CERN-LHCb/2003-127, PHYS
September 11, 2003

$B_S^0 \to D_S^{\pm} K^{\mp}$ and $B_S^0 \to D_S^- \pi^+$ event selection

A. Golutvin[1], R. Hierck[2], J. van Hunen[2], M. Prokudin[1], R. White[3]

[1] *ITEP, Moscow, Russia*
[2] *NIKHEF, Amsterdam, The Netherlands*
[3] *Imperial College, London, UK*

Abstract

The decay channels $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\pm} K^{\mp}$ are used to extract the physics parameters Δm_s, $\Delta \Gamma_s$ and $\gamma + \phi_s$. Simulation studies showed that with one year of data taking a total of (82±26)k $B_s^0 \to D_s^- \pi^+$ and (5.4±2.2)k $B_s^0 \to D_s^{\pm} K^{\mp}$ events are reconstructed, triggered and selected.

The B/S for $B_s^0 \to D_s^- \pi^+$ originating from B-inclusive events is expected to be (0.32±0.10), while for $B_s^0 \to D_s^{\pm} K^{\mp}$ a 90% CL upper limit for the B-inclusive background is obtained: B/S <0.5. For $B_s^0 \to D_s^{\pm} K^{\mp}$ an extra source of background, $B_s^0 \to D_s^- \pi^+$, gives an additional contribution to B/S = (0.111±0.056)%.

Contents

1 Introduction

This note describes the selection of the two specific $B_s^0 \to D_s^- h^+$ decays, $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$. Since both decays are very similar, the selection criteria for both decays are almost identical. The analysis presented here is an update of the analysis described in Ref. [1]. The update includes the use of version v8r3p4 of the DaVinci package, increased statistics for signal and background and improved Particle ID performance. The new particle identification algorithms now combine the information of the various subdetectors to make combined global likelihoods for all particle hypothesis.

The flavour specific $B_s^0 \to D_s^- \pi^+$ decay is used to measure the B_s-mixing and decay parameters Δm_s and $\Delta\Gamma_s$, while interference in the $B_s^0 \to D_s^\pm K^\mp$ decays are used to measure the CP parameter $\gamma + \phi_s$. The final results obtained in this note are used in the physics sensitivity studies, described in Ref. [2].

2 $B_s^0 \to D_s^- h^+$ event topology

Figure 1 illustrates the typical topology of a $B_s^0 \to D_s^- \pi^+$ or $B_s^0 \to D_s^\pm K^\mp$ event. Since the B_s-meson has a life time τ_{B_s} of 1.5 ps and typically traverses a few centimeters before decaying, a secondary vertex can be observed. The pion/kaon that accompanies the D_s is referred to as the bachelor π/K. This particle has a relative high (transverse) momentum. The D_s-meson has a life time τ_{D_s} of 0.5 ps and therefore also here a detached vertex can be observed. The decay of the D_s^- into the final state $K^-K^+\pi^-$ has been selected for its significant branching ratio. Almost all of these 3-body final state are produced through the decay of resonances, for instance $D_s^- \to \phi(\to K^+K^-)\pi^-$ or $D_s^- \to K^{0*}(\to \pi^-K^+)K^-$. The specific masses of these resonances can also be used in the event selection. The event selection described in this note does not explore the properties of the resonances, therefore this method is applicable to all $D_s \to KK\pi$ final states.

Table 1 summarizes the contributions to the branching ratio (BR) of the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$ decay. The total $B_s^0 \to D_s^\pm K^\mp$ branching ratio is a factor 12 smaller than the $B_s^0 \to D_s^- \pi^+$ branching ratio. These numbers will be used as input for the calculation of the total expected event yield for signal and background as described in the Section 6.

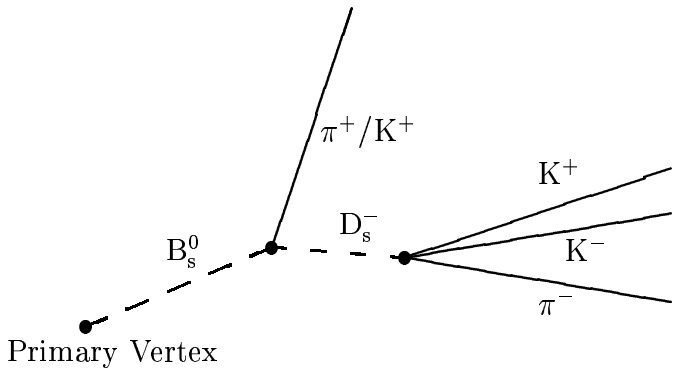

Figure 1: *Topology of a $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^- K^+$) decay.*

Process	Assumption/measured	Branching fraction
Probability(b-quark$\to$$B_s$)	measured	0.100
$B_s^0 \to D_s^- \pi^+$	=BR($B_d^0 \to D^- \pi^+$)	$(2.76 \pm 0.25) \times 10^{-3}$
$B_s^0 \to D_s^- K^+$	=BR($B_d^0 \to D^- K^+$)	$(2.0 \pm 0.6) \times 10^{-4}$
$B_s^0 \to D_s^+ K^-$	=BR($B_d^0 \to D_s^+ \pi^-$)	$(2.7 \pm 1.0) \times 10^{-5}$
$D_s^- \to K^- K^+ \pi^-$	measured	$(4.4 \pm 1.2) \times 10^{-2}$
Total $B_s^0 \to D_s^- (\to K^+ K^- \pi^+) \pi^+$		1.21×10^{-4}
Total $B_s^0 \to D_s^\pm (\to K^\pm K^- \pi^+) K^\mp$		$10. \times 10^{-6}$

Table 1: *Expected or assumed branching fractions for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$. If a branching ratio has not been measured, the specific assumption is mentioned. The same numbers hold for the charged conjugated decays. All numbers are obtained from Ref. [3].*

3 Monte Carlo event samples

To tune the event selection a sample of signal and a sample of background events are used. The tuning aims to preserve high efficiency for selecting the signal B-decays, while rejecting backgrounds, particularly from other B-decays. Two types of background can be distinguished.

