Kaons (SS-Kaon) with respect to Opposite-Side-Kaons (OS-Kaon) is justified by the fact that a requirement on impact parameter significance with respect to the primary vertex is imposed on OS-Kaons to be greater than 3. The subsample of tracks which are not pointing to the primary vertex includes a higher fraction of tracks reconstructed with random hits combinations.

3.2 Opposite Side taggers

3.2.1 Single particle taggers

Opposite-side tag algorithms use the charge of the lepton from semileptonic b decay and of the kaon from the b $\rightarrow$ c $\rightarrow$ s decay chain to tag the flavour of the B meson.

In the case of opposite-side muon tagger, a momentum $p > 5\,\mathrm{GeV}/c$ and $p_\mathrm{T} > 1.2\,\mathrm{GeV}/c$ is required, reducing the contribution from b $\rightarrow$ c $\rightarrow \ell$ decays which would tag the wrong charge. Figure 3 shows how the optimal cut on p_T is evaluated. The upper plot shows the distributions for correct-tag and wrong-tag decisions of the muon tagger. The mistag ω, calculated in each separate bin, is also shown in the middle plot. Finally the effective efficiency ε_eff is calculated as a function of minimum p_T required. The cut value indicated by the arrow is set to the beginning of the plateau.

An additional algorithm, called "Non Shared Hits" (NSH), has been developed in order to avoid the identification of fake muons (actually true pions) due to the fact that close-by tracks may share the same hits of a true muon track, giving rise to an extra muon that can be eligible to be an opposite tagger. This algorithm is particularly useful in all decay channels where there is a muon in the final state. Its output is used to veto the use of such tracks as muon taggers.

To select OS-Kaon candidates, a momentum $p > 3\,\mathrm{GeV}/c$, $p_\mathrm{T} > 0.4\,\mathrm{GeV}/c$ and an impact parameter with respect to the primary vertex with significance $\mathrm{IP}/\sigma_\mathrm{IP} > 3.5$ are required. These cuts enhance the contribution of kaons from b decays with respect to kaons produced in the fragmentation. Figure 4 shows the distributions in $\mathrm{IP}/\sigma_\mathrm{IP}$ for right and wrong tags and the optimal cut value with the same meaning as explained for Figure 3.

For electron taggers, the cut on transverse momentum becomes $p_\mathrm{T} > 1\,\mathrm{GeV}/c$. In addition, a cut on the ionization charge deposited in the silicon layers of the Vertex Locator is also applied. This helps reducing the background components coming from

---

[1]Tracks reconstructed by different algorithms are defined in [4]. Essentially long tracks have hits along all the tracking devices while upstream tracks only up to the Trigger stations.
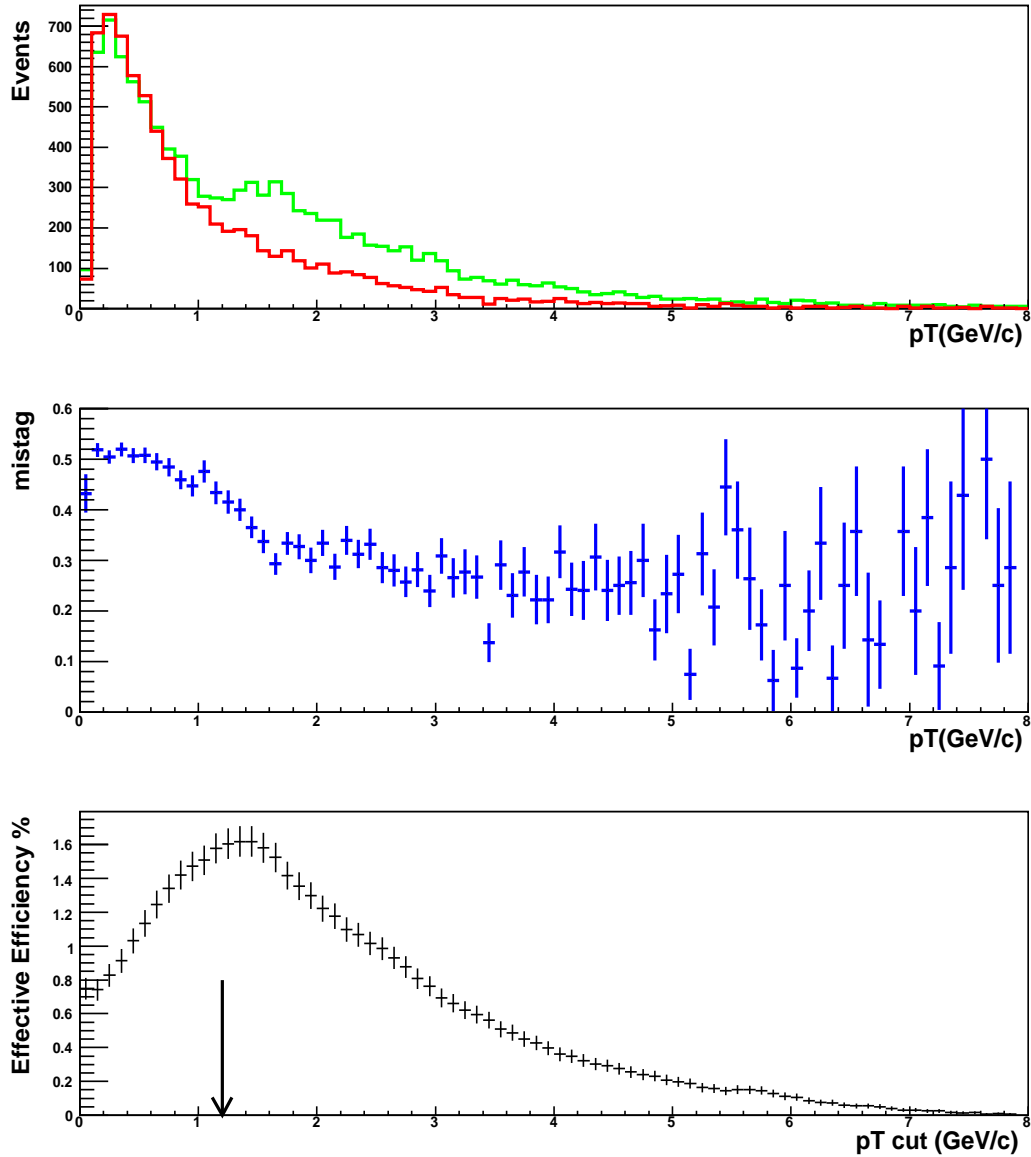
Figure 3: Distributions for the p_T of the muon tagger. Upper plot: light histogram (green) is for events correctly tagged, dark histogram (red) is for wrong tag decision. Middle plot: the wrong tag fraction ω evaluated bin by bin of the above plot. Bottom plot: effective efficiency calculated as a function of the cut applied on the p_T. Each bin corresponds to the integral from that p_T value to infinity. The region to the left of the arrow is excluded by the cut. Error bars in this plot are correlated.
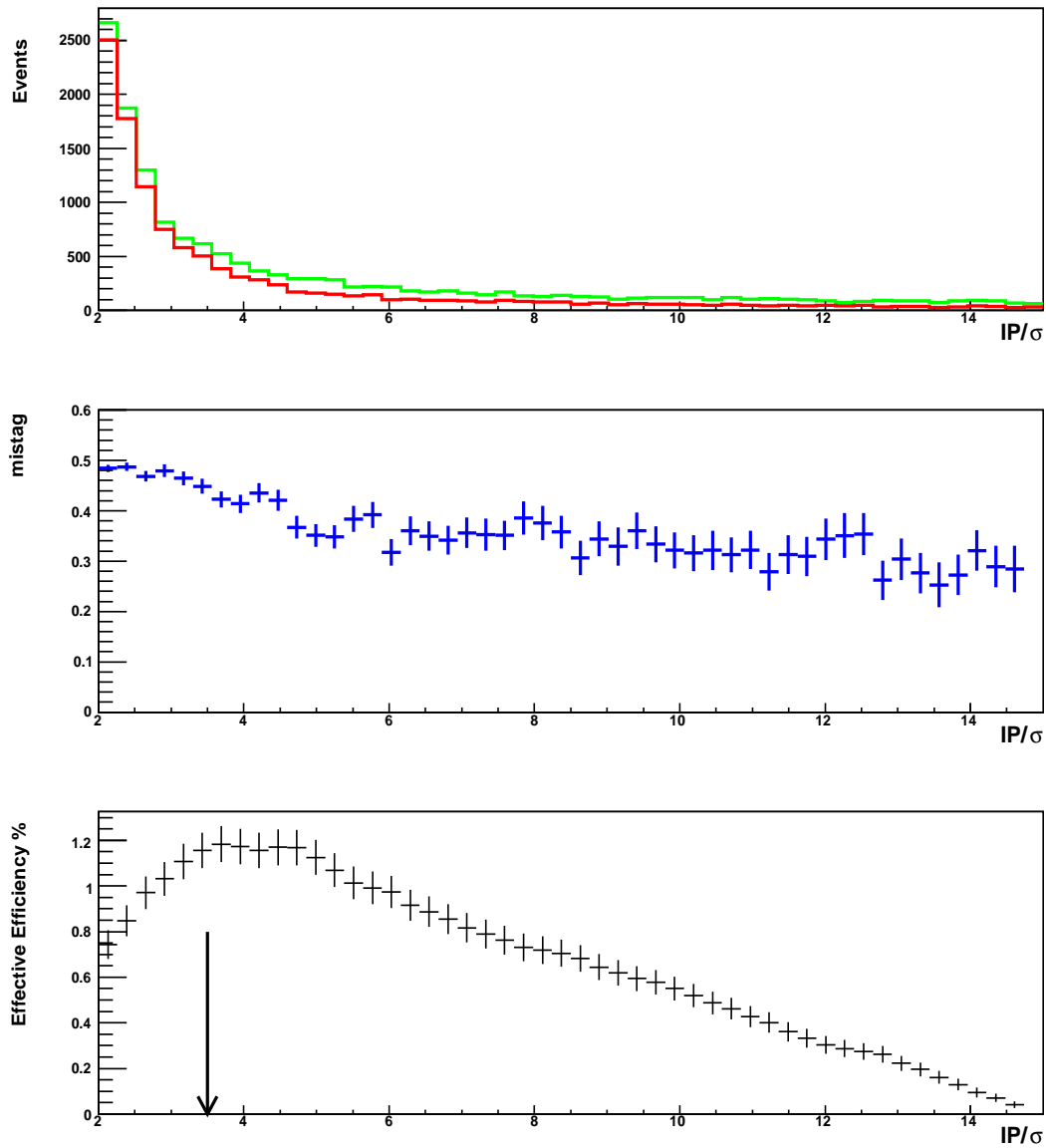
Figure 4: Distributions for the impact parameter significance IP/σ_{IP} of the OS-Kaon tagger, with respect to the primary vertex. Upper plot: light histogram (green) is for the correct tags, dark histogram (red) is for wrong tag decision. Middle plot: the wrong tag fraction ω evaluated bin by bin from the above plot. Bottom plot: effective efficiency calculated as a function of the cut applied to IP/σ_{IP}. Each bin corresponds to the integral from that p_T value to infinity. The region to the left of the arrow is excluded by the cut. Error bars in the plot are correlated.
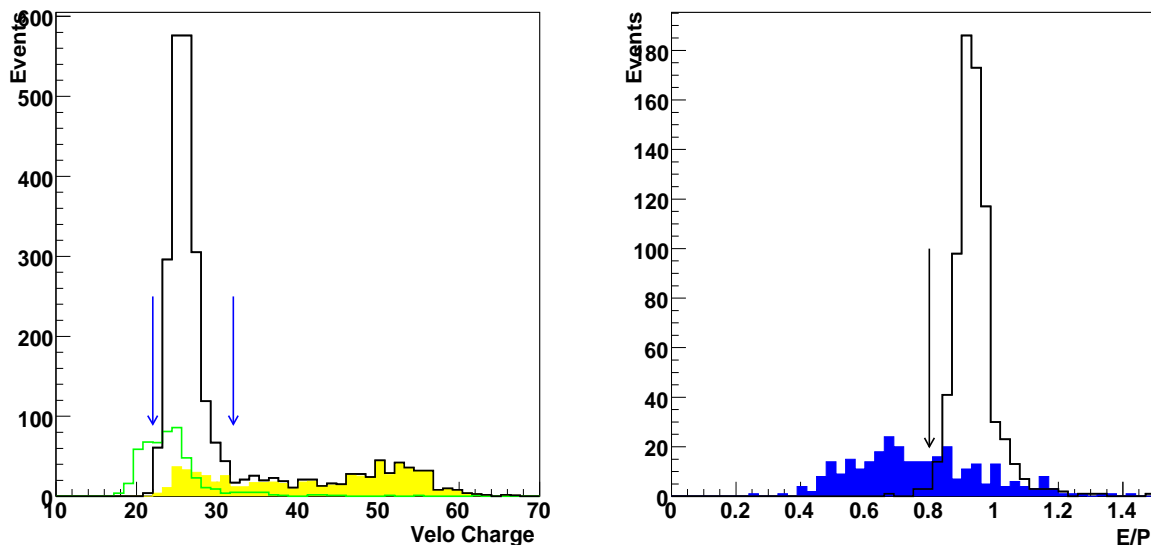
Figure 5: Left: distribution of the ionization charge deposited in the silicon layers of the Vertex Locator. Shaded histogram (yellow) indicates the electrons coming from photon conversions, while the open (green) histogram represents the hadron background component. Black open histogram corresponds to true electrons. Velo Charge units are arbitrary. Right: distribution of energy over momentum of the electron candidate. The shaded histogram corresponds to the mis-ID component. Cuts are indicated by the arrows.

photon conversions near the interaction point and from misidentified hadrons. Two cuts on this variable are applied and are indicated by the arrows in Figure 5. One other useful variable is the ratio E/p of energy measured in the electromagnetic calorimeter over momentum of the candidate electron, measured with the tracking system. The corresponding distribution is shown in the same figure on the right plot. The cut $E/p >$ 0.8 is applied on this quantity. No cut on impact parameter is required for both muon and electron taggers.

3.2.2 Vertex Charge tagger

An inclusive reconstruction of the opposite b-hadron is performed by means of the inclusive reconstruction of secondary vertex. The general description of how this is constructed is given in [1]. A secondary vertex is found in 44.7% of the events which passed Level-0 and Level-1 trigger selection. The mean number of charged tracks at the secondary vertex is 3.1, 2.6 of which are true b-hadron decay products.

Figure 6 shows the difference in ϕ with respect to the Monte Carlo truth of the b-hadron flight directions, determined from the primary to secondary vertex, using generated vertices coordinates or reconstructed vertices coordinates. The core resolution in ϕ is about 45 mrad which corresponds to about 80% of the total events.

The inclusive reconstruction of the accompanying b decay vertex is used to determine the b-hadron charge. The weighted vertex charge is defined as the normalized sum of the

	$p>$ (GeV/c)	$p_\mathrm{T}>$ (GeV/c)	$\chi^2/\mathrm{ndf}<$ (track)	IP/σ_IP	PID cuts	extra cuts
μ	5	1.2	–	–	$\Delta\mathrm{LL}(\mu-\pi)>-3$ NSH	–
e	5	1.0	–	–	$\Delta\mathrm{LL}(\mathrm{e}-\pi)>3$ $21<$veloch<32 $E/p>0.8$	–
$\mathrm{K_{opp}}$	3	0.4	2.5 (L-F) 1.4 (L-M) 2.5 (U)	>3.5	$\Delta\mathrm{LL}(\mathrm{K}-\pi)>3$ $\Delta\mathrm{LL}(\mathrm{K}-\mathrm{p})>-4$	IP<2 mm IP<1 mm IP<1 mm
$\mathrm{K_{same}}$	4	0.4	2.5 (U)	<2.5	$\Delta\mathrm{LL}(\mathrm{K}-\pi)>3$ $\Delta\mathrm{LL}(\mathrm{K}-\mathrm{p})>-4$	$\Delta\phi<1.1$ $\Delta\eta<1.0$ $\Delta m>1.5\,\mathrm{GeV}/c^2$
π_same	2	0.2	2.5 (U)	<3	–	$\Delta\phi<1.1$ $\Delta\eta<1.1$ $\Delta m>1-3\,\mathrm{GeV}/c^2$

Table 1: Summary of the cuts used to define individual tagger candidates. L-F, L-M and U refer respectively to Long-Forward, Long-Matched and Upstream tracks. The impact parameter significance is computed with respect to the primary vertex. An additional cut on IP is applied for OS-Kaons to reduce the *ghost* fraction. The velo charge units are arbitrary (see Fig. 5). The other cuts are described in the text.

charges of all tracks associated to the vertex weighted with p_T^κ:

$$Q_\mathrm{vtx} = \frac{\sum_i p_\mathrm{T}^\kappa(i) Q_i}{\sum_i p_\mathrm{T}^\kappa(i)} \tag{3}$$

The κ parameter is optimized together with a cut on the non-discriminating central part of the Q_vtx distribution, in order to maximize the effective tagging efficiency. This leads to $\kappa = 0.4$ and to consider events with $|Q_\mathrm{vtx}| < 0.3$ as untagged. Figure 6 shows the distribution of the weighted vertex charge when the decaying b-hadron is charged. The peaks at $Q_\mathrm{vtx} = \pm 1$ are populated by events with vertices formed by tracks of the same charge.

We have tried to improve this estimator by adding tracks from fragmentation in a cone around the opposite b-hadron flight direction. The cone radius is

$$\rho_j = \sqrt{(\phi_j - \phi_{\mathrm{B^{opp}}})^2 + (\eta_j - \eta_{\mathrm{B^{opp}}})^2},$$

where j is an iterator on the charged tracks and η is the pseudo-rapidity. The optimization gives a maximum radius $\rho = 0.2$, where only those tracks which satisfy $\rho_j < \rho$ are included. Adding more tracks would dilute the information brought by the tracks associated to the inclusive vertex. Since the improvement obtained is marginal, only the vertex charge defined by Eq. (3) is finally used.
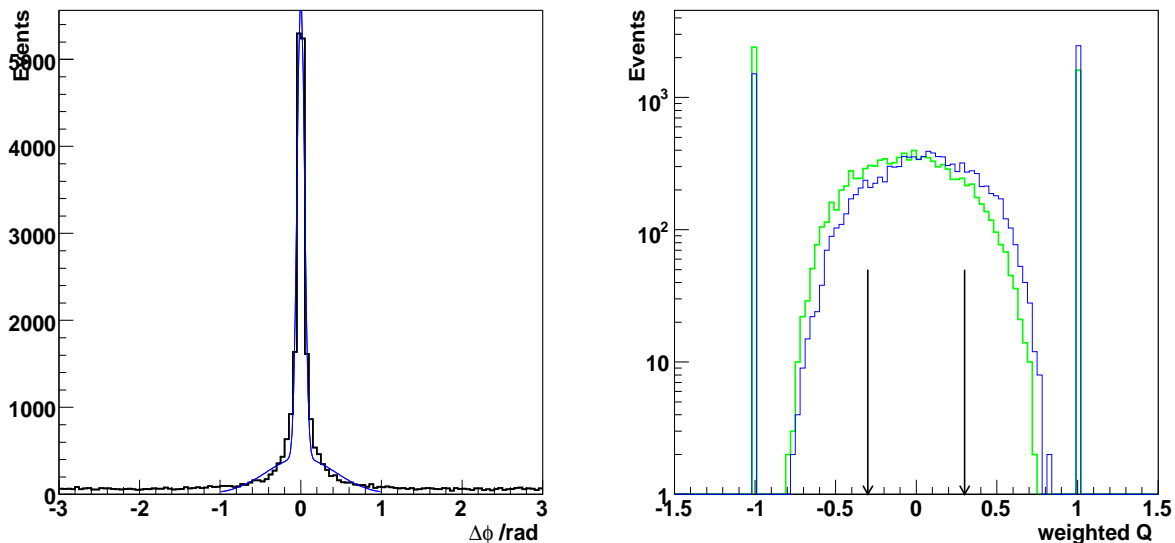
Figure 6: Left: Difference in azimuthal angle ϕ of the opposite b-hadron flight directions, determined from the primary to secondary vertex direction, using generated vertices coordinates and reconstructed vertices coordinates. Right: inclusive weighted vertex charge distribution in logarithmic scale for decaying anti-b-hadrons (lighter histogram shifted to the left) and b-hadrons.

3.3 Same Side taggers

3.3.1 Kaon Same Side tagger

The same side tagging algorithms determine the flavour of the signal B_s^0 meson by exploiting the correlation in the fragmentation decay chain. If a B_s^0 ($\bar{b}s$) is produced in the fragmentation of a $\bar{b}$ quark, an extra $\bar{s}$ is available to form a hadron, which leads to a charged K in about 50% of the cases[2]. These kaons are selected requiring an impact parameter with respect to the primary vertex with a significance $IP/\sigma_{IP} < 2.5$, a difference in pseudo-rapidity with respect to the reconstructed B_s^0 $|\Delta\eta| < 1$, a difference in ϕ angle $|\Delta\phi| < 1.1$ rad and $\Delta m < 1.5\,\mathrm{GeV}/c^2$, where Δm is the difference between the mass of the $B_s^0 K$ combination and the mass of the reconstructed B_s^0. Figure 7 on the left shows the difference between the azimuthal angle of the signal B meson and the same side kaon tagger. For this tagger we require also $p > 4\,\mathrm{GeV}/c$ and $p_{\mathrm{T}} > 0.4\,\mathrm{GeV}/c$.

3.3.2 Pion Same Side tagger

In B^0 events the same principle is applied as for SS-kaons, but in this case a pion is originated in the fragmentation chain or from excited state of the B^0. In the Pythia event generator the following decays are included: $B^{0*} \rightarrow B^0\gamma$, $B^{0**} \rightarrow B^{0(*)}\pi^0/\pi^+\pi^-$, $B^{0**} \rightarrow B^{+(*)}\pi^-(\pi^0)$, $B^{+**} \rightarrow B^{0(*)}\pi^+(\pi^0)$ and $B^{+**} \rightarrow B^{+(*)}\pi^0/\pi^+\pi^-$.

---

[2]Even if a K^{*0} is produced, it leads to a right-sign K^+. However, if a ϕ is produced, the probability to have a correct tag is 50%.
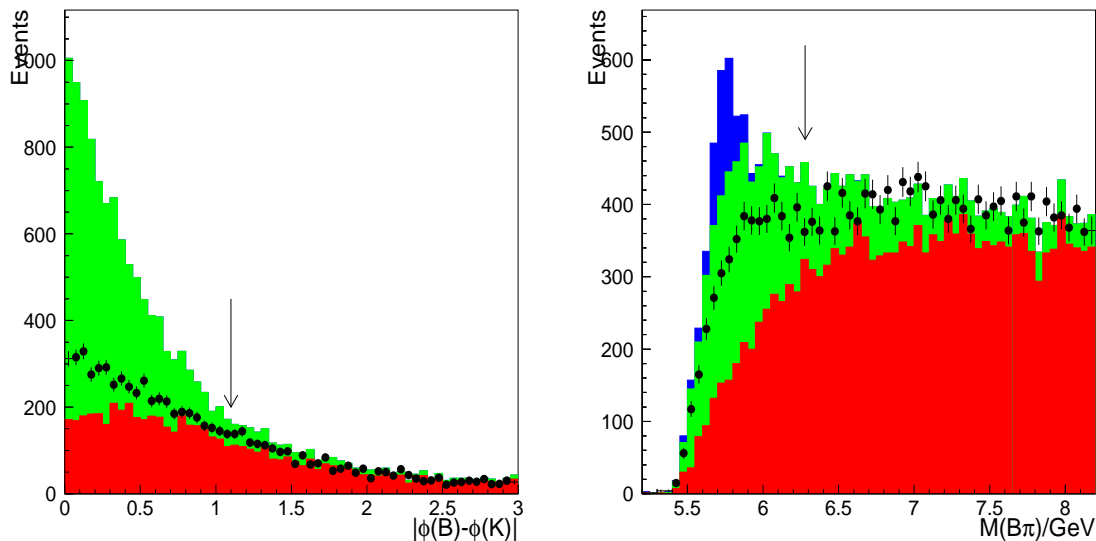
Figure 7: Left: distribution of the difference between the azimuthal angle of the signal B meson and the kaon, for same side kaon taggers in $B_s^0 \to D_s^{\mp} K^{\pm}$ events. Histograms show right-tagging kaons coming from the same fragmentation string of the B_s^0 (green) and from other sources (red). Points with error bars indicate wrong-tagging kaons. Right: distribution of the invariant mass of the Bπ system in $B^0 \to \pi^+\pi^-$ events for candidate same side pion tagger passing kinematic cuts. In blue pions from B^{+**} decays $B^{+**} \to B^{0(*)}\pi^+(\pi^0)$ which always bring the correct tag for the flavour of the B, in green (grey) are π coming from the same fragmentation string of the B and bringing the correct tag, in red (dark grey) all the other right-tagging pions. Points with error bars correspond to the distribution of wrong tagging pions. Cut values are indicated by the arrows.

The masses, the widths and the fractions of B-mesons are set to reproduce the experimental data available from LEP and CDF [5]. The relative fractions of weakly decaying B-mesons result in 21.0%, 62.9% and 16.2% for B^0, B^{0*} and B^{0**} respectively.

Figure 7 shows the distribution of the invariant mass of the $B^0\pi$ system $M(B^0\pi) = m_B + \Delta m$ in $B^0 \to \pi^+\pi^-$ events for candidate same side pion taggers passing kinematic cuts. Candidates are required to have a $p_T > 0.2 \, \text{GeV}/c$, $p > 2 \, \text{GeV}/c$, $\text{IP}/\sigma_{\text{IP}} < 3$ and $\Delta m < 3 \, \text{GeV}/c^2$. Among all these, the highest p_T particle is chosen. Finally an additional requirement on this pion $\Delta m < 1 \, \text{GeV}/c^2$ is imposed. Of all pions coming in the B^{0**} decay, 69% are in the LHCb acceptance and 19% are selected. The width of the peak is about 80 MeV/c^2.

4   Combination of taggers

The final tagging decision on the production flavour of the reconstructed B candidate can be taken following different strategies. In a previous study reported in [1], the final decision is made by looking at the type of the particle (whether it is a muon, electron or kaon) used as a tagger for the event. Events are therefore subdivided into different

categories depending upon all the possible combination of tagging particles. Another possibility, which turns out to improve the final performance by about 25%, is to assign a individual probability p_i of being correct to each separate tagger i, and to sort events by looking at the combined probability of correct tagging. The probability p_i is a function of the kinematic properties of each of the single taggers, and it is evaluated by means of a *neural net* which has been trained on Monte Carlo events. The only true information which the net needs to access during the learning phase is the correct flavour of the B meson. On real data, the learning phase will be performed on self tagging control channels, large amount of which will be available, as will be discussed in Section 6.

Figure 8 shows two examples of neural net output for the case of the OS-Muon and SS-Kaon taggers. The neural net is a *multi-layer perceptron* which, in the case of the muon, takes as an input the charged track multiplicity of the event (see Fig. 11), the p_T of the track, its momentum (see Fig. 3), and the impact parameter significance (in the
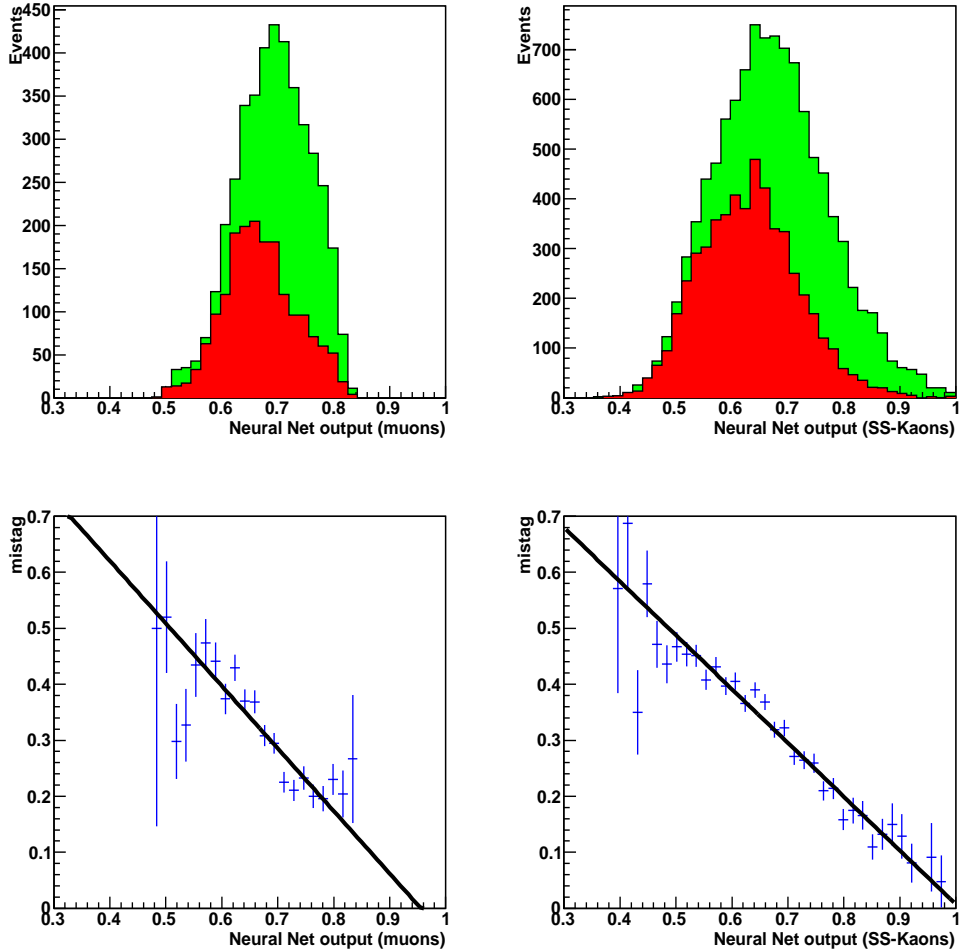

Figure 8: Neural net output for the case of the OS-Muon (left column) and SS-Kaon taggers (right column) in $B_s^0 \rightarrow D_s^- \pi^+$ events. Light shaded (green) is for right-tag while dark shaded (red) histogram is for wrong-tag assignment. In the second row, the corresponding mistag value ω is fitted from the neural net output with a first order polynomial.
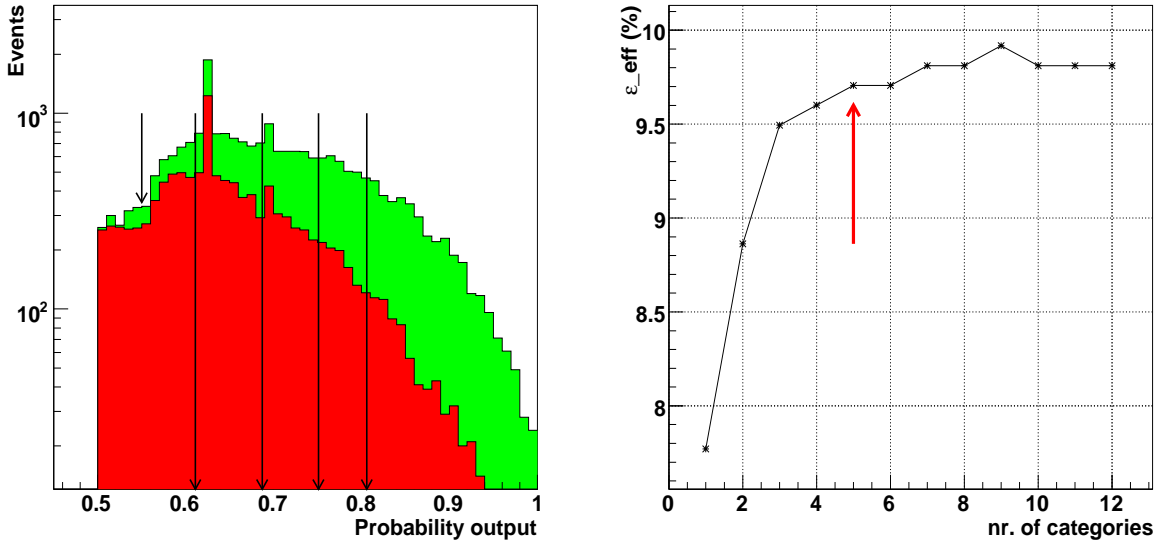
Figure 9: Left: distribution of the probability of the final tagging decision evaluated from the output of the neural net. Light (green) and dark (red) histograms represent the cases where the event is tagged with correct flavour or wrong flavour respectively. Arrows indicate the subdivision into five tagging categories. Events below the first short arrow are considered untagged. Right: effective efficiency as a function of the number of categories. The arrow indicate the number of categories eventually chosen.

case of same-side taggers also $\Delta\phi$, $\Delta\eta$ and Δm are included). It is constituted by an intermediate layer of six neurons, and one output neuron which gives the probability for the input tagger to tag a b or a $\bar{\text{b}}$. A corresponding probability of wrong tag is fitted as function of this output by looking at the true flavour of the B-meson to be tagged, eventually on control channels.

This procedure is repeated for each tagger separately, ending up in five decisions, where each tagger "votes", independently in the first approximation[3], for the flavour of the B. These five separate decisions need to be combined into a global decision.

If p_i is the probability associated to the i-th tagger, and $q_i = \pm 1$ is the decision taken by the i-th tagger, which was made based on the charge and the kinematic properties of the particle, then the combined probability $\mathcal{P}(b)$ that the B meson contains a b-quark is calculated as

$$\mathcal{P}(b) = \frac{p(b)}{p(b) + p(\bar{b})}, \quad \text{and} \quad \mathcal{P}(\bar{b}) = 1 - \mathcal{P}(b) \tag{4}$$

where

$$p(b) = \prod_i \left(\frac{1 - q_i}{2} + q_i p_i \right), \quad \text{and} \quad p(\bar{b}) = \prod_i \left(\frac{1 + q_i}{2} - q_i p_i \right) \tag{5}$$

---

[3]The five taggers considered so far are indeed correlated. In particular the Q_{vtx} tagger is rather correlated with all other opposite taggers. On the contrary the same side tagger is not significantly correlated with any other tagger. These correlations can cause a loss in the discriminating power of the flavour tagging algorithms, but are not expected to cause any bias in the determination of the mistag, as in any case ω will be measured on real data events on control channels.

12

The final decision on the flavour is therefore taken based on whether $\mathcal{P}(b) > \mathcal{P}(\bar{b})$ or vice versa. Those events where $0.45 < \mathcal{P}(b) < 0.55$ are considered untagged. Only for the SS-Pion tagger, if $0.48 < p_\pi < 0.52$ the tagger is suppressed in order not to dilute the other taggers with high values of mistag probability of pions.

Figure 9 on the left shows in logarithmic scale the outcome for such a probability for right and wrong tagged events. If one would not consider how events distribute in that plot, but use them as such to evaluate CP asymmetries, then "bad" events (i.e. high-ω) would mix up with "good" events (i.e. low-ω) yielding an effective tagging efficiency that corresponds to the "Average" line of Table 2 and to the first point in the plot on the right of Figure 9. Grouping together events with similar ω, estimated from the neural net output following the procedure described before, produces a noticeable improvement in the global performance of tagging, as can be seen from the "Combined" line of Table 2 and from the indicated plot, where a *plateau* region starts at about 5, which is the number of *tagging categories* eventually chosen. The performance is therefore determined in each category, and the total effective efficiency is the sum of the effective efficiencies determined in each category separately. These five categories correspond to the five arrows drawn on the left plot of figure 9. Events falling below 0.55 (first shorter arrow) are untagged, as said. The two visible spikes in the distributions are due to the discrete probability values assigned to the Q_{vtx} tagger, in the cases where all tracks used have the same charge. The obvious drawback is that one has to perform CP asymmetry measurements in each of these separate tagging categories independently. One other possibility is to use directly the fitted dependence of the neural net output as an event-by-event ω, with an uncertainty given by the fit itself, as it will be further discussed in Section 6.4.

Table 2 summarizes the results obtained for the $B_s^0 \to D_s^- \pi^+$ channel (52k triggered events), for both the individual taggers inclusively and for the five exclusive tagging categories. The combined value at the bottom can exceed the direct sum of the individual tagger performances. More results on other channels are given in the next Section.

	$\varepsilon_{\mathrm{eff}}$ %	$\varepsilon_{\mathrm{tag}}$ %	ω %
μ	1.76±0.11	11.53±0.14	30.5±0.6
e	0.55±0.06	4.10±0.09	31.7±1.0
K_{opp}	2.38±0.13	30.82±0.20	36.1±0.4
K_{same}	3.26±0.15	30.63±0.20	33.7±0.4
Q_{vtx}	1.34±0.10	23.97±0.19	38.2±0.4
cat#1	0.74±0.07	22.38±0.18	40.9±0.5
cat#2	1.67±0.11	15.95±0.16	33.8±0.5
cat#3	2.05±0.12	10.26±0.13	27.7±0.6
cat#4	1.95±0.11	6.05±0.10	21.6±0.7
cat#5	3.26±0.12	6.22±0.11	13.8±0.6
Average	7.77±0.34	60.86±0.21	32.13±0.26
Combined	9.67 ± 0.35	60.86 ± 0.21	30.07 ± 0.26

Table 2: Results of flavour tagging obtained for $B_s^0 \to D_s^- \pi^+$ events passing Level-0, Level-1 and HLT trigger selections (52k events), for both the individual taggers and for the five tagging categories. Uncertainties are statistical.

5 Results

Results for tagging efficiencies, wrong tag probabilities and effective efficiencies are shown in Table 3 for the indicated signal samples passing the Level-0, Level-1, and HLT triggers. The performance is shown for the combined tagging decision.

Channel	ε_tag (%)	ω (%)	ε_eff (%)
$B^0 \to \pi^+\pi^-$	52.48 ± 0.19	34.09 ± 0.25	5.31 ± 0.23
$B^0 \to K^+\pi^-$	53.24 ± 0.17	33.50 ± 0.22	5.79 ± 0.22
$B^0 \to J/\psi(\mu\mu)K^0_S$	52.95 ± 0.08	35.37 ± 0.10	4.53 ± 0.09
$B^0 \to D^{*-}\mu^+\nu_\mu$	56.22 ± 0.50	35.35 ± 0.63	4.82 ± 0.60
$B^+ \to J/\psi(\mu\mu)K^+$	54.05 ± 0.11	35.31 ± 0.15	4.66 ± 0.13
$B^0_s \to D^\mp_s K^\pm$	61.11 ± 0.34	30.02 ± 0.42	9.76 ± 0.56
$B^0_s \to D^-_s \pi^+$	60.86 ± 0.21	30.07 ± 0.26	9.67 ± 0.35
$B^0_s \to J/\psi(\mu\mu)\phi$	54.46 ± 0.16	31.97 ± 0.21	7.08 ± 0.23

Table 3: Performance of the combined tag for different signal decays passing trigger and offline cuts. Uncertainties are statistical.

Table 4 shows the details of how the application of the Level-1 trigger improves the tagging performances on signal events.

	Level-0			Level-0 and Level-1		
	ε_eff(%)	ε_tag(%)	ω(%)	ε_eff(%)	ε_tag(%)	ω(%)
μ	1.48 ± 0.08	9.96 ± 0.10	30.7 ± 0.5	1.67 ± 0.09	11.26 ± 0.12	30.8 ± 0.5
e	0.52 ± 0.04	3.81 ± 0.06	31.6 ± 0.8	0.58 ± 0.05	4.16 ± 0.08	31.3 ± 0.9
K_opp	2.02 ± 0.09	29.31 ± 0.15	36.9 ± 0.3	2.28 ± 0.11	31.06 ± 0.17	36.5 ± 0.3
K_same	3.30 ± 0.11	30.58 ± 0.15	33.6 ± 0.3	3.23 ± 0.13	30.53 ± 0.17	33.7 ± 0.3
Q_vtx	1.16 ± 0.07	22.98 ± 0.14	38.7 ± 0.3	1.33 ± 0.08	24.53 ± 0.16	38.4 ± 0.4
Combined	8.99 ± 0.25	59.21 ± 0.16	30.52 ± 0.20	9.48 ± 0.30	60.99 ± 0.18	30.28 ± 0.23

Table 4: Tagging performance for $B^0_s \to D^-_s \pi^+$ signal events after Level-0 (left) and after Level-0 and Level-1 triggers (right). Uncertainties are statistical.

6 Systematic Effects

The precise knowledge of ω and related systematic effects cannot rely only on Monte Carlo but need to be evaluated on data themselves to avoid all the uncertainties due to the modeling, as for example the b-production mechanism, the details of trigger and reconstruction, and the possible dependence of the tagging on the initial flavour of the B, because of the different interactions with matter of its decay products. Tagging algorithms need therefore to be calibrated by using flavour specific channels. Possible control channels, with their annual yield corresponding to $2\,\mathrm{fb}^{-1}$ integrated luminosity are indicated in Table 5 [4,7–11]. The neutral B decays generally need a proper time dependent fit to extract

the mistag fraction ω, since neutral B can oscillate to their flavour conjugate. This is not the case if a double tagging technique is used instead [12], but in that case, only the same-side mistag rate can be measured.

Channel	$2\,\mathrm{fb}^{-1}$ yield (k)	B/S
$B^+ \to J/\psi K^+$	1740	0.4
$B^+ \to \overline{D}^0 \pi^+$	1000	0.1
$B^+ \to \overline{D}^{(*)0} \mu^+ \nu_\mu$	2400	0.7
$B^0 \to J/\psi K^{*0}$	1017	0.16
$B^0 \to K^+ \pi^-$	135	0.16
$B^0 \to D^{*-} \mu^+ \nu_\mu$	9000	0.26
$B^0_s \to D^-_s \pi^+$	179	0.4
$B^0_s \to D^{(*)-}_s \mu^+ \nu_\mu$	1930	0.36

Table 5: For some control channels, expected number of events after Level-0 and Level-1 trigger and offline selection, and B/S estimated from inclusive b$\bar{\mathrm{b}}$ sample.

In this Section we consider some of the possible issues involved in the measurement of the mistag fractions in the control channels, and outline a general approach on how to deal with them.
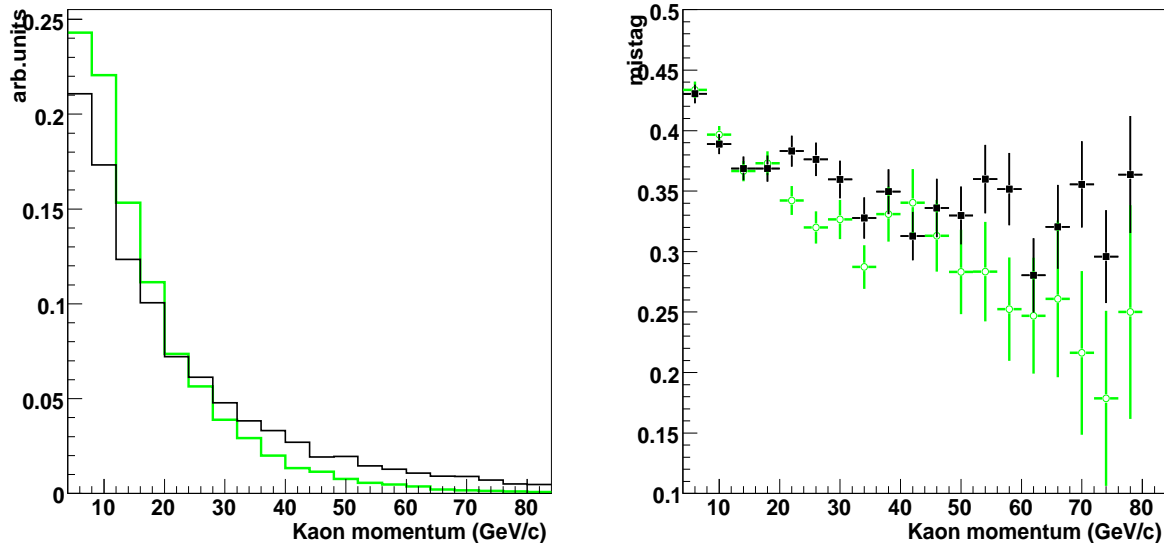

Figure 10: Left: distribution of the momentum of kaon opposite side candidate (black) and kaon same side candidate (green). Right: wrong tag fraction as a function of momentum in the two cases.

6.1 Dependence on flavour

The differences in K^-/K^+ interactions in the material of the detector and kaon production in secondary interactions have a non negligible effect on tagging. The cross-section of

	anti-$\bar{B}_s^0$ events only			B_s^0 events only		
	$\varepsilon_{\mathrm{eff}}(\%)$	$\varepsilon_{\mathrm{tag}}(\%)$	$\omega(\%)$	$\varepsilon_{\mathrm{eff}}(\%)$	$\varepsilon_{\mathrm{tag}}(\%)$	$\omega(\%)$
μ	1.57±0.13	11.31±0.17	31.4±0.7	1.77±0.13	11.22±0.17	30.1±0.7
e	0.56±0.08	4.23±0.11	31.8±1.2	0.61±0.08	4.09±0.11	30.7±1.2
K_{opp}	2.44±0.16	31.33±0.25	36.0±0.5	2.11±0.15	30.79±0.25	36.9±0.5
K_{same}	2.90±0.17	30.49±0.24	34.6±0.5	3.58±0.19	30.57±0.25	32.9±0.5
Q_{vtx}	1.55±0.13	24.72±0.23	37.5±0.5	1.12±0.11	24.34±0.23	39.3±0.5
Combined	9.54±0.42	61.15±0.26	30.25±0.32	9.45±0.42	60.82±0.26	30.29±0.32

Table 6: Tagging performance for $B_s^0 \to D_s^- \pi^+$ signal events after Level-0 and Level-1 triggers for B-decays with different initial flavour. Uncertainties are statistical.

kaons with ordinary matter (p,n,d) is ~20% higher for K^- than for K^+ [6]. Table 6 shows the tagging performances for the $B_s^0 \to D_s^- \pi^+$ channel after Level-0 and Level-1 triggers in the two cases when the true initial flavour is a b ($\bar{B}_s^0$) or a $\bar{b}$ (B_s^0). For the kaon taggers there is a significant change in the wrong tag fraction, particularly for the same side tagger. The $\bar{B}_s^0$-mesons same-side tagging particle is a K^- which interacts more with matter, and the tagging effective efficiency is 2.9%, while for B_s^0-mesons it is 3.6%. On the other hand, for $\bar{B}_s^0$-mesons the effective efficiency of the opposite-side kaon tagger (which is a K^+) is 2.4%, which drops to 2.1% for B_s^0-mesons. The relative difference in the tagging performance between same and opposite kaon taggers is due to their different momentum spectra, because soft kaons tend to interact more with matter. Figure 10 shows that the spectrum for same-side kaons is sensibly softer with respect to opposite-side kaons. Also the dependence of the mistag on the kaon momentum is more pronounced in the case of same-side kaons.

For the other taggers, the effective efficiency remains the same within statistics in the two cases, except for Q_{vtx} tagger which can contain kaons from the opposite B. It is therefore necessary to treat the two cases separately in the estimation of the mistag fractions on CP and control channels.

	TOS events only			TIS events only		
	$\varepsilon_{\mathrm{eff}}(\%)$	$\varepsilon_{\mathrm{tag}}(\%)$	$\omega(\%)$	$\varepsilon_{\mathrm{eff}}(\%)$	$\varepsilon_{\mathrm{tag}}(\%)$	$\omega(\%)$
μ	0.20±0.07	1.32±0.09	30.4±3.2	5.81±0.51	35.95±0.55	29.9±0.9
e	0.46±0.10	2.09±0.11	26.4±2.4	0.85±0.20	7.94±0.31	33.7±1.9
K_{opp}	1.47±0.19	23.87±0.34	37.6±0.8	5.78±0.52	45.56±0.57	32.2±0.8
K_{same}	4.00±0.30	28.44±0.36	31.2±0.7	2.75±0.36	32.21±0.53	35.4±1.0
Q_{vtx}	1.12±0.16	17.29±0.30	37.3±0.9	1.98±0.31	37.53±0.55	38.5±0.9
Combined	7.74±0.57	50.30±0.40	30.39±0.53	16.86±1.20	81.00±0.45	27.19±0.57

Table 7: Tagging performance for $B_s^0 \to D_s^- \pi^+$ signal events after Level-0 and Level-1 triggers for TIS and TOS type events. Fraction of TOS events in the total sample is 31%, while for TIS events it is 14.8%. Uncertainties are statistical.
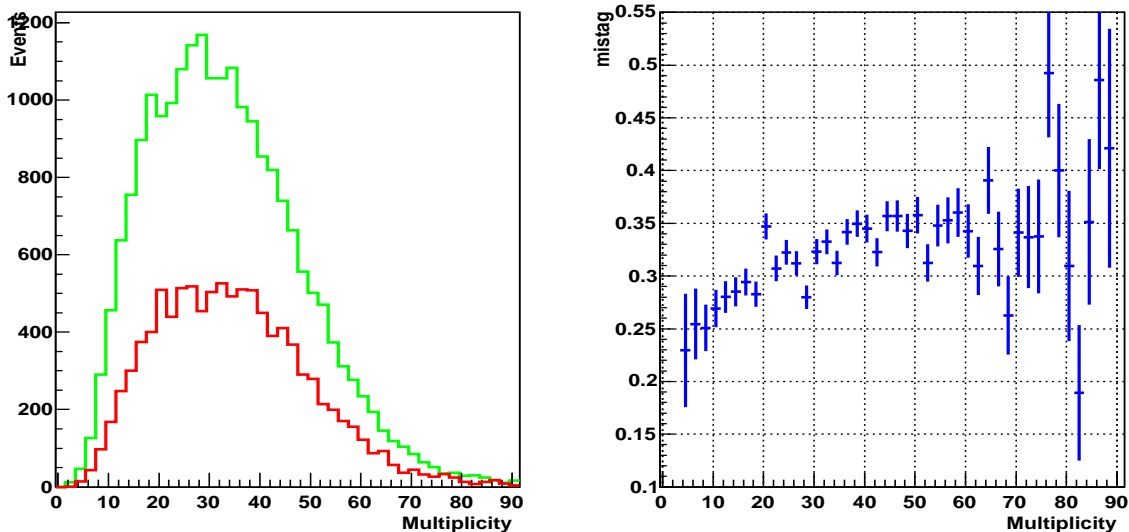
Figure 11: Left: Multiplicity of charged tracks in $B_s^0 \to D_s^- \pi^+$ events. Light (green) histogram is for correct tags, while dark (red) is for wrong tags. Right: dependence of the mistag fraction ω for the same events as function of the multiplicity.

6.2   Interplay with trigger

The mistag measured in the control channels cannot be blindly used for the CP channels, due to possible different bias introduced by trigger and offline selections [13]. The largest bias comes from the fact that these events are triggered in many different ways. Opposite side taggers are reliable if the opposite B-meson be in the LHCb geometrical acceptance, which is more likely to happen if the trigger has fired because of a track not belonging to the signal B (**T**rigger **I**ndependent of **S**ignal, TIS). On the other hand, if the event has been triggered because of a particle of the decay chain of the signal (**T**rigger **O**n **S**ignal, TOS), which is usually the case for decay modes with muons, then there is nothing *a priori* which grants for the opposite B-meson to be in the acceptance. From this consideration, and from other effects like possible bias on the B momentum induced by the trigger, it follows that decay modes with different TIS/TOS fractions will have different tagging performances.

For example in $B_s^0 \to D_s^- \pi^+$ events, the fraction of TIS events (at both Level-0 and Level-1 trigger) is 14.8%, and among these events, 89% have the opposite B in the acceptance (12 mrad$< \theta_B <$ 250 mrad). The fraction of TOS events is 30%. For $B^+ \to J/\psi K^+$ events instead, the TIS fraction is 7.5% (83% with the opposite B in the acceptance). The TOS component is 61% of the total sample were the fraction of opposite B in the acceptance is only 50%. Table 7 shows the tagging performance for $B_s^0 \to D_s^- \pi^+$ events after Level-0 and Level-1 when considering only-TIS and only-TOS type events. The remaining part of the events are combinations of Level-0 and Level1 TIS-TOS, TOS-TIS and events where both B need to be present for the event to be triggered.

17

6.3 Dependence on signal B momentum

Another effect to be considered is the fact the two B mesons in the proton-proton collision emerge with a strong angular correlation. The production angle is also correlated to the momentum (small angles correspond to higher momenta), which means that a difference in the signal B momentum distribution between two different channels, one of which used as control channel, can bias the opposite B momentum giving as a result a difference in the tagging performances. This is illustrated in Figure 12 which shows a comparison between the p_T of the B-meson in $B_s^0 \rightarrow D_s^- \pi^+$ and in $B^0 \rightarrow \pi^+ \pi^-$ events. The top plots correspond to the signal B, while the two bottom plots correspond to the opposite B which is never completely reconstructed and therefore its momentum is unknown. The ratio of the distributions on the right, shows how the opposite B distribution is influenced by the signal B following a similar pattern in $p_T(B)$.

Because of the presence of kinematic cuts on the taggers, the performance is dependent on the momenta. For example in $B_s^0 \rightarrow D_s^- \pi^+$ events when $p_T(B)< 8\,\mathrm{GeV}/c$ $\varepsilon_{\mathrm{eff}}(\mathrm{K_{same}})=1.8\pm0.2\%$, while for $p_T(B)> 8\,\mathrm{GeV}/c$ it becomes $\varepsilon_{\mathrm{eff}}(\mathrm{K_{same}})=4.6\pm0.3\%$. The ratio shown in the upper right plot can be used as a weight for the event to correct the bias.

6.4 Possible approaches to the mistag determination

The actual value of the mistag fraction ω in a CP channel and in the corresponding control channel is dependent on all the effects described above. The general idea is therefore to measure ω in different subsamples of the trigger (TIS/TOS splitting) also taking into account the differences in the mistag rate for different signal-B spectra and B-meson flavour.

This can be done in three different ways:

1. Choose a control channel which is topologically very similar to the CP channel considered. Examples are $B^0 \rightarrow J/\psi(\mu\mu)K^{*0}$ for $B^0 \rightarrow J/\psi K_S^0$ events, or $B_s^0 \rightarrow D_s^- \pi^+$ for $B_s^0 \rightarrow D_s^\mp K^\pm$. In this case the mistag measured in the control channel is directly applicable to the CP channel, while residual differences due to channel selection can be treated as systematics.

2. Evaluate the mistag in each *tagging category*, trigger sample, and in each B-signal phase space bins of the control channel.

3. Evaluate the mistag per event, as a function of the trigger sample, *the tagger* kinematic variables and the signal B phase space.

In the second case, one needs to calculate ω_{ijk} where $i = 1-5 = i$-th tagging category, $j = 1-5 = $ (TIS-TIS, TOS-TOS, TIS-TOS, TOS-TIS, Other) are all the possible combinations of Level-0 and Level-1 trigger subsamples and $j = 1-3 = $ bins of $p_T(B)$ [4].

This ω_{ijk}, measured in the control channel, can be used as the mistag of the CP channel in the same i, j, k bin. In Monte Carlo one can check that the correct ω_{ijk} is

---

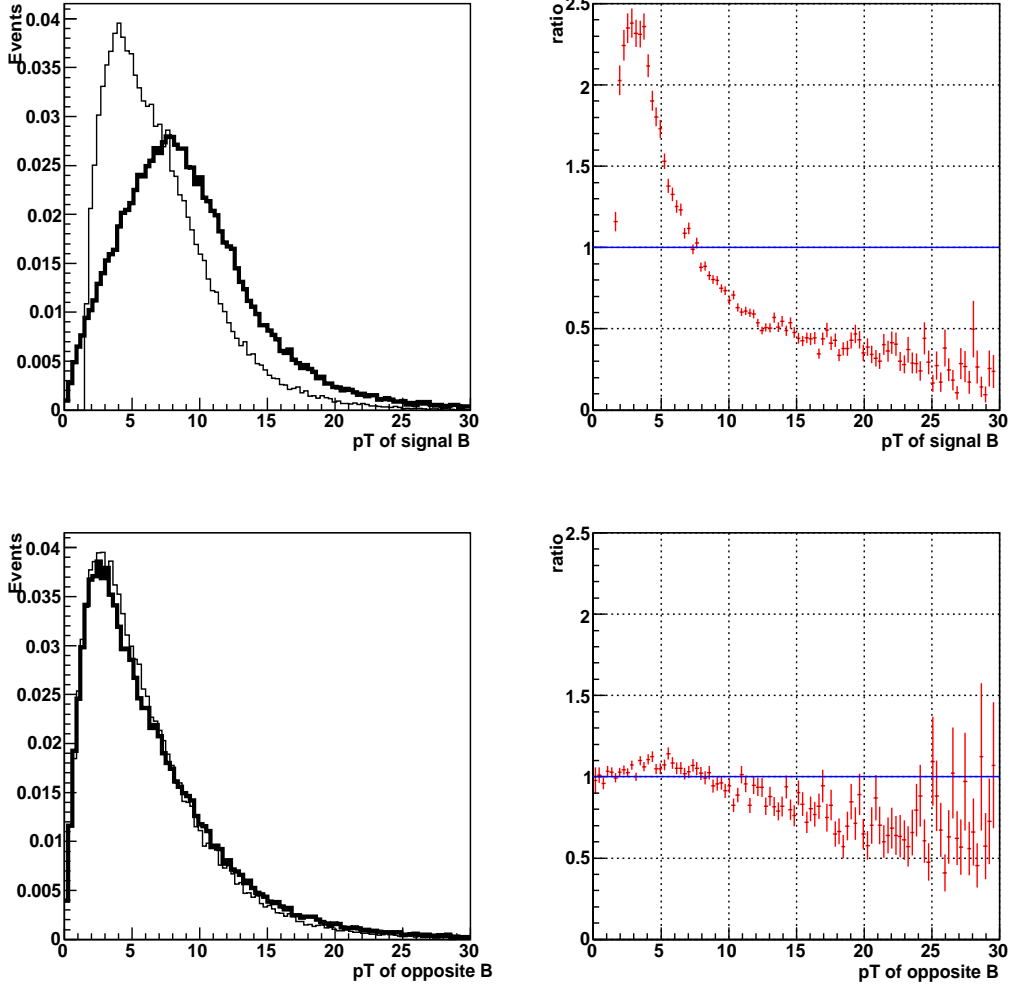[4]Three bins in $p_T(B)$ is a compromise between precision and statistics, no special study has been done yet.

Figure 12: Comparison between the B-meson p_T in $\mathrm{B_s^0 \to D_s^- \pi^+}$ (thicker line) and $\mathrm{B^0 \to \pi^+ \pi^-}$ events. On the right: ratio of the two distributions. For signal B, the reconstructed p_T is displayed, while for the opposite B, we plot the true p_T.

reproduced by looking at the Monte Carlo truth. In data, ω_{ijk} can be compared in two different control channels. In each tagging category of the considered CP channel one can estimate $N^{(\mathrm{flavour=b})} = \Sigma_{ij}(1 - \omega_{ij})N_{ij}^{(\mathrm{tag=b})}$ with $\omega_{ij} = 1/N^{\mathrm{CP}}\Sigma_k\omega_{ijk}N_k^{\mathrm{CP}}$ where N_k^{CP} is the number of events in the CP channel in the k-th bin of B momentum.

Concerning statistics available in control channels, if we take as an example the $\mathrm{B^+ \to J/\psi K^+}$ channel, which corresponds to 1.74M events in one year data taking, assuming to have 150 bins $= 5(\mathrm{tag}) \cdot 5(\mathrm{trigger}) \cdot 3(p_\mathrm{T}^B) \cdot 2(\mathrm{flavour})$ one ends up with about 350-38000 events per bin depending on the specific bin, which allow for a relative precision on ω_{ijk}, $\sigma(\omega_{ijk}) \approx 0.7 - 7\%$ with this channel alone. Background components need to be subtracted statistically from right-tag and wrong-tag data before calculating the mistag rate.

In the third case, which is actually a variant of the second, one can use the dependence of the *neural net* output from the kinematic properties of the tagger. As we have seen, a neural net is defined for each tagger, and its output is a function of the momentum, impact parameter of the track and also of the signal-B phase space.

Using the control channel one is able to construct distributions of the neural net output for the right and wrong-tagged events and their ratios to get the mistag $\omega = N_W/(N_R + N_W)$. This can be fitted as a function of the neural net output, as shown in the bottom plots of Figure 8 for the muon and kaon taggers, in each of the five mentioned trigger subsamples.

In this way we can extract from the control channel one function per tagger giving the mistag per event, as a function of the tagger $\omega(\Phi_i, \Phi_j)$ where $i = $ e, μ, K_{opp}, Q_{vtx}, $K/\pi_{\mathrm{same}} = 5$ taggers (not exclusive), $j = $ trigger subsample, $\Phi_K = (p^K, p_{\mathrm{T}}^K, IP/\sigma_{IP}, ...)$ the phase space of the tagger and $\Phi_B = (p^B, p_{\mathrm{T}}^B, ...)$ the signal-B phase space. If more than one tagger per event is available, the ω_i are combined into a global mistag per event using probabilities as described in Section 4.

Once it is verified that on data events the function $\omega(\Phi_i, \Phi_j)$ is the same through various control channels, then ω can be exported to the CP channel under study. It must be pointed out that there is also the possibility to simplify the whole procedure by merging some of the i, j bins where the ω fractions are comparable.

7   Conclusions

Results on flavour tagging performances in LHCb have been presented, using same side and opposite side algorithms. The effective efficiency, which is the figure of merit of the tagging power, range from about 4.5% in B^0 and B^+ decay modes, up to about 9.7% in B_s^0 modes where also the kaon same side tagger is available.

A number of effects on tagging performances have been studied involving the trigger and other possible biases due to selection and interaction with matter. Finally a general procedure on how to deal with control channels to determine the mistag fraction directly from data has been outlined.

References

[1] M. Calvi *et al.*, "LHCb Flavour Tagging Performance", LHCb 2003-115.

[2] T. Sjöstrand *et al.*, Computer Physics Commun. **135** (2001) 238.

[3] For the event reconstruction and analysis Brunel v23r7, DaVinci v12r18 and FlavourTagging v6r7 have been used.

[4] LHCb Collaboration "LHCb Reoptimized Detector Technical Design and Performance", CERN/LHCC 2003-030, 9 September 2003.

[5] ALEPH Collaboration, Phys.Lett B 425 (1998) 215. OPAL Collaboration, Z.Phys C 66 (1995) 19. DELPHI Collaboration, Phys. Lett. B345 (1995) 598.
CDF Collaboration, Phys Rev D 64 (2001) 072002 and CDF Note 7938 13/10/2005.

[6] W.-M. Yao *et al.*,"Review of Particle Physics" J. Phys. G 33, 1 (2006).

[7] E. Aslanides *et al.*, "Selection of $B^+ \to \overline{D}^0\pi^+$ and $B^+ \to \overline{D}^0\mu^+\nu_\mu$ to control flavour tagging in LHCb", LHCb note 2006-058.

[8] J. Borel *et al.*, "$B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ selections" LHCb-2007-017.

[9] O. Leroy *et al.*, "Selection of $B_s^0 \to D_s^-\mu^+\nu_\mu$ events in LHCb", LHCb note 2007-029.

[10] M. Calvi, "Selection of $B^0 \to D^{*-}\mu^+\nu_\mu$ events and flavour tagging studies", LHCb note 2007-036.

[11] S. Amato *et al.*, "Update to the LHCb sensitivity to $\sin 2\beta$ from $B^0 \to J/\psi(\mu\mu)K_S^0$ CP-asymmetry in $B^0 \to J/\psi(\mu\mu)K_S^0$ decays", LHCb note 2007-045.

[12] O. Awunor, "Double tagging technique in LHCb", LHCb note in preparation.

[13] H. Dijkstra *et al.*, "Some Remarks on Systematic Effects of the Trigger and Event Generator Studies", LHCb/2003–157.