Combination of same-side with opposite-side flavour tagging

Public Note

Issue: 1
Revision: 0

Reference: LHCb-PUB-2009-027
Created: January 20, 2010
Last modified: February 11, 2010

Prepared by: Miriam Calvo[a], Marc Grabalosa[b], Marco Musy[b]
[a]Enginyeria La Salle, Universitat Ramon Llull
[b]Universitat de Barcelona

LHCb-PUB-2009-027
11/02/2010

Abstract

A considerable number of CP violation measurements require the most possible accurate knowledge of the flavour at production of the reconstructed B meson. The performance of different flavour tagging methods will be measured from control channels. One of the possible tagging methods that can be used is the one known as *Same Side kaon*. In this note we present how to calibrate the response of the tagging algorithms using the data and combine the result with the so called *Opposite Side tagging*. Also trigger and selection effects are briefly discussed.

Contents

1 Introduction

Flavour tagging algorithms in LHCb are described in detail in [1]. In summary, for any B decay a set of *taggers*, i.e. generic tagging methods, can be used. The so-called Opposite Side (OS) tagging exploits the presence of a muon, electron, kaon or secondary vertex in the event coming from the opposite (non-signal) B. Same Side (SS) kaon tagging in the case of a B_s (or pion tagging for a B^0), exploits the presence of a particle from the fragmentation chain which is correlated in the phase-space with the signal decay itself.

The final decision for the flavour of the signal at production is taken according to the charge of the tagging particles. To improve the global performance of the tagging the sample is subdivided into a number of *tagging categories*. This splitting can be made following different criteria. One possibility is to sort the events based on the type of *tagger* (electron, μ or kaon particle) which contributes to form the tag decision (*PID approach*).

One other possibility is to consider a number of kinematic variables for each tagger and use them as the input to a Neural Net that estimates a probability of the tagger decision to be correct. These probabilities are then combined treating them as independent (*Neural Net approach*). This last way to proceed gives the advantage of a higher effective tagging efficiency[a] $\epsilon_{\rm eff}$. In the case of the Neural Net approach, which is the one considered in this note, the signal samples are sorted into 5 categories of decreasing mistag probability.

[a]The effective tagging efficiency is defined as $\epsilon_{\rm eff} = (1-2\omega)^2\epsilon_{\rm tag}$, where ω represents the fraction of wrongly tagged events and $\epsilon_{\rm tag}$ the tagging efficiency or fraction of tagged events.

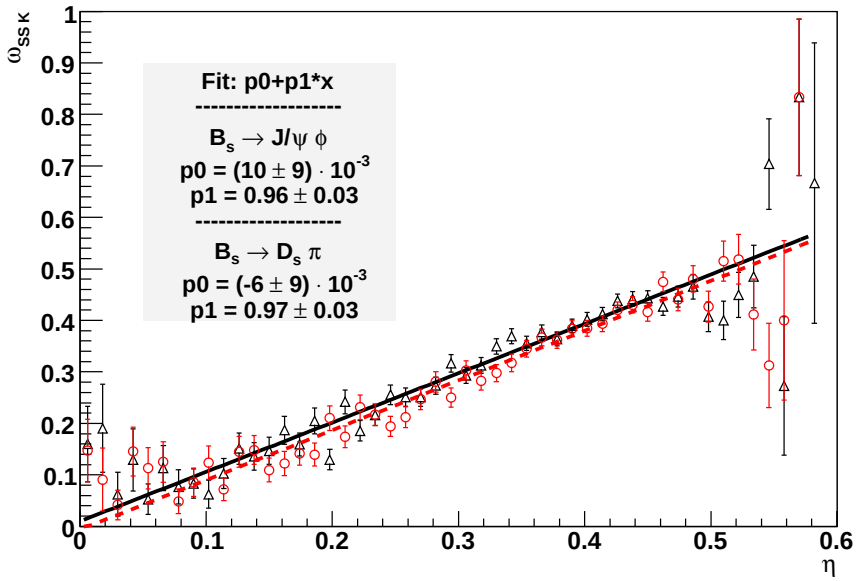

Figure 1 Same Side kaon true mistag rate vs Neural Net output probability η for $B_s \to J/\psi\phi$ (black triangles, solid fit) and $B_s \to D_s\pi$ (red circles, dashed fit) offline selected events after the first-level (L0) trigger.

Such probability per tagger and per event, obtained from the Neural Net output η, needs to be calibrated from data to get the real mistag or wrong tag fraction ω, to avoid relying on Monte Carlo (MC) predictions. For each tagger the dependence of this probability on η can be parametrized as $\omega = p_0 + p_1 \cdot \eta$. The parameters p_0 and p_1 can be fitted in control channels, where the tag of the B meson is known from the charge of the particles identified in the final state, so that it can be compared with the tagging decision.

The same side tagging almost doubles the global tagging power, so that it would be used in all analyses that need to tag the flavour of a B_s meson at production (in the case of B_d and B_u the same side tag will not be so effective, due to the presence of a pion instead of a kaon from the fragmentation chain). The procedure that will be discussed in this note is rather general and the same principle can be applied to different decay modes. In what follows we will refer to the specific case of the measurement of the B_s mixing phase β_s through the decays of $B_s \to J/\psi\phi$, as this mode has been studied in detail [2].

As control channels for this decay, for what concerns the OS taggers, it is planned to use the $B^+ \to J/\psi K^+$and $B^0 \to J/\psi K^*$ channels, due to their similar topology to $B_s \to J/\psi\phi$ and their high statistics [3]. For SS kaons, a B_s decay has to be used as control channel to perform the calibration. The $B_s \to D_s\mu\nu_\mu$ mode would be a good candidate to measure ω of the SS kaon tagger, as its annual yield is estimated to be approximately 1 million events for an integrated luminosity of 2 fb^{-1} [4]. However, due to the presence of the invisible neutrino, it is not straightforward to extract the mistag from the B_s oscillations, as the reconstruction of the decay requires a good proper time resolution. More studies will be needed to assess the feasibility of a measurement of the mistag with this channel. For $B_s \to D_s\mu\nu^\mu$ events a *double tagging* method can be used to obtain a global $\omega_{\rm SS}$ for each of the 5 tagging categories. In this case one considers events where both OS and SS taggers are present: from the knowledge of the OS tagging performance is it then possible to extract the SS mistag (see [5] for details).

For $B_s \to J/\psi\phi$ we expect to obtain $\epsilon_{\rm eff}$ = 3.4% when only using OS taggers (with the Neural Net approach, sorting events into 5 categories) . Including the SS kaon tagger as well in the analysis, the effective efficiency increases up to 6.4%. Therefore there is a strong physics motivation to combine the

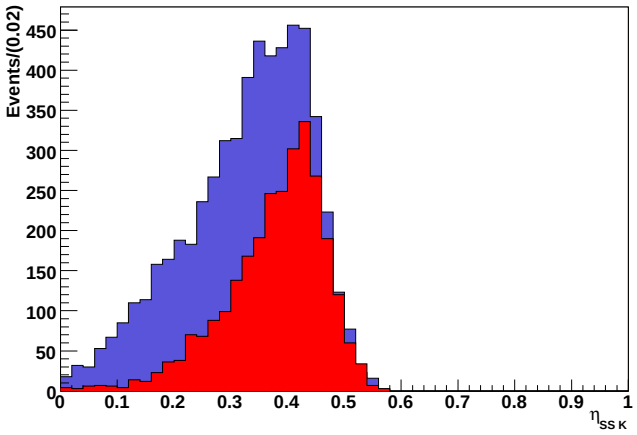

Figure 2 SS kaon η distribution for right tag (darker blue histogram) and wrong tag (lighter red histogram) L0 triggered and offline selected $B_s \to D_s\pi$ events.

decision of SS with OS taggers as soon as a calibration procedure for the SS tagger is available.

This note is organized as follows. Section 2 explains how to calibrate the SS kaon tagger with the data using the $B_s \to D_s\pi$ mode. Section 3 discusses the combination of the individual tagger responses, calibrated on control samples. Section 4 shows briefly how to deal with selection and trigger effects. Finally, conclusions are given in Section 5.

2 Same Side kaon calibration with $B_s \to D_s\pi$ events

As already introduced, we plan to use $B_s \to D_s\pi$ events to calibrate the Neural Net output for the SS kaon tagger, fitting the Neural Net response with a simple linear polynomial $\omega = p_0 + p_1 \cdot \eta$. As B_s mesons mix, the free parameters p_0 and p_1 need to be extracted in a simultaneous fit to the oscillations of the B_s, taking also into account the presence of physics backgrounds.
The detailed procedure is explained later in this section.

A required condition is that the estimation of ω as a function of the response of the Neural Net has to be the same for the CP signal and the control channel, so that a calibration using $B_s \to D_s\pi$ events can be used for the $B_s \to J/\psi\phi$ events. Figure 1 shows the result of this proof-of-principle check (which is possible by accessing the MC truth information[b]). The $B_s \to J/\psi\phi$ channel is indicated by black triangles while $B_s \to D_s\pi$ by red circles. For the $B_s \to J/\psi\phi$ channel a "lifetime unbiased" event selection has been used [6], starting with 92k fully simulated and reconstructed Monte Carlo events. For the $B_s \to D_s\pi$ channel the available data sample consisted of 63k events. The results of the linear fit in the Figure shows that the SS kaon tagging calibrations are compatible for the two channels. All selected events are after the first-level (L0) trigger [7].

The offline selection of $B_s \to D_s\pi$ events is described in [8]. The expected annual yield is around 175k events after the L0 trigger and the background over signal ratio, B/S, is 0.4. The background composition is mainly coming from 2 sources: from combinatorial background (50%), evaluated on an inclusive $b\bar{b}$ sample, and from the specific decay mode $B^0 \to D\pi$ (35%). Other sources of background (15%) include $B_s \to D_s\rho$, $\Lambda_b \to \Lambda_c\pi$ and $\Lambda_b \to D_s\pi$ events.

Tables 7 and 8 summarize the tagging performance of the two channels $B_s \to J/\psi\phi$ and $B_s \to D_s\pi$ respectively after the L0 trigger. They are subdivided into three sections. The first block corresponds to the performance of each individual tagger method, regardless on the number of active taggers

[b] All results shown in this note correspond to the so-called "Data Challenge 2006" Monte Carlo samples.

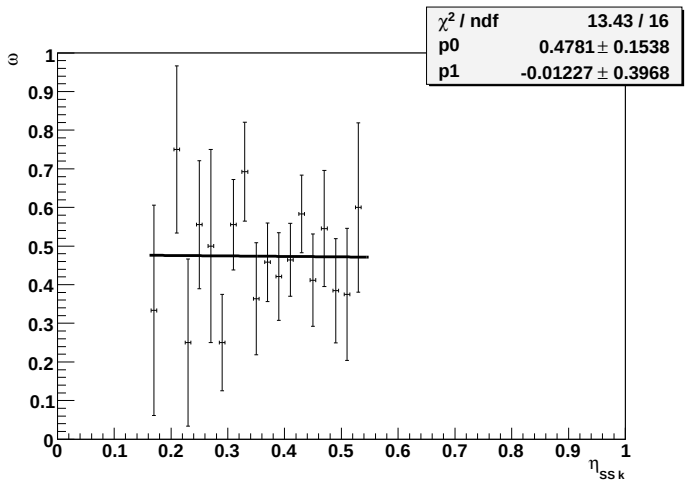

Figure 3 True mistag vs η for SS kaons in events that are background for the $B_s \to D_s\pi$ channel. A simple linear fit is superposed.

in a given event. In the second block, the information from the tagger responses are combined to form a final flavour decision. The events are sorted into five bins of decreasing wrong tag fraction. The effective efficiencies in each of these bins can be summed up directly and corresponds to the global tagging power achievable, indicated by the "combined" label. Its value corresponds to 6.46% for $B_s \to J/\psi\phi$ and 8.67% in the case of $B_s \to D_s\pi$ events. The third block is the same as above with the exclusion of the SS kaon tagger. The label "average" refers to the results where all events are put together, as if only one tagging category existed. The main differences between the two modes come from the effect of the trigger and from the different signal spectra as it will be discussed in the last section.

The response from the Neural Net for the SS kaon tagger representing the raw estimate for the probability of mistag has the distribution shown in Figure 2, for right and wrong tags. The true mistag has to be measured along with the oscillation amplitude and can not be extracted as easily as in the $B^+ \to J/\psi K^+$ case. In order to do so, data are fitted to extract p_0 and p_1 with the help of the RooFit toolkit [9], including both signal and backgrounds. In the signal component the wrong tag fraction is about 33%, while in the background component it is approximately 50%, as can be seen from Figure 3.

An unbinned fit was performed with a $B_s \to D_s\pi$ data sample where the background has been generated, due to the lack of statistics, with parameters extracted from the available background samples ($b\bar{b}$, $B_d \to D\pi$, $\Lambda_b \to D_s p$ and $\Lambda_b \to \Lambda_c\pi$ events).

The probability density function (PDF) consists of the exponential decay time distribution of the signal B_s, a double gaussian for the mass dependence, a term for the decay time acceptance and a PDF for the expected eta distribution (see Figure 2). The explicit expression of the final PDF is given by:

$$PDF_{sig}(m) = f_m \cdot G(m; M_B, \sigma_{m1}) + (1 - f_m) \cdot G(m; M_B, \sigma_{m2})$$

$$PDF_{sig}(t, \eta, \mathrm{q}) = A(t) \cdot \frac{1}{2\tau_{B_s}} \cdot exp(-\Gamma_s t') \cdot [\cosh(\frac{\Delta\Gamma_s}{2} t') + \mathrm{q} \cdot (1 - 2 \cdot (p_0 + p_1 \cdot \eta)) \cdot cos(\Delta m_s \cdot t')] \otimes G(t - t'; \sigma_{t(sig)})$$

$$PDF^{tot}_{sig}(m, t, \eta, \mathrm{q}) = PDF_{sig}(m) \cdot PDF_{sig}(t, \eta, \mathrm{q}) \cdot PDF(\eta)$$

where q is +1 for unmixed B_s and -1 if mixed; the proper time acceptance is taken as $A(t) = \frac{a \cdot t^3}{1 + a \cdot t^3}$ while $\Delta\Gamma_s = 0.06852\ \mathrm{ps}^{-1}$, $\Delta m_s = 20\ \mathrm{ps}^{-1}$, and $\tau_B = 1.461\ \mathrm{ps}$ are kept fixed in the fits.

For background parametrization, an exponential term for the mass, acceptance for proper time and the observed PDF for η are considered:

$$PDF_{bkg}(m) = exp(m; -\alpha)$$

$$PDF_{bkg}(t, \mathrm{q}) = A(t) \cdot \frac{1}{2\tau_{bkg}} \cdot exp(t'; \tau_{bkg}) \cdot [1 + \mathrm{q} \cdot (1 - 2\omega_{bkg})] \otimes G(t - t'; \sigma_{t(bkg)})$$

$$PDF^{tot}_{bkg}(m, t, \eta, \mathrm{q}) = PDF_{bkg}(m) \cdot PDF_{bkg}(t, \mathrm{q}) \cdot PDF(\eta)$$

The total PDF for signal and background is then:

$$PDF^{tot}(m, t, \eta, \mathrm{q}) = f_{sig} \cdot PDF^{tot}_{sig}(m, t, \eta, \mathrm{q}) + (1 - f_{sig}) \cdot PDF^{tot}_{bkg}(m, t, \eta, \mathrm{q})$$

where f_{sig} is the fraction of signal events.

Figure 4 shows the fit to the B_s mass distribution (signal + background), Figure 5 shows the fit to the proper time of unmixed and mixed events, and Figure 6 the corresponding asymmetry. In the top plot of Figure 7 shows the same oscillation plot in the region of proper time up to 5 ps. It can be noticed that the amplitude of the oscillations is slightly increasing with the proper time. This effect is due to the dilution of the background which has a stronger impact at the small proper times, as indicated by the bottom plot of Figure 7 from which the background component has been excluded.

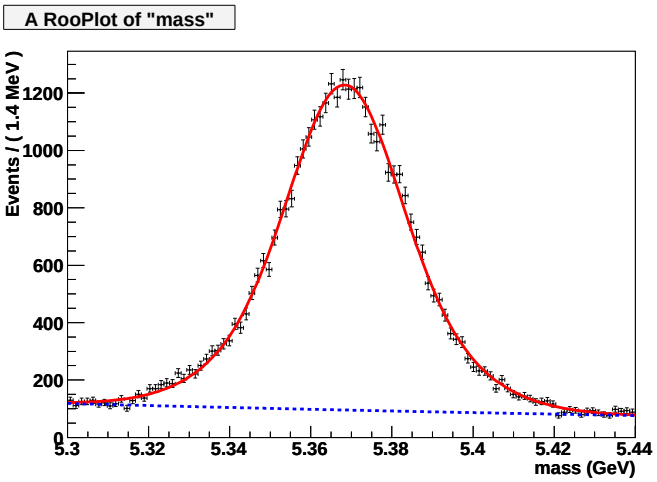

Figure 4 B_s mass distribution for SS kaon tagged $B_s \to D_s\pi$ events including background. The solid red line shows the result of the fit of the signal+background components, while the blue dotted line indicates the background component alone.

Results of the fitted parameters are shown in Table 1. The proper time resolution of signal has been fixed (38 fs) to properly determine the contribution of the mistag to the overall dilution. Around 37k events with a SS kaon tag have been used to perform the fit, which is equivalent to an integrated luminosity of 1.4 fb^{-1}. The tagging efficiency in this sample corresponds to $\epsilon_{\mathrm{eff}} = 26\%$. The values obtained for p_0 and p_1, which are compatible with the true ones, are $p_0 = -0.006 \pm 0.009$ and $p_1 = 0.97 \pm 0.03$, see Figure 1.

To assess the statistical sensitivity on the mistag parameters after one year of nominal LHCb data taking (2 fb^{-1}), a full toy study has been performed using 1348 toy Monte Carlo samples, each corresponding to 50k events. The resulting uncertainties correspond to ± 0.0028 on parameter p_0, and ± 0.016 on p_1. The first two plots of Figure 8 show the pull distributions for the fitted parameters p_0 and p_1.

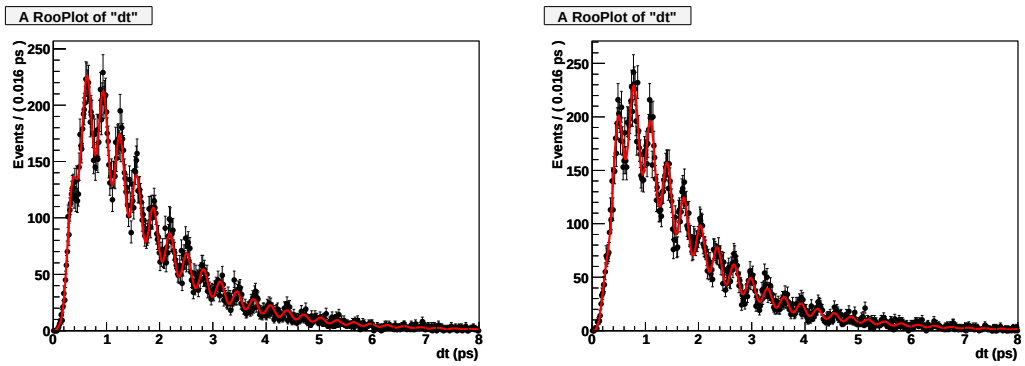

Figure 5 B_s proper time distribution for SS kaon tagged $B_s \to D_s\pi$ events including background. Unmixed events (left plot) and mixed events (right plot).

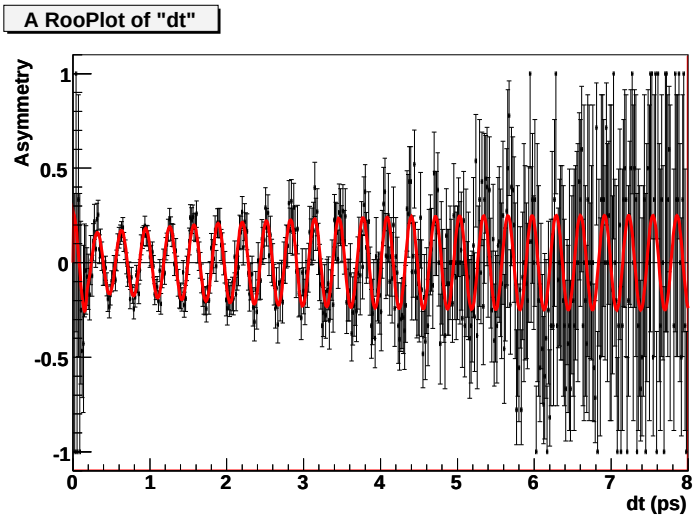

Figure 6 The B_s proper time asymmetry, $\frac{N_{unmixed}-N_{mixed}}{N_{unmixed}+N_{mixed}}$, for SS kaon tagged $B_s \to D_s\pi$ events including background in the range of the proper time up to 8 ps.

The same procedure can be used to evaluate the sensitivity to the global mistag rate $\omega_{\rm SS}$, which corresponds to ± 0.005 (i.e. 1.6% relative error). This estimate for the error has been obtained from 920 toy experiments, including background. The third plot of Figure 8 shows the pull distribution for $\omega_{\rm SS}$.

Future studies will be needed to evaluate the possible systematics associated to the fit.

3 Combination of the taggers' decisions

The combination of SS kaon tagger with OS taggers can increase $\epsilon_{\rm eff}$ by almost a factor 2. This can be seen in Tables 7 and 8, where the inclusion of the SS kaon tagger augments the combined effective tagging efficiency from 3.4% to 6.46% in the case of $B_s \to J/\psi\phi$, and from 4.53% to 8.67% in the case of the $B_s \to D_s\pi$ mode. Since the SS kaon tagger has a higher purity with respect to the other taggers, it tends to contribute to the tagging categories from 2 to 5, i.e. the ones of lower mistag.

The function $\omega(\eta)$ for OS taggers can be calibrated directly with $B^+ \to J/\psi K^+$. We need therefore a general way to combine the probabilities of all taggers, OS and SS, to get a reliable event-per-event probability of mistag to apply in $B_s \to J/\psi\phi$ events. However, due to the correlations among taggers,

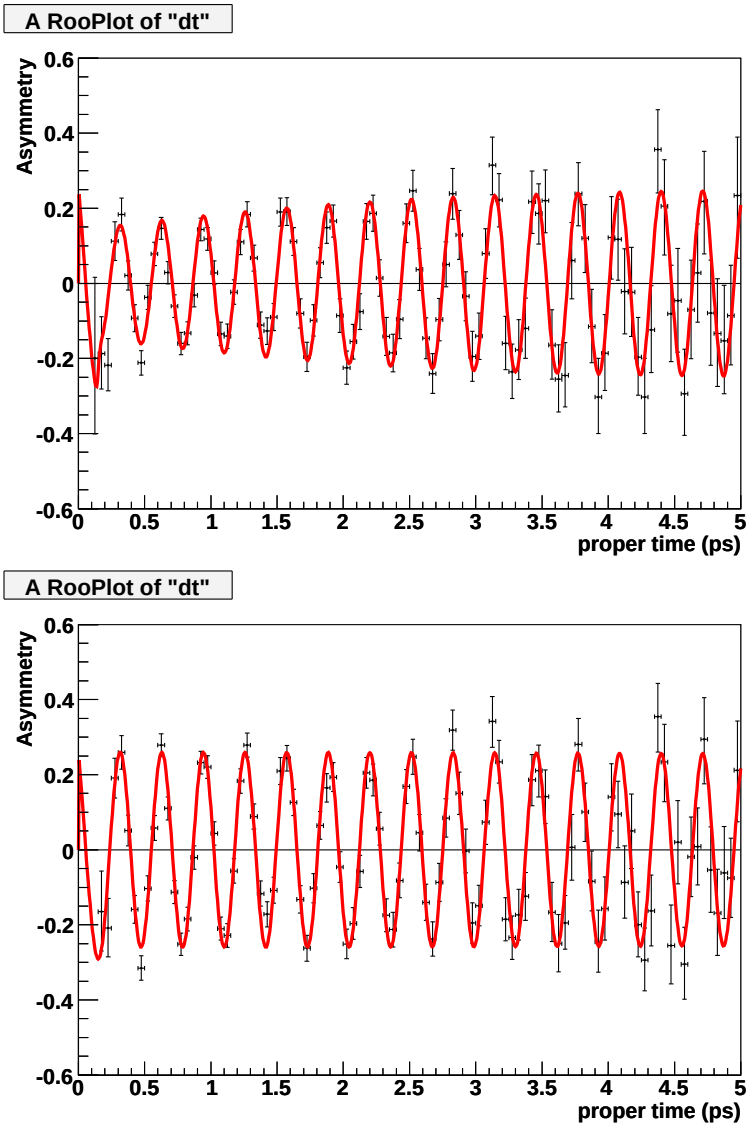

Figure 7 The B_s proper time asymmetry, $\frac{N_{unmixed}-N_{mixed}}{N_{unmixed}+N_{mixed}}$, for SS kaon tagged $B_s \to D_s\pi$ events including background in the range of the proper time up to 5 ps. The plot below corresponds to the case with no background.

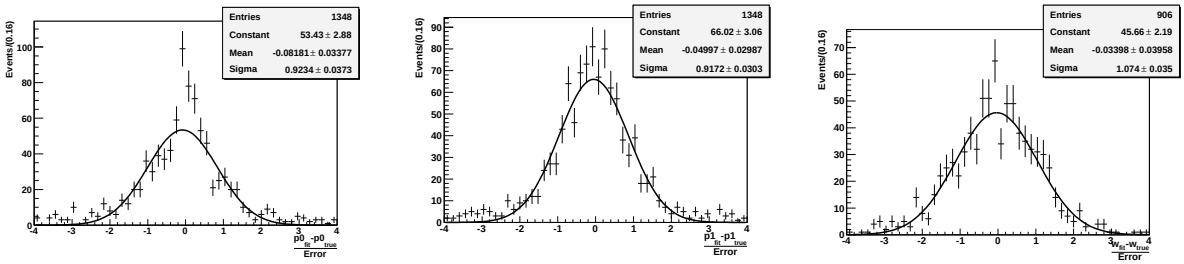

Figure 8 Pull distributions of the mistag parameters p_0, p_1 and $\omega_{\rm SS}$ respectively.

parameter	fitted value
B/S	0.306±0.005
f_m	0.379±0.010
M_B (GeV/c^2)	5.3683±8·10^{-5}
σ_{m1} (GeV/c^2)	0.0119±0.0001
σ_{m2} (GeV/c^2)	0.0204±0.0002
α (GeV/c^2)$^{-1}$	3.4±0.2
a (ps^{-3})	7.4±0.4
p_0	0.015±0.005
p_1	0.920±0.014
$\tau^{\rm bkg}$ (ps)	0.647±0.012
$\sigma_t^{\rm bkg}$ (ps)	0.10±0.07
$\omega^{\rm bkg}$	0.510±0.005

Table 1 Value of the fitted parameters on $B_s \to D_s\pi$ signal + background events.

this can not be done straightforwardly, as the real mistag would become underestimated. Figure 9 shows the deviation of the estimated mistag from the true mistag when combining OS taggers.

	μ	e	OS K	SS K	$Q_{\rm vtx}$
μ	1.00	0.01	0.03	0.02	0.10
e		1.00	0.03	0.01	0.07
OS K			1.00	0.03	0.19
SS K				1.00	0.02
$Q_{\rm vtx}$					1.00

Table 2 Correlations of tagger decisions for $B_s \to J/\psi\phi$ decays.

Table 2 shows the correlation in the decision of the taggers in $B_s \to J/\psi\phi$ events. The main source of correlation is the secondary vertex tagger, $Q_{\rm vtx}$, with the other OS taggers. Note also that correlations among OS taggers and SS kaon are much smaller.

A few strategies to correct or remove these correlation effects can be envisaged, from the most simple to the one which offers the highest $\epsilon_{\rm eff}$:

1. Remove the Secondary Vertex Charge tagger from the combination.
2. Measure *a posteriori* a correction to apply to the event-per-event ω from $B^+ \to J/\psi K^+$ events.
3. Measure $\omega_{\rm OS}$ in each one of the 5 categories from $B^0 \to J/\psi K^*$ events and combine with the SS kaon tagger using the fact that the correlation with OS is very small.
4. Use an additional Neural Net which takes as input the response of the five individual taggers.

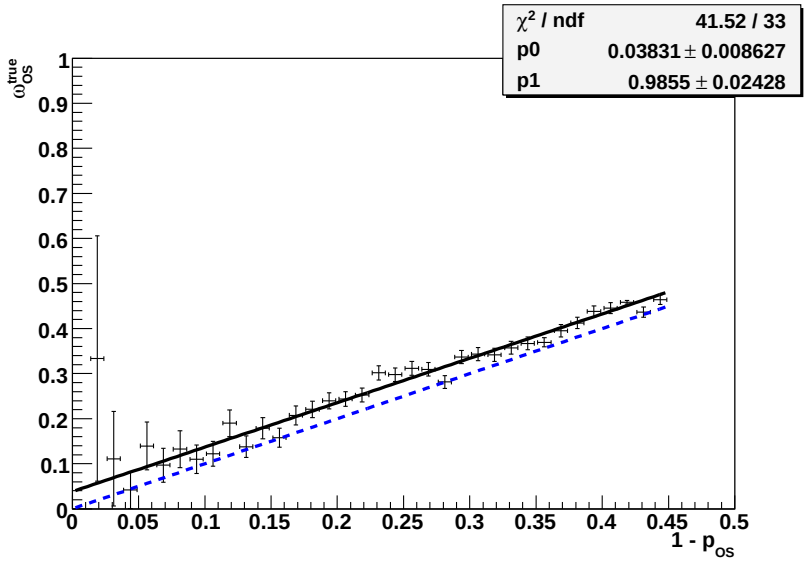

Figure 9 True wrong tag fraction of OS taggers versus the estimated probability to be wrong obtained as a product of individual OS taggers probabilities. For $B_s \to J/\psi\phi$ offline selected events after the L0 trigger. Full line corresponds to a linear fit and dashed line the expected behavior without correlations.

3.1 Excluding the Secondary Vertex Charge tagger

The effect of removing the Secondary Vertex Charge (SVC) tagger from the combination of taggers is shown in Figure 10, to be compared with Figure 9 where the tagger is activated. The total effective efficiency in this case is reduced from 6.46% to 6.05% for $B_s \to J/\psi\phi$ events.

$B_s \to J/\psi\phi$			
	$\epsilon_{\rm tag}$ (%)	ω (%)	$\epsilon_{\rm eff}$ (%)
cat 1	26.98±0.15	42.7±0.3	0.57±0.05
cat 2	12.19±0.11	34.3±0.4	1.21±0.07
cat 3	7.34 ±0.09	26.8±0.5	1.57±0.08
cat 4	3.95 ±0.06	21.0±0.7	1.32±0.07
cat 5	3.17 ±0.06	13.9±0.6	1.65±0.07
average	53.6±0.2	34.1±0.2	4.93±0.15
combined	53.6±0.2	32.8±0.2	6.33±0.15

Table 3 Tagging performance in $B_s \to J/\psi\phi$ decays after the L0, using $Q_{\rm vtx}$ only if it is the only available tagger or in presence of the SS tagger.

If we remove the SVC tagger when one of the other taggers is present, the effective efficiency decreases to 6.26%. Since the SS tagger response shows a negligible correlation with the SVC tagger, we can use SVC tagger also when SS tagger is present. The effective efficiency in this case goes from 6.26% to 6.33% in $B_s \to J/\psi\phi$ events. This can be done without introducing a bias in the correlation plot of Figure 10. Table 3 shows the tagging performances obtained in this case.

This option has the advantage of being very simple, with a modest reduction of the tagging performance with respect to the initial 6.46%. For its simplicity we plan to take this option as the baseline for tagging. As seen from Table 2 correlations, albeit small, still exist among the other OS taggers, so that also other possibilities are worth investigating.

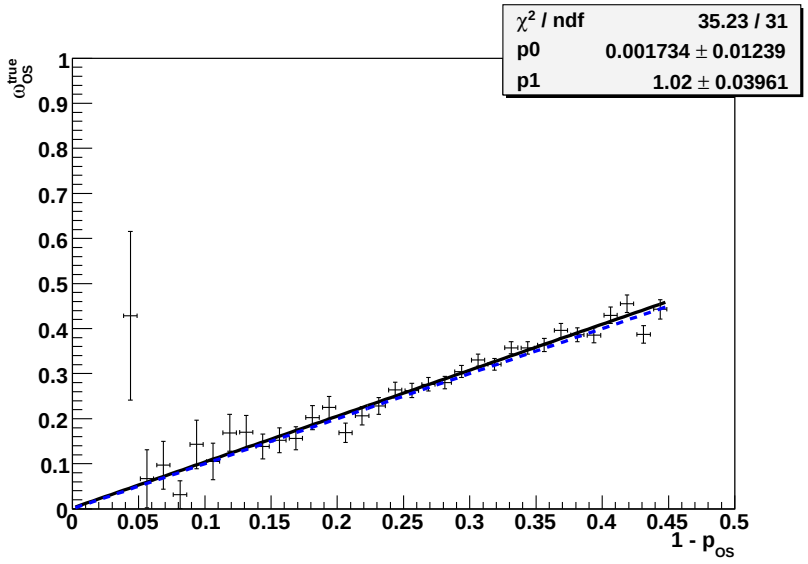

Figure 10 True wrong tag fraction of OS taggers versus the estimated probability to be wrong excluding the OS vertex charge tagger.

3.2 A posteriori correction of the OS mistag

One possible way to proceed is to use the observed bias in the determination of ω in a control channel to recorrect event-per-event *a posteriori* the mistag in the CP signal channel. In the measurement of the B_s mixing phase β_s [2] one could use $B^+ \to J/\psi K^+$ events as a control channel, and export such a correction to the CP signal channel (see top of Figure 11, where the difference $\omega_{\rm estimated} - \omega_{\rm true}$ is plotted as a function of $\omega_{\rm estimated}$). This is possible provided that the CP signal channel, in this case $B_s \to J/\psi\phi$, shows the same dependence on $\omega_{\rm estimated}$ (see bottom of Figure 11). This is indeed the case as the results of the fits are comparable for the two channels.
Table 4 shows the results for the tagging performances with the applied correction for $\omega_{\rm estimated} - \omega_{\rm true}$ evaluated on $B^+ \to J/\psi K^+$ events. The column labeled as "ω from $J/\psi K^+$" shows the comparison of the mistag rates in the five categories calculated using the correction evaluated with the control channel $B^+ \to J/\psi K^+$. All estimated mistags are compatible with the $\omega_{\rm true}$ values.

$B_s \to J/\psi\phi$				
	$\epsilon_{\rm tag}$ (%)	$\omega_{\rm true}$ (%)	ω from $J/\psi K^+$	$\epsilon_{\rm eff}$ (%)
cat 1	18.43±0.13	40.4±0.4	40.6±0.4	0.67±0.05
cat 2	9.90±0.09	33.3±0.5	33.1±0.5	1.01±0.06
cat 3	6.95±0.08	27.6±0.6	27.2±0.6	1.39±0.07
cat 4	4.04±0.06	20.4±0.7	21.3±0.7	1.41±0.07
cat 5	3.45±0.06	13.5±0.6	12.6±0.6	1.84±0.07
average	41.96±0.14	32.60±0.24	–	5.08±0.14
combined	41.96±0.14	30.57±0.24	–	6.34±0.15

Table 4 OS (corrected correlations) + SS flavour tagging performance in $B_s \to J/\psi\phi$ decays after the L0.

3.3 Use of $B^0 \to J/\psi K^*$ events to measure OS mistag

One other possible recipe is to first tune each individual tagger on a control channel, like $B^+ \to J/\psi K^+$, and then build an estimator of the combined probability treating the taggers as independent,

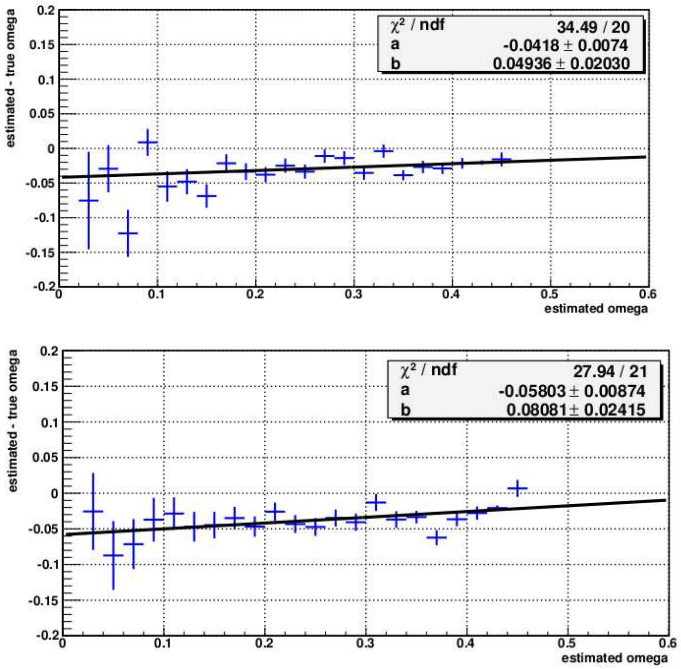

Figure 11 The difference $\omega_{\rm estimated} - \omega_{\rm true}$ as a function of $\omega_{\rm estimated}$ on $B^+ \to J/\psi K^+$ events (top) and on $B_s \to J/\psi\phi$ events (bottom) [2].

even if they are not, in order to sort them into the usual 5 tagging categories. At this point one can use the $B^0 \to J/\psi K^*$ control channel to measure the correct ω for the events in each of the 5 categories. In this way any possible existing correlation among the initial taggers becomes irrelevant, since one measures directly the combined OS mistag from the data in a control channel for a specific tagging category.

Once an unbiased mistag evaluation for the combination of the OS taggers is obtained, the OS mistag probability can be combined with the SS mistag probability, exploiting the fact that SS and OS taggers are uncorrelated. An event-per-event mistag probability can be obtained by means of the simple expression:

$$\omega_{\rm TOT}(\pm 1) = \omega_{\rm OS}(\pm 1) \cdot \omega_{\rm SS}(\pm 1)$$

$$\omega = \frac{max(\omega_{\rm TOT}(+1), \omega_{\rm TOT}(-1))}{\omega_{\rm TOT}(+1) + \omega_{\rm TOT}(-1)}$$

The final probability ω constitutes therefore an unbiased estimation of the mistag in the control channel which can be used for the CP channel. In [3] it has been shown that the results of the tagging algorithms are comparable within errors for the two channels $B^0 \to J/\psi K^*$ and $B^+ \to J/\psi K^+$, so that the final performance achievable with this method is equivalent to the method described in the previous section.

3.4 Use of a Neural Net for the combination of taggers

One valuable feature of Neural Nets is that they are insensitive to correlations in the input variables, so that if, for example, one of the tagger decisions were an exact duplication of another one (i.e. 100% correlated), its weight would be automatically set to zero by the Neural Net in the training phase.

The simple idea is therefore to feed the Neural Net inputs with the output ω_i from the individual tagger probability estimation. In the case that the tagger decision is $\bar{B}$, then $1 - \omega_i$ is given as input.

Treating B-tagged events as "signal" and $\bar{B}$-tagged events as "background" is a convenient way to exploit one of the multi-variate tools like TMVA [10]. Events which are tagged by one single tagger do not obviously need any special treatment, so the Neural Net is only trained and applied on events with multiple active taggers. As SS tagging is different for $B_{d,u}$ and B_s channels, two different trainings for the Neural Net combination would be needed in principle. As a matter of fact this turns out to be unnecessary, as applying the tuning of a B_s channel to a B_d channel shows the same tagging performance that one obtains with a dedicated tuning. This is not so surprising if one considers that in the case of the OS taggers, the correlations are similar in the B_s and $B_{d,u}$ modes, while for the OS versus SS they are anyway small. As a result the Neural Net is able to deal with both cases at once.

The results with this method are comparable with or even better than the two previous methods, (see Table 5) giving a $\epsilon_{\rm eff}$=(6.55 ± 0.15)% on $B_s \to J/\psi\phi$ events, and $\epsilon_{\rm eff}$=(8.80 ± 0.21)% for $B_s \to D_s\pi$. Other channels have been studied without changing the tuning of the Neural Net itself, obtaining similar or better tagging performances and successfully correcting the correlation issues. Figure 12 shows the true mistag as a function of the estimated mistag for the three channels $B_s \to J/\psi\phi$, $B_s \to D_s\pi$ and $B^+ \to J/\psi K^+$. This procedure allows to have an event-per-event correct estimation of ω which can then be used as input to the CP fits.

$B_s \to J/\psi\phi$			
	$\epsilon_{\rm tag}$ (%)	ω (%)	$\epsilon_{\rm eff}$ (%)
cat 1	21.41±0.14	44.0±0.4	0.30±0.04
cat 2	13.29±0.11	37.2±0.4	0.87±0.06
cat 3	9.69 ±0.10	30.1±0.5	1.54±0.08
cat 4	6.24 ±0.08	22.9±0.6	1.83±0.08
cat 5	4.01 ±0.06	14.5±0.6	2.02±0.07
average	54.63±0.2	34.2±0.2	5.17±0.15
combined	54.63±0.2	32.7±0.2	6.55±0.15

Table 5 Performance in $B_s \to J/\psi\phi$ decays after the L0, obtained combining the taggers using a dedicated neural net.

One additional possibility that has been explored is the application of a linear transformation to the η_i of the taggers defined by $\omega' = (1 - k_i)\eta_i + k_i/2$ where k_i are 5 free parameters that minimize $|\omega_{\rm true} - \omega_{\rm estimated}|$. The minimum is found at $k_e = k_\mu = k_k = 0$, $k_{kSS} = 0.1$ and $k_{vtx} = 0.3$. The tagging performances are similar to those in the previous case, but the correction of the bias is not as good as in the case of the combination with the Neural Net.

4 Trigger and selection effects

The calibration of the Neural Net outputs is independent on the offline selection differences among decays, so when computing an event-per-event probability of the tagging decision to be right, there is no need for further corrections. The reason is that given a point in the phase space of the reconstructed B the mistag is found to be the same. See for instance Figure 13 which shows the mistag dependence of OS and SS kaons with the transverse momentum, p_T of the B_s. The linear fit is compatible among different decays.

However, since there is a dependence of the mistag with the $p_T(B)$, which can be different for each of the taggers, differences on ω have to be taken into account if it is measured globally (i.e. on events already sorted into tagging categories), because of the differences in the spectrum of the control channels and in the $B_s \to J/\psi\phi$ spectrum. Figure 14 shows the $p_T(B_s)$ distributions before and after the L0 trigger. This differences in spectra can be introduced by offline selections and also by the trigger. For instance, most $B_s \to J/\psi\phi$ pass the L0 trigger due to the muons while $B_s \to D_s\pi$ events due to hadrons. To cope with these differences, the samples can be split into Triggered Independently of Signal (TIS) and Triggered On Signal (TOS) categories. Table 6 gives the fraction of TIS and TOS events in

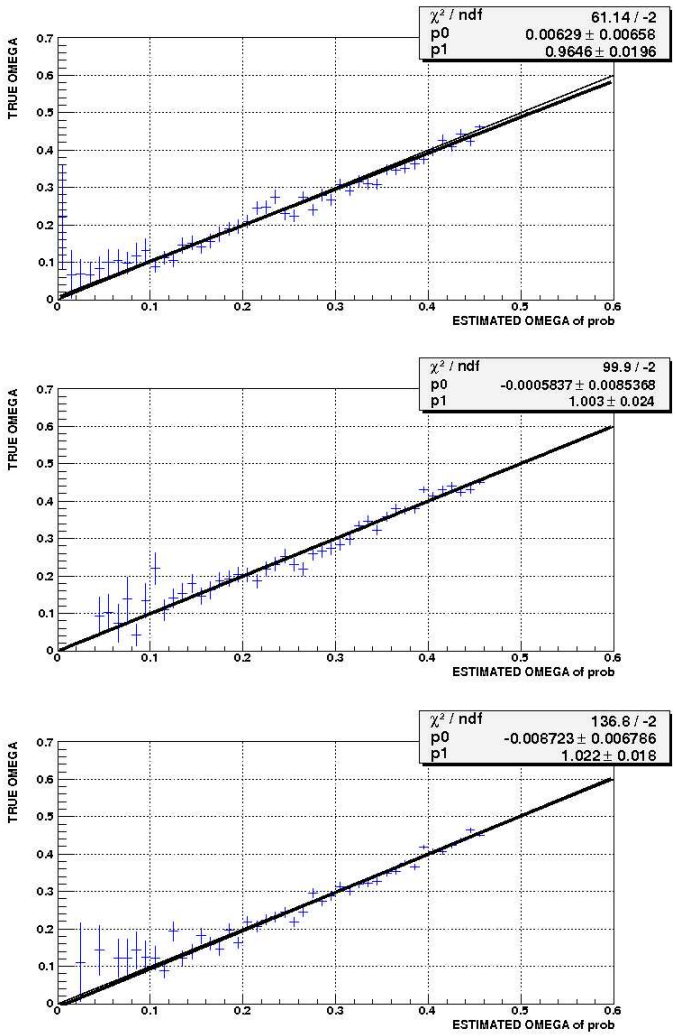

Figure 12 True wrong tag fraction versus the estimated wrong tag fraction from the Neural Net combination for the three channels $B_s \to D_s\pi$ (top), $B_d \to \pi\pi$ (center) and $B^+ \to J/\psi K^+$ (bottom).

different decays, where TIS also includes TIS+TOS events ($B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ channels have similar fractions as $B_s \to J/\psi\phi$).

Figure 15 shows the $p_T(B)$ distributions of TIS and TOS samples. The performance of tagging on TIS events is expected to be similar among different decays and higher than in TOS events, because in the case of TIS events there is a higher probability that the trigger was fired by some decay products of the opposite B. This explains why for example $\epsilon_{\rm eff}$ is 6.4% for $B_s \to J/\psi\phi$ and 8.7% for $B_s \to D_s\pi$. Because of the selection cuts in the signal B, TOS events can exhibit larger differences among different decays, that can be corrected for knowing the $\omega_{p_T(B)}$ dependence of each tagger, as explained in detail in [11, 12, 13]. OS taggers ω increases with p_T of B_s, while SS just does the opposite. When estimating the global mistag of each one of the 5 categories including all taggers, this dependence is compensated, and as a result, similar values of ω are found for $B_s \to J/\psi\phi$ and $B_s \to D_s\pi$ despite their differences in spectrum.

	L0 TIS (%)	L0 TOS (%)
$B_s \to J/\psi\phi$	23.5	74.3
$B_s \to D_s\mu\nu_\mu$	36.8	61.0
$B_s \to D_s\pi$	55.1	44.2

Table 6 Fraction of Triggered Independently of Signal (TIS) and Triggered On Signal (TOS) events. For the $B^0 \to J/\psi K^*$ and $B^+ \to J/\psi K^+$ modes the corresponding fractions are the same as for the $B_s \to J/\psi\phi$ channel due to the same topology of the decay.

5 Conclusions

A way to calibrate the response of the Same Side kaon tagger using $B_s \to D_s\pi$ events has been presented along with different possible procedures to combine Same Side with Opposite Side tagging. Each one may show advantages or disadvantages depending on the specific physics analysis involved. In all cases, the inclusion of the Same Side implies a very significant improvement, increasing the final effective tagging efficiency by almost a factor 2.
The absolute precision that can be achieved with $B_s \to D_s\pi$ events on the determination of $\omega_{\rm SS}$ using the number of events equivalent to one year data taking is ± 0.005.

In order to combine the tagger algorithms to give a final decision for the flavour of the B, the baseline choice would be to remove the secondary vertex charge in the presence of other OS taggers, because of the simplicity of the method. On the other hand, the Neural Net method provides a better final tagging performance in different decay modes with the same initial tuning. Both ways to proceed can correct the bias issue between the estimated and the true wrong tag fraction and will allow to have an event-per-event correct estimation of ω that can be used as input to the CP fits.

$B_s \to J/\psi\phi$			
	$\epsilon_{\rm tag}$ (%)	ω (%)	$\epsilon_{\rm eff}$ (%)
Individual taggers			
μ	6.02±0.08	31.1±0.6	0.86±0.06
e	2.91±0.06	31.0±0.9	0.42±0.04
OS K	15.60±0.12	35.3±0.4	1.34±0.07
SS K	26.75±0.15	34.9±0.3	2.43±0.10
$Q_{\rm vtx}$	34.64±0.16	40.9±0.3	1.14±0.07
Combination of all taggers			
cat 1	27.46±0.15	42.8±0.3	0.57±0.05
cat 2	11.06±0.10	35.1±0.5	0.98±0.06
cat 3	7.61±0.09	28.9±0.5	1.35±0.07
cat 4	5.14±0.07	23.2±0.6	1.48±0.07
cat 5	4.20±0.07	14.8±0.6	2.08±0.08
average	55.47±0.16	35.8±0.2	4.50±0.13
combined	55.47±0.16	32.93±0.2	6.46±0.15
Combination of OS taggers only			
cat 1	27.18±0.15	44.9±0.3	0.28±0.04
cat 2	6.97±0.08	35.5±0.6	0.59±0.05
cat 3	5.02±0.07	30.8±0.7	0.74±0.05
cat 4	3.68±0.06	26.0±0.8	0.84±0.06
cat 5	2.13±0.05	16.7±0.8	0.95±0.05
average	45.0±0.2	39.0±0.2	2.19±0.10
combined	45.0±0.2	36.2±0.2	3.40±0.11

Table 7 Flavour tagging performance for offline selected $B_s \to J/\psi\phi$ events passing the Level-0 trigger, for the individual taggers and for their combination. Results using opposite side tagger only are shown as well. Uncertainties are statistical.

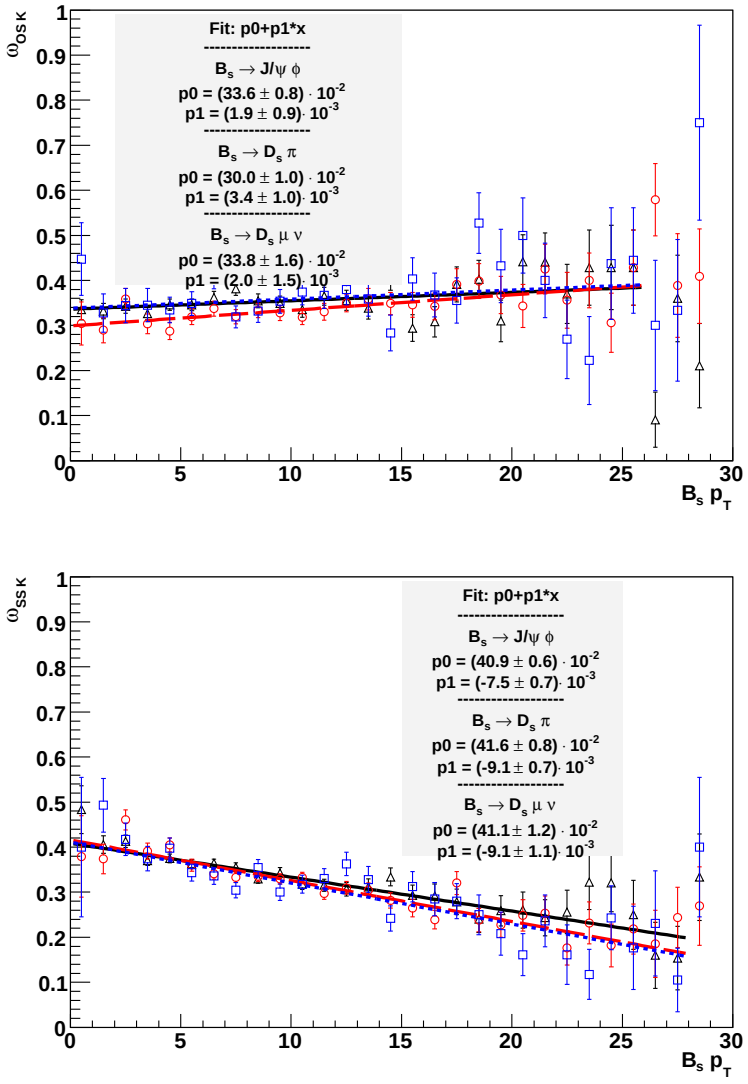

Figure 13 Wrong tag fraction calculated using MC information of OS kaon tagger (top Figure) and SS kaon tagger (bottom Figure) vs p_T of the reconstructed B_s. Results for offline selected $B_s \to J/\psi\phi$ (black triangles, full line fit), $B_s \to D_s\pi$ (red circles, dashed line fit) and $B_s \to D_s\mu\nu_\mu$ (blue squares, dotted line fit) events after the L0 trigger.

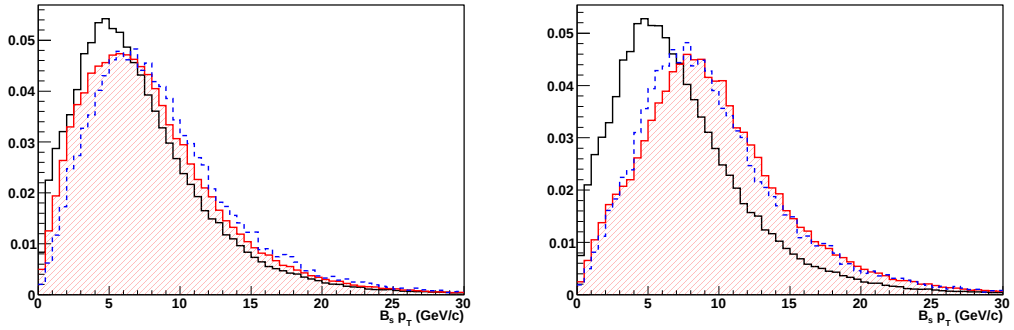

Figure 14 p_T of the reconstructed $B_s \to J/\psi\phi$ (black empty histogram), $B_s \to D_s\pi$ (red dashed pattern) and $B_s \to D_s\mu\nu_\mu$ (blue dashed line empty histogram) events before (left plot) and after (right plot) the L0 trigger.

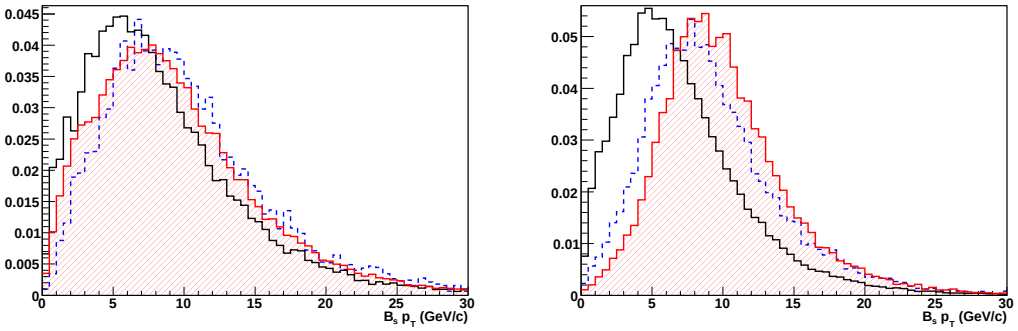

Figure 15 p_T of the reconstructed $B_s \to J/\psi\phi$ (black empty histogram), $B_s \to D_s\pi$ (red dashed pattern) and $B_s \to D_s\mu\nu_\mu$ (blue dashed line empty histogram) events classified as TIS (left plot) and TOS (right plot).

$B_s \to D_s\pi$			
	$\epsilon_{\rm tag}$ (%)	ω (%)	$\epsilon_{\rm eff}$ (%)
Individual taggers			
μ	9.2±0.1	29.0±0.6	1.6±0.1
e	2.9±0.1	30.6±1.1	0.43±0.05
OS K	15.0±0.1	31.6±0.5	2.0±0.1
SS K	26.0±0.2	30.9±0.4	3.8±0.2
$Q_{\rm vtx}$	39.5±0.2	40.0±0.3	1.6±0.1
Combination of all taggers			
cat 1	27.81±0.18	43.1±0.4	0.53±0.06
cat 2	10.77±0.12	33.3±0.6	1.21±0.08
cat 3	8.22±0.11	27.3±0.6	1.70±0.10
cat 4	6.18±0.10	23.7±0.7	1.70±0.09
cat 5	6.83±0.10	14.0±0.5	3.54±0.12
average	59.8±0.2	33.8±0.2	6.25±0.19
combined	59.8±0.2	31.0±0.2	8.67±0.20
Combination of OS taggers only			
cat 1	27.88±0.18	44.2±0.4	0.38±0.05
cat 2	7.93±0.11	34.4±0.7	0.77±0.07
cat 3	5.95±0.09	28.9±0.7	1.06±0.07
cat 4	4.33±0.08	25.0±0.8	1.09±0.07
cat 5	3.11±0.07	18.4±0.9	1.24±0.07
average	49.2±0.2	37.4±0.3	3.10±0.14
combined	49.2±0.2	34.8±0.3	4.53±0.15

Table 8 Flavour tagging performance for offline selected $B_s \to D_s\pi$ events passing the Level-0 trigger, for the individual taggers and for their combination. Also results using opposite side tagger only are shown. Uncertainties are statistical.

6 References

[1] M. Calvi *el al., Flavour Tagging Algorithms and Performances in LHCb.* LHCb note 2007-058.

[2] *"Road map for the measurement of mixing induced CP violation in* $B_s \to J/\psi\phi$ *at LHCb."*, LHCB-ROADMAP3-001.

[3] M. Calvi *et al., Calibration of Flavour Tagging with* $B^+ \to J/\psi K^+$ *and* $B^0 \to J/\psi K^{*0}$ *control channels at LHCb.* LHCb note 2009-020.

[4] O. Leroy, F. Muheim and S. Poss, *Selection of* $B_s \to D_s \mu\nu_\mu$ *events in LHCb.* LHCb note 2007-029.

[5] O. Awunor and U. Egede, *Using Double Tagging to measure the performance of the Same Side Kaon Tagger in data.* LHCb note 2007-127.

[6] M. Calvi et al., *Lifetime Unbiased Selection of* $B_s \to J/\psi\phi$ *and related control channels:* $B_d \to J/\psi K^*$ *and* $B^+ \to J/\psi K^+$*"*, LHCb note 2009-025.

[7] LHCb Collaboration, R. Antunes et al., *LHCb Trigger System Technical Design Report*, CERN/LHCC 2001-0040.

[8] V. Gligorov, *Reconstruction of the decay modes* $B^0_{d,s} \to D^{\pm}_{d,s}(\pi, K)^{\mp}$ *at LHCb.* LHCb note 2009-003.

[9] http://roofit.sourceforge.net

[10] http://tmva.sourceforge.net

[11] B. Souza de Paula, *Sensibilidade ao Parametro* ϕ_s *com o Canal de Decaimento* $B_s \to \phi\phi$ *e Efeitos Sistematicos no Experimento LHCb.* CERN-THESIS 2007-020.

[12] L. Somerville, *Performance of the LHCb RICH Photon Detectors and Tagging Systematics for CP Violation Studies.* CERN-THESIS 2006-020.

[13] M. Calvo, *Backsplash Effects on SPD and Flavour Tagging in LHCb.* CERN-THESIS 2006-011.