The first source of background consists of exclusively reconstructed B-decays which resemble the signal decay and have a significant branching ratio. Examples are $B_d^0 \to K^\pm \pi^\mp$ in the $B_d^0 \to \pi^+ \pi^-$ selection or $B_s^0 \to D_s^- \pi^+$ in the $B_s^0 \to D_s^\pm K^\mp$ selection. In these particular cases a mistake in the particle identification of a single track gives rise to background. Misidentifying the bachelor π in a $B_s^0 \to D_s^- \pi^+$ as a K can result in the selection of a background $B_s^0 \to D_s^\pm K^\mp$ event. Since the $B_s^0 \to D_s^- \pi^+$ branching ratio is a factor 12 larger than the $B_s^0 \to D_s^\pm K^\mp$, $B_s^0 \to D_s^- \pi^+$ is considered as a background for $B_s^0 \to D_s^\pm K^\mp$, while the other way around is expected to be negligible.

The second type of background is pure combinatorial background. Random combinations of tracks can lead to a B-meson candidate. This type of background can, in principle, occur in any interaction (minimum bias events). Since the event selection as well as the trigger requires the presence of a detached secondary vertex and high p_t tracks, it is assumed here that only events in which a $b\bar{b}$-quark pair is produced can lead to such background events. This sample of events is referred to as the B-inclusive event sample. For both the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$ event selection, the B-inclusive sample is considered as a possible source of background.

To save CPU time in the Monte Carlo simulation, an angular acceptance cut (ε_θ) at the event generator level has been applied. This cut already removes a large fraction of events which do not have the final state tracks in the detector acceptance. For the signal events the B-meson of interest has to be produced with an angle smaller than 400 mrad [1]. This angular acceptance criterion selects ε_θ=34.7% of all signal events. The B-inclusive background must have at least one of the two produced B-particles within 400 mrad. With this criterion

[1] All particles produced with a larger angle are not detected by the LHCb setup, since the LHCb acceptance only reaches up to 300 mrad.

ε_θ=43.2% of the generated events are accepted.

Even for this reduced signal event sample only 15.6% of the events are *reconstructible* events. A reconstructible event is defined to have all four decay particles (KK$\pi\pi$ or KKπK) reconstructible [2] as long, upstream or downstream track. The reconstructibility efficiency therefore is equal to $\varepsilon_{\rm accept}$=15.6%. The track reconstruction inefficiency again reduces the number of reconstructible events to the number of *reconstructed* events. Of all reconstructible $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$ events respectively 80.6% and 82.0% are fully reconstructed.

The Monte Carlo generated event samples used in the tuning of the event selection are: 196.5k $B_s^0 \to D_s^- \pi^+$ event, 1.07M $B_s^0 \to D_s^\pm K^\mp$ events and 5.4M B-inclusive events. In order to give a systematically unbiased estimate of the background, the tuned selection criteria are applied to another sample of 5.2M B-inclusive events. These events are used for the background estimate.

As will become clear in the discussion on the background event yields, the available statistics is insufficient to determine a reliable background estimate. Several methods are applied to obtain a better estimate. One method to artificially increase the background statistics is to apply a loose B_s mass cut for the background event. For this method the selected background events in the broad mass window are scaled down to the tight mass window. In this selection factor a 10 times larger mass window ($\pm$500 MeV/c^2) is used. In Section 6.3 two other methods are described.

4 Pre-selection of $B_s^0 \to D_s^- h^+$

Before fine tuning all the selection criteria, all B_s candidates are selected with loose pre-selection criteria such that reduced event samples remain. Table 2 summarizes the pre-selection cuts which are applied on the B_s candidates. It should be noted that both an unconstrained vertex fit and a vertex fit with the D_s mass constraint is applied for D_s candidates and that the unconstrained D_s vertex is used to create the B_s vertex.

Cut	value
Track type	Long
Unconstrained D_s vertex χ^2	<20
Constrained D_s vertex χ^2	<20
D_s mass window	50 MeV/c^2
Unconstrained B_s vertex χ^2	<20
B_s mass window	500 MeV/c^2

Table 2: *B_s and D_s candidate pre-selection criteria on the reconstructed and fitted decay vertices.*

In addition to B_s and D_s vertices also a primary vertex is reconstructed (more than one reconstructed primary vertex is possible in the case that there are multiple interactions in the event). In the case that there is more than one reconstructed primary vertex it is assumed

[2]The definition of reconstructible tracks can be found in the TDR [4]

that the B_s candidate originates from the vertex with the highest track multiplicity [3].

Figure 2 shows the primary vertex resolution for $B_s^0 \to D_s^- h^+$ events. Double Gaussian distributions are fitted through the data points. In each figure the width of the first Gaussian is quoted as the core resolution.

Although in LHCb the events are boosted in the forward direction, the primary vertices have good resolution in z. This is due to the fact that the reconstructed vertex contains many tracks, typically 60, including tracks with a large opening angle in the laboratory frame.

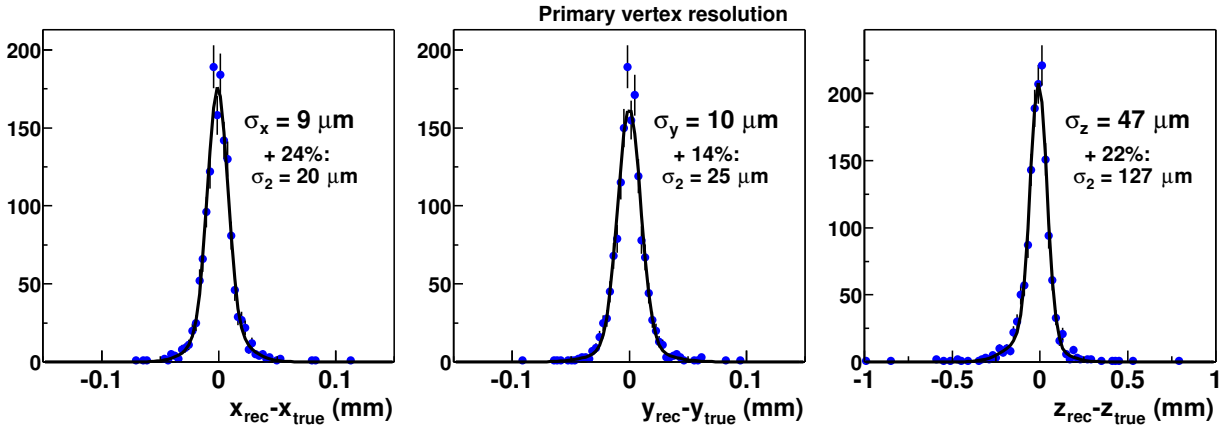

Figure 2: *Primary vertex position resolutions. A double Gaussian is fitted through the distributions. In each figure the width of the first Gaussian, the fraction of events in the second Gaussian and the width of the second Gaussian are given.*

5 $B_s^0 \to D_s^- h^+$ event selection

This section describes the tuning of the kinematic cuts to select a clean sample of $B_s^0 \to D_s^- h^+$ events. The selection is separated in three parts: the individual track selection, the selection of D_s candidates and the selection of the B_s candidates. The figures in the next sections show the kinematic distributions of the signal events (the true B_s decay tracks) and the combinatorial background from the B-inclusive events (the tracks selected by the pre-selection criteria). In each figure the cut values are indicated.

5.1 Track selection

Fitting a track yields a χ^2 per degree of freedom, indicating the goodness of the fit. For a precise reconstruction of the B_s and D_s decay vertices, well reconstructed tracks are required. Each track must therefore have a $\chi^2/\mathrm{NDF} < 4$.

The particles which originate from a B-decay have relative high (transverse) momentum. The momentum of all the reconstructed particles which are used for the reconstruction of

[3] This leaves room for improvement since this method sometimes selects the wrong primary vertex in events with multiple interactions.

the B_s decay are required to be above 2 GeV/c. Figure 3 shows the transverse momentum distributions for the D_s candidate decay products and the bachelor π candidates. The p_t of the bachelor pion is required to be above 700 MeV/c and the p_t for each of the D_s decay products is required to be above 300 MeV/c.

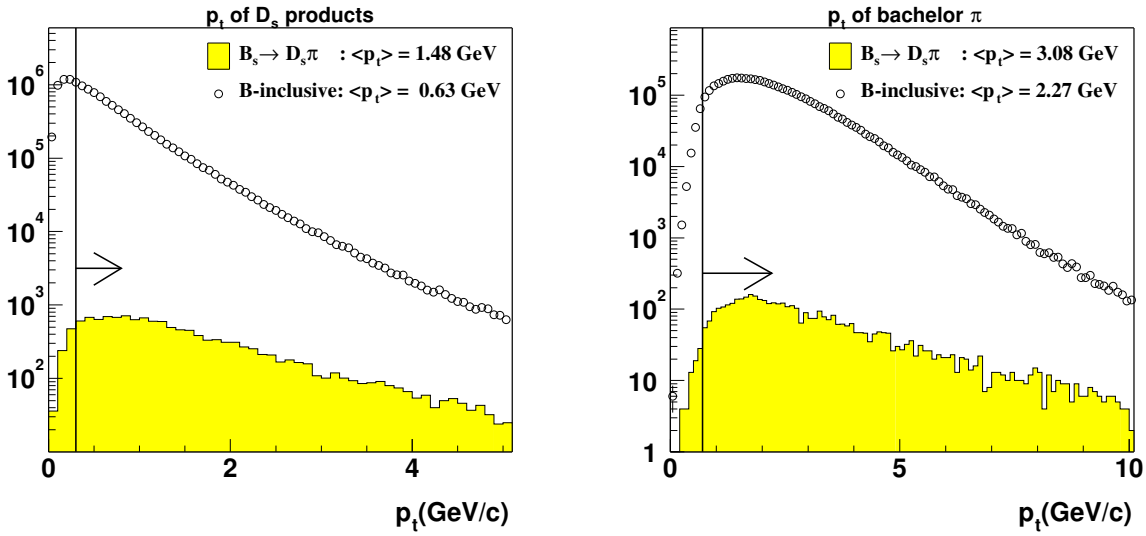

Figure 3: *Transverse momentum distribution of the D_s decay products and the bachelor pion.*

Figure 4 shows the sum of the p_t of the three D_s decay products. The sum for the particles from the true D_s is much higher than for the random combinations. Therefore it is required that $\sum_{KK\pi} p_t$ >2200 MeV/c.

Good π/K separation is required for background rejection and separation of the $B_s^0 \rightarrow D_s^{\pm}K^{\mp}$ and $B_s^0 \rightarrow D_s^{-}\pi^{+}$ channels. The particle identification algorithms combine the information from the various subdetectors (RICH, ECAL, HCAL, MUON), to assign a global likelihood for each particle hypothesis [4]. For this analysis, each particle must be identified by the RICH and the likelihood separation for the kaons must be $\Delta\mathcal{L}_{K\pi}$ >-5 and for the pions $\Delta\mathcal{L}_{\pi K}$ >-5 is required.

Figure 5 shows the $\Delta\mathcal{L}_{K\pi}$ separation for the bachelor kaons and pions from the $B_s^0 \rightarrow D_s^{\pm}K^{\mp}$ and $B_s^0 \rightarrow D_s^{-}\pi^{+}$. The much lower branching ratio of $B_s^0 \rightarrow D_s^{\pm}K^{\mp}$, requires more stringent cuts on the bachelor particle. In order to obtain a sufficient background rejection both the K/π and K/e separation in $B_s^0 \rightarrow D_s^{\pm}K^{\mp}$ events is required to be: $\Delta\mathcal{L}_{K\pi}$ >2 and $\Delta\mathcal{L}_{Ke}$ >2.

At this point the presented cuts are applied to the signal and background event samples, such that reduced samples remain for the tuning of the D_s vertex selection discussed in the next section.

5.2 D_s vertex selection

To each KKπ track combination a D_s vertex is fitted with a constrained mass fit. In order to obtain good vertices, the χ^2 of the fit must be <10.

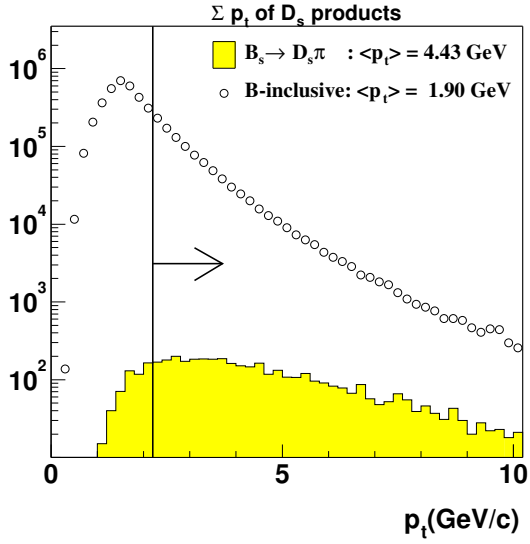

Figure 4: *The sum of the transverse momenta of the D_s decay products.*

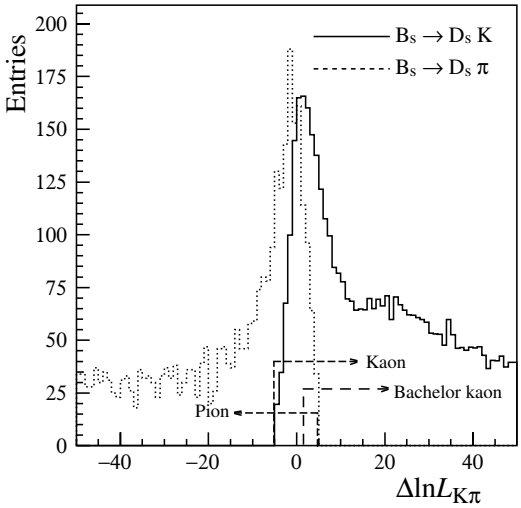

Figure 5: *Kaon pion separation for the bachelor particles from $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$ events.*

One can calculate the distance of closest approach of a track to a (primary) vertex. This so-called Impact Parameter (IP) is schematically illustrated in Fig. 6. Using the uncertainty on the primary vertex position and the error on the track parameters x, y, and z of the reconstructed track, an impact parameter significance ($S_{\rm IP}$) can be calculated,

$$S_{\rm IP} = \frac{\rm IP}{\sigma_{\rm IP}} \quad . \tag{1}$$

Since the decay products of the D_s do not originate from the primary vertex, their track states do not necessarily point back in the direction of the primary vertex. These tracks therefore can have a significant impact parameter. All particles, which make up the D_s vertex, are required to have a $S_{\rm IP}(D_s \text{ prod.})>1$. The distributions for the signal and background are shown in Fig. 7.

The reconstructed mass should be close to the true D_s mass (1969 MeV/c^2). Figure 8 shows the D_s mass distribution from the unconstrained vertex fit for signal and background. The mass distribution of the signal D_s is fitted with a double Gaussian function. The core resolution of the first Gaussian is 5.1 MeV/c^2. A $\sim 3\sigma$ mass window cut is applied: $|\Delta m| <15$ MeV/c^2. Around the D_s mass the background distribution is flat, while further from the correct mass, the distribution drops rapidly. This is caused by the pre-selection cut on the χ^2 of the constrained mass fit.

In a similar way as the impact parameter of the individual particle, also the impact parameter for the reconstructed D_s is calculated. From Fig. 6 it can be seen that a large impact parameter for the D_s can be expected. Figure 9 shows the impact parameter significance of the D_s candidates from the signal and background. The D_s impact parameter significance is required to be $S_{\rm IP}(D_s)>2$.

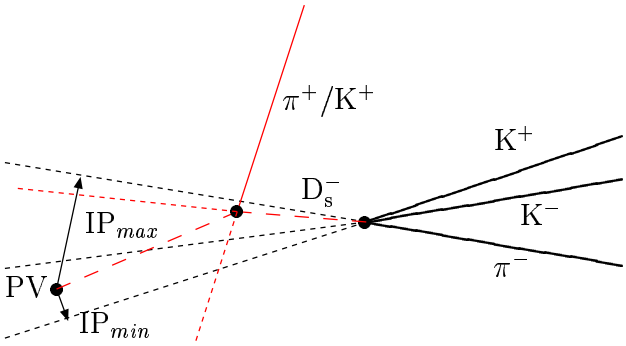

Figure 6: *Schematic illustration, only in 2D, of the impact parameter definition.*

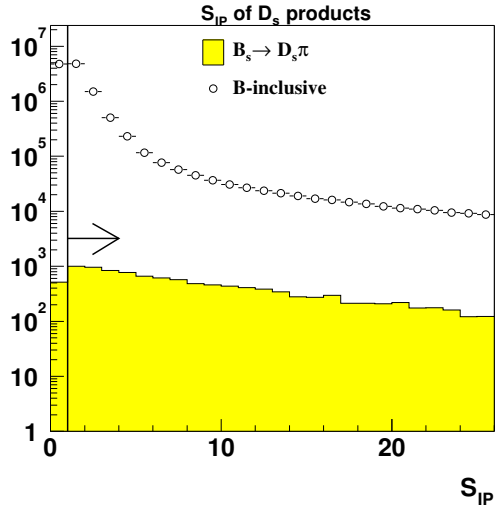

Figure 7: *The significance of the impact parameter of the particles which are used to construct the D_s candidates.*

In the signal events, the 3-particle vertices of the D_s, obtained using an unconstrained vertex fit, have a resolution of the spatial parameters as indicated in Fig. 10. Because the products from the decay are mainly produced under small angles in the forward direction, the resolution in z is not as good as that for the primary vertex (see Fig. 2). The x and y resolutions are less affected.

The D_s vertex should be separated from the primary vertex. Figure 11 shows the significance of the z-position separation from the primary vertex. This is calculated using,

$$S_z(\mathrm{D_s}, \mathrm{PV}) = \frac{z_{\mathrm{D_s}} - z_{\mathrm{PV}}}{\sqrt{\sigma^2_{z,\mathrm{D_s}} + \sigma^2_{z,\mathrm{PV}}}} \quad . \tag{2}$$

The D_s vertex is required to be more than 4.5 standard deviations away from the primary vertex: $S_z(\mathrm{D_s},\mathrm{PV}) > 4.5$.

At this point the presented cuts are applied to the signal and background event samples, such that reduced samples remain for the tuning of the B_s vertex selection discussed in the next section.

5.3 B_s vertex selection

The B_s vertex is obtained by combining the reconstructed D_s candidate with a bachelor pion/kaon track. The typical position resolutions of the B_s vertex are shown in Fig. 13. The relatively large opening angle between the bachelor h and the D_s, due to the large mass difference between initial and final state, provides significantly better resolutions than those for the D_s vertex.

The first requirement on the reconstructed B_s vertex is that the χ^2 of the unconstrained fit has to be smaller than 4 (see Fig. 12).

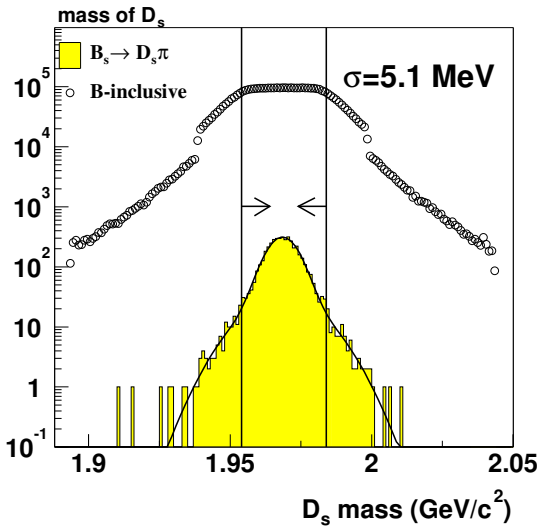

Figure 8: *The mass distribution of the selected signal D_s and combinatorial background. A double Gaussian is fitted through the signal distribution and the core of the first Gaussian is quoted as the resolution. The discontinuities in the background distribution at 1.94 and 1.99 GeV/c^2 are an artifact of the pre-selection.*

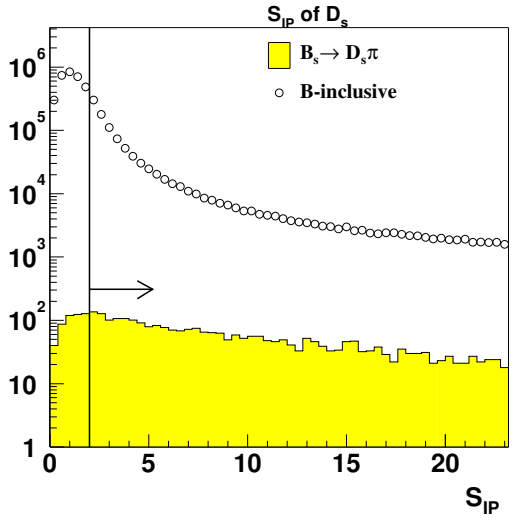

Figure 9: *The impact parameter significance of the reconstructed D_s candidates.*

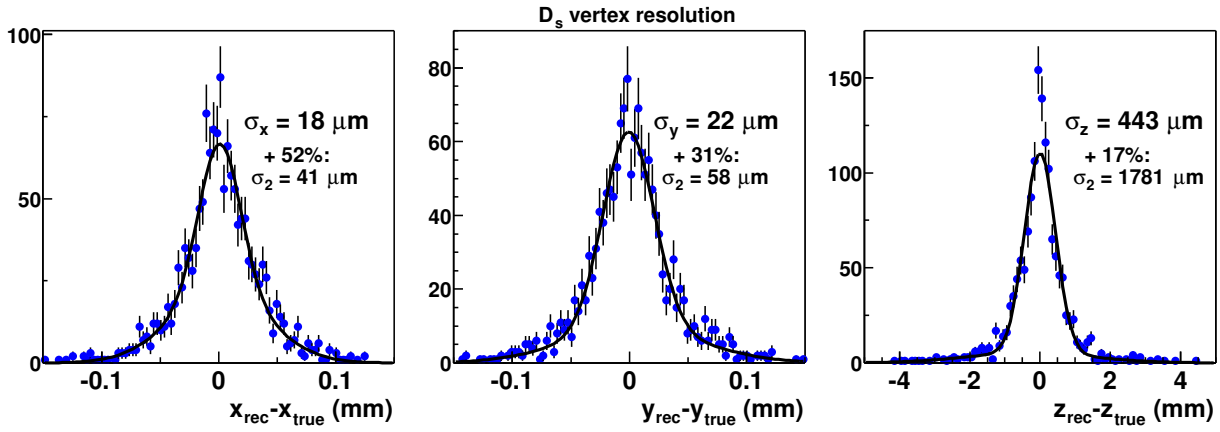

Figure 10: *D_s vertex position resolution. A double Gaussian is fitted through the data points, the width of the first Gaussian, the fraction in the second Gaussian and the width of the second Gaussian are given.*

As can be seen from Fig. 6 the impact parameter from the bachelor pion in the B-decay is expected to be large. On the other hand, the reconstructed B-meson should point back to the primary vertex (small impact parameter). The distributions for these two parameters

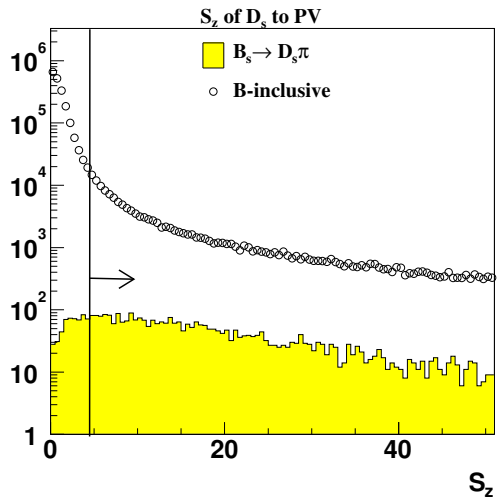

Figure 11: *The significance of the separation of the reconstructed primary vertex and the D_s vertex.*

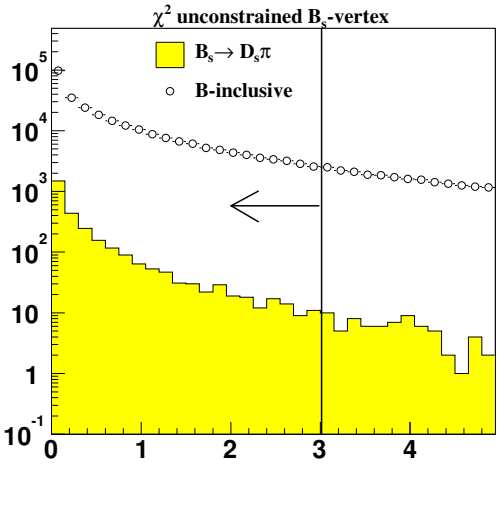

Figure 12: *The χ^2 from an unconstrained fit of the B_s vertex.*

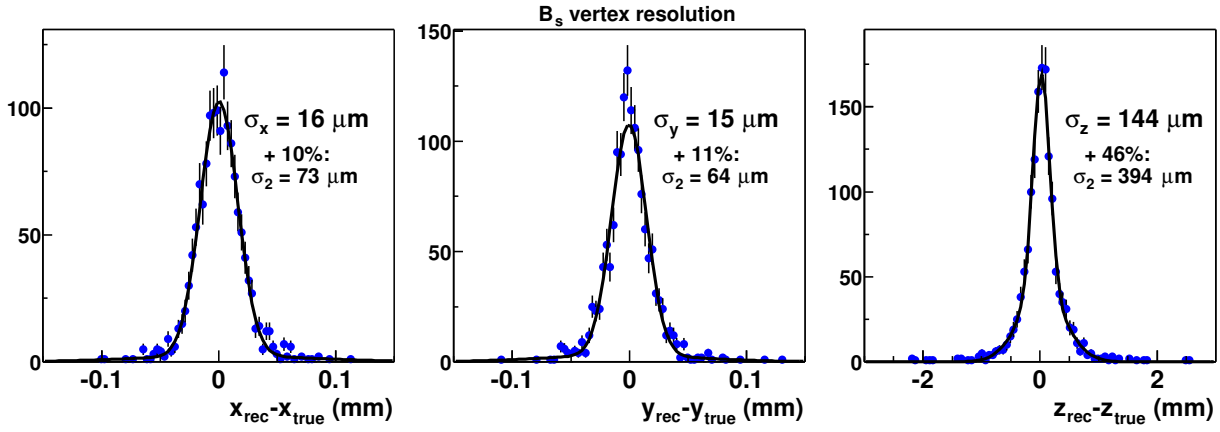

Figure 13: *B_s vertex position resolutions using the unconstrained vertex fit. Double Gaussians are fitted through the data points, the width of the first Gaussian, the fraction in the second Gaussian and the width of the second Gaussian are given.*

are shown in Fig. 14. In order to reduce the background the significance of the bachelor h must be $S_{\rm IP}(h)$ >4, while the B_s is required to have an impact parameter significance $S_{\rm IP}(B_s)$ <3.

As shown in Section 5.2 the D_s selection algorithm already required that the D_s vertex is downstream of the primary vertex ($S_z(D_s, {\rm PV}) > 4.5$). The position of the B_s vertex should be between the primary vertex and the D_s vertex. Figure 15 shows the distance between the B_s and D_s vertex. A small fraction of the signal decays has a reconstructed D_s vertex upstream of the reconstructed B_s vertex ($\Delta z < 0$ mm). This is caused by the limited vertex

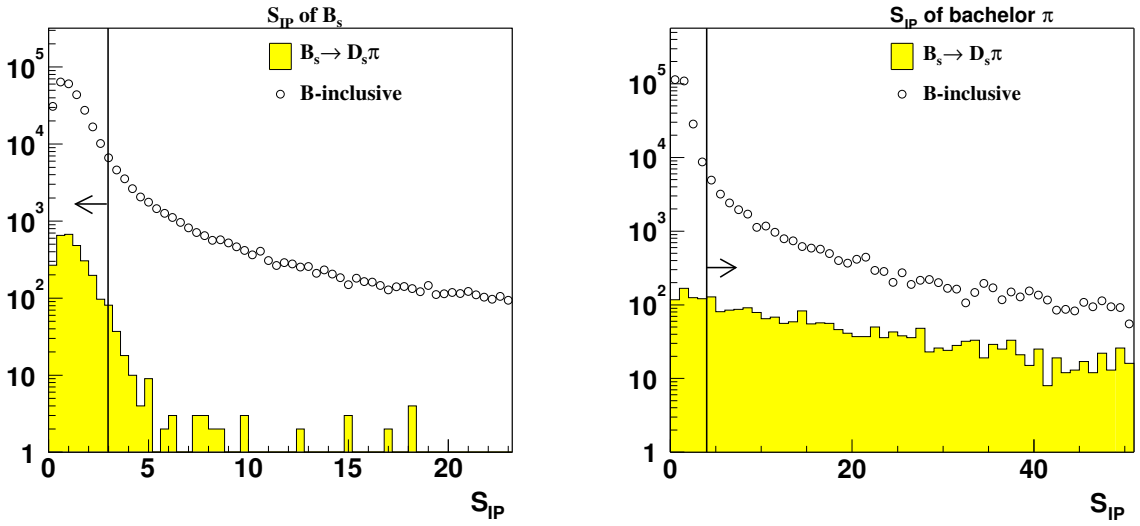

Figure 14: *Impact significance of the B_s (left) and the bachelor π/K (right).*

resolution (see Figures 10 and 13). The selection requires that the reconstructed D_s vertex is downstream of the reconstructed B_s vertex ($\Delta z > 0$ mm).

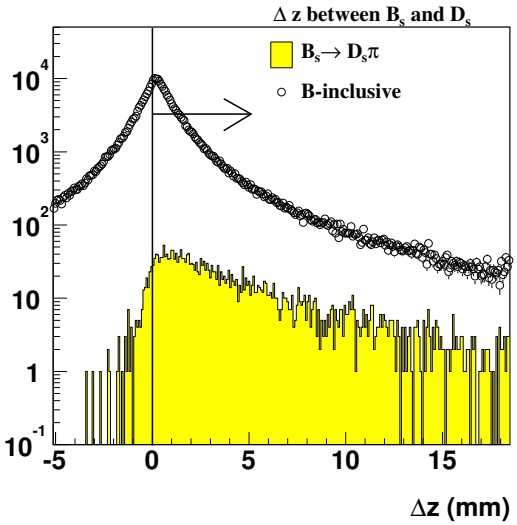

Figure 15: *Distance between the B_s and D_s vertex in mm.*

The fact that the B_s-meson should have a momentum vector pointing back toward the primary vertex is further exploited. The vector $\vec{\mathbf{r}}$ is constructed from the positions of the primary and secondary vertex. This vector is expected to be collinear with the momentum vector $\vec{\mathbf{p}}$ from the B_s. Figure 16 schematically illustrates the angle θ, defined as

$$\cos(\theta) = \frac{\vec{\mathbf{p}} \cdot \vec{\mathbf{r}}}{|\vec{\mathbf{p}}||\vec{\mathbf{r}}|} \quad . \tag{3}$$

Figure 17 shows the $\cos(\theta)$ distribution for the reconstructed signal and background B_s

candidates. A cut at $\cos(\theta)$ >0.99997 removes a large fraction of the background, while retaining most of the signal events.

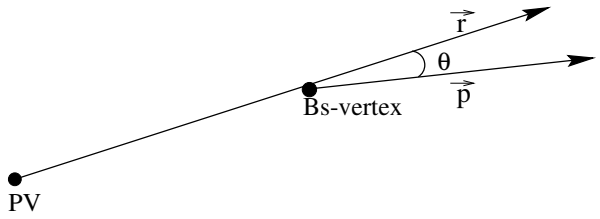

Figure 16: *Schematic illustration of the definition of the angle θ.*

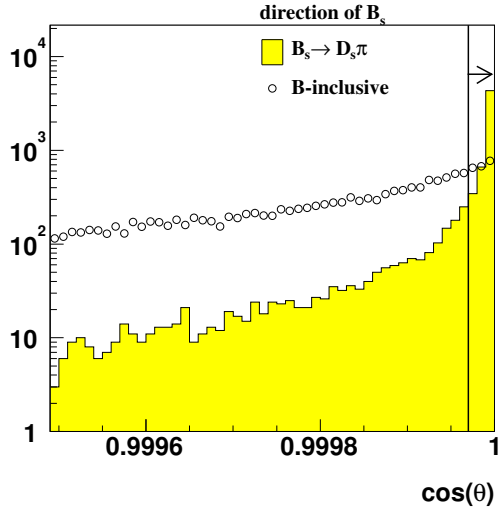

Figure 17: *The cos(θ) distribution for the selected signal and background B_s candidates.*

The most efficient background rejection criterion has not been discussed so far. The reconstructed B_s-meson is expected to have a mass close to the true B_s-mass (m_{B_s}=5370 MeV/c^2). Requiring a reconstructed mass within $\pm$50 MeV/c^2 of the true mass selects most of the signal events. Since the statistics for the background are limited, the mass cut for the background is not put at $\pm$50 MeV/c^2 but at $\pm$500 MeV/c^2. In Fig. 18 the mass resolution of the correctly reconstructed B_s-mesons is shown. In section 6.2.2 this figure will be discussed in more detail. For both the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$ a mass resolution of 13.8 MeV/c^2 is obtained.

Table 3 summarizes all the selection criteria discussed in this section.

6 Event yield calculation and background estimates

The following sections discuss the selection efficiency of signal and background, the L0 and L1 trigger efficiency, the background-to-signal ratio and the expected annual event yields.

The total offline selection efficiency can be factorized in several individual contributions. Firstly the detector acceptance $\varepsilon_{\rm det}$, consisting of the angular cut and the fraction of events that are reconstructible ($\varepsilon_{\rm det}$=$\varepsilon_\theta \times \varepsilon_{\rm accept}$), then there is the reconstruction efficiency $\varepsilon_{\rm rec/det}$ and finally the selection efficiency, $\varepsilon_{\rm sel/rec}$. The total offline selection efficiency is therefore expressed as

$$\varepsilon_{\rm off} = \varepsilon_{\rm det} \times \varepsilon_{\rm rec/det} \times \varepsilon_{\rm sel/rec} \quad . \tag{4}$$

In addition to the reconstruction and selection efficiency, also the trigger performance is taken into account. The total efficiency is therefore defined as

$$\varepsilon_{\rm tot} = \varepsilon_{\rm off} \times \varepsilon_{\rm trg/sel} \quad . \tag{5}$$

Selection requirements	$B_s^0 \to D_s^- h^+$
All products p	2 GeV/c
D_s product p_t	300 MeV/c
D_s product $\sum p_t$	2200 MeV/c
bachelor K/π p_t	700 MeV/c
pions: $\Delta\mathcal{L}_{\pi K}$	>-5
kaons: $\Delta\mathcal{L}_{K\pi}$	>-5
χ^2/NDF	< 4
Constrained D_s vertex χ^2	<10
D_s mass window	15 MeV/c^2
$S_{IP}(D_s$ prod)	>1
$S_{IP}(D_s)$	>2
$S_z(D_s, PV)$	>4.5
Unconstrained B_s vertex χ^2	<3
$S_{IP}(h)$m	>4
$S_{IP}(B_s)$	<3
cos(θ)	0.99997
$z_{B_s} - z_{D_s}$	>0 μm
B_s mass window (B-inclusive)	$\pm$50 (500) MeV/c^2
specific $B_s^0 \to D_s^\pm K^\mp$ selection criteria	
bachelor K: $\Delta\mathcal{L}_{K\pi}$	>2
bachelor K: $\Delta\mathcal{L}_{Ke}$	>2

Table 3: *Summary of the cuts used for the selection of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$ events.*

With the studied simulated events, all these efficiency numbers can be obtained. From the sample of simulated events the total offline selection efficiency is calculated by

$$\varepsilon_{\text{off}} = \frac{N_{\text{sel}}}{N_{\text{gen}}} * \varepsilon_\theta \quad , \tag{6}$$

where N_{gen} is the number of generated events, N_{sel} is the number of selected events and ε_θ is the fraction of events accepted by the angular acceptance cut. The detection efficiency and reconstruction efficiency are obtained by analyzing the simulated events.

To calculate the annual event yield for LHCb, the total number of produced events must be multiplied with the total efficiency. The number of produced events [4] is calculated by

$$N_{\text{prod}} = \mathcal{L}_{\text{int}} \times \sigma_{b\bar{b}} \times (2 \times \text{Prob}(\bar{b}\text{-quark} \to B_s)) \times \text{BR}_{\text{vis}} \quad , \tag{7}$$

where $\mathcal{L}_{\text{int}}$ is the annual integrated luminosity ($2{\times}10^{32}$ cm^{-2}s^{-1} for 10^7 s per year), $\sigma_{b\bar{b}}$ is the assumed $b\bar{b}$ production cross-section of 500 μb, Prob($\bar{b}$-quark $\to B_s$) is the fraction of b-quarks which produce a B_s-meson, the factor 2 includes the fact that per event two b-quarks have the probability to produce a B_s. The last factor in the equation is the visible branching ratio to the final state.

[4] This is the sum of both the B_s^0 and the $\overline{B_s^0}$ decay.

The annual yield of offline selected and triggered events is calculated by

$$N_{\rm phys} = \varepsilon_{\rm tot} \times N_{\rm prod} \quad . \tag{8}$$

6.1 $B_s^0 \to D_s^- \pi^+$ event yield

In this section the expected $B_s^0 \to D_s^- \pi^+$ event yields are calculated. These estimates are obtained from the analysis of $N_{\rm gen}$=196.5k signal events and 5.2M B-inclusive events (cuts have been tuned on 5.4M B-inclusive events). Applying all the presented selection criteria results in a total of $N_{\rm sel}^{\rm signal}$= 6,125 selected signal events in the tight mass windows and $N_{\rm sel}^{\rm B\text{-}inclusive}$=32 selected background events in the wide mass window.

6.1.1 $B_s^0 \to D_s^- \pi^+$ signal selection

For the signal $B_s^0 \to D_s^- \pi^+$ events, by using Eq. (6) the total selection efficiency is calculated to be (1.082±0.014)%. This efficiency is subdivided (Eq. 4) into

$$\varepsilon_{\rm off} = \quad \varepsilon_\theta \quad \times \varepsilon_{\rm accept} \times \varepsilon_{\rm recons} \times \varepsilon_{\rm sel} = \tag{9}$$

$$0.01082 = \quad 0.347 \quad \times 0.156 \quad \times \quad 0.806 \quad \times \quad \mathbf{0.250} \quad . \tag{10}$$

Substituting the B_s production probability and the final state branching ratios from Table 1 in Eq. (7), the number of produced signal events can be calculated. This calculation gives a total of $N_{\rm prod}$=(24.3±7.7)M $B_s \to D_s(\to KK\pi)\pi$ events per year [5]. With a selection efficiency of $\varepsilon_{\rm off}$= (1.082±0.014)%, the expected number of offline selected events per year is (262.8±82.8)k.

The L0 and L1 trigger algorithms are tuned using offline selected events from various decay channels. In the tuning the number of triggered and selected events is maximized taking into account the constraints of the 1 MHz and 40 kHz output rate on the L0 and L1 trigger respectively [6]. For the $B_s^0 \to D_s^- \pi^+$ decay the efficiency of the tuned L0 trigger is $\varepsilon_{\rm L0}$=(49.4±0.6)% and the efficiency for the L1 trigger for L0 accepted events is $\varepsilon_{\rm L1}$=(63.0±0.9)%. The combined L0/L1 efficiency is (31.1±0.6)%. This results in a total of $N_{\rm phys}$=(81.7±25.8)k selected and triggered events per year.

6.1.2 B-inclusive background selection

Before the various numbers for the background are calculated, the individually selected background events are inspected in more detail. Since a large mass windows is used, some selected events from the B-inclusive event sample should not be counted as real background events. An example of such a reconstructed decay is $B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$, where the γ is not used in the invariant mass calculation. The mass of this reconstructed B-meson will always be on the lower side of the mass window. These events will never pass the tight mass window cut and should therefore not be counted as combinatorial background. Table 4 subdivides the 32 selected background events in six different groups.

Below the different groups are discussed in more detail.

[5] The error on the number of produced is events is dominated by the large uncertainty in the branching fraction.

nr.	event type	# events	remarks
1	$B_s^0 \rightarrow D_s^- \pi^+$	5	is true signal
2	true D_s and π same B	11	
3	true D_s and π other B	**2**	
4	$D^{\pm}(\rightarrow K\pi\pi)\pi$	**2**	misidentification $\pi \rightarrow$K
5	$D^{\pm}(\rightarrow K\pi\pi)\pi X$	6	misidentification $\pi \rightarrow$K and incomplete reconstruction
6	Random combinations	**6**	
	Total	32	

Table 4: *The subdivision of the 32 selected events from the B-inclusive sample. The numbers in bold are real background event, the rest is only introduced due to the wide mass window and will never appear in the tight mass window.*

1. The true $B_s^0 \rightarrow D_s^- \pi^+$ signal events in the B-inclusive event sample should not be counted as background.

2. A correctly reconstructed D_s, originating from a B-decay, is combined with a π originating from the same B-decay. Decays of this type are for example $B_s^0 \rightarrow D_s^{-*}(\rightarrow D_s^- \gamma)\pi^+$, $B_s^0 \rightarrow D_s^- \rho(\rightarrow \pi^- \pi^0)$ or $B_s^0 \rightarrow D_s^{-*}(\rightarrow D_s^- \gamma)\pi^+ \pi^0$. These events will have a systematically underestimated mass and should not be counted as background.

3. A correctly reconstructed D_s, originating from a B-decay, is combined with a π originating from the other B-decay. This should be counted as background.

4. A misidentification of a π for a K, reconstructs a $D^{\pm}(\rightarrow K\pi\pi)$ as a $D_s(\rightarrow KK\pi)$. The pion also originates from the same B-decay: $B^0 \rightarrow D^{\pm}(\rightarrow K\pi\pi)\pi$. The reconstructed mass is close to the true B_s mass and is therefore counted as background.

5. Similar as the previous type of events, but the reconstruction is incomplete. This reconstructed mass will be underestimated and should therefore not be counted as background.

6. The last type is true random combinatorics and should be counted as background. In one event a ghost track has been used in the reconstruction.

Taking this analysis into account, a total of 10 background events in the broad mass window are real background events. The factor 10 for the broad mass window gives an estimate of 1.0±0.3 background event in the tight mass window. When using the tight mass window cut also 1 events is observed.

For the background event yield a similar calculation as for the signal events can be done. The number of produced B-inclusive events is equal to $N_{\text{prod}} = \mathcal{L}_{\text{int}} \times \sigma_{b\bar{b}}$, which gives a total of 10^{12} $b\bar{b}$ events per year.

The 10 events in the broad mass window gives a selection efficiency of $(8.3 \pm 2.6) \times 10^{-8}$. This translates into a total of (82.8±26.2)k selected background events per year. With the

expected event yields for the signal and background events an estimation of the background-to-signal ratio can be made.

$$\left(\frac{B}{S}\right)^{\mathrm{B-inclusive}}_{\mathrm{B_s\to D_s\pi}} = 0.32 \pm 0.10 \tag{11}$$

The limited number of selected background events (only 10) makes it impossible to give a reliable estimate of the trigger efficiency for the selected background events. Using the same trigger efficiency for the background as for the signal is a conservative estimate of the trigger performance.

Table 5 summarizes all the efficiency numbers for the $B_s^0 \to D_s^- \pi^+$ selection.

		Factors (in %) contributing to $\varepsilon_{\mathrm{tot}}$						
	N_{prod}	$\varepsilon_{\mathrm{det}}$	$\varepsilon_{\mathrm{rec/det}}$	$\varepsilon_{\mathrm{sel/rec}}$	$\varepsilon_{\mathrm{trg/sel}}$	$\varepsilon_{\mathrm{tot}}$	N_{phys}	B/S
$B_s^0 \to D_s^- \pi^+$	(24±8)M	5.4	80.6	25.0	31.1	0.337	(82±26)k	0.32±0.10

Table 5: *The various efficiency factors which contribute to the total efficiency. Also the expected event yield and background level are given.*

6.2 $B_s^0 \to D_s^\pm K^\mp$ event yield

The $B_s^0 \to D_s^\pm K^\mp$ event yield is estimated from a sample of 1.07M signal events. After all selection criteria $N_{\mathrm{sel}}^{\mathrm{signal}}$=28,223 events remain.

Since the branching fraction of $B_s^0 \to D_s^\pm K^\mp$ is a factor of 12 lower than $B_s^0 \to D_s^- \pi^+$, more detailed studies of the different backgrounds are required. From the 10.6M used B-inclusive background events, 2 events pass the selection criteria in the broad mass window. In a dedicated study using the 196.5k signal $B_s^0 \to D_s^- \pi^+$ decays, 47 have been accepted by the $B_s^0 \to D_s^\pm K^\mp$ selection in the tight mass window.

In the next few subsections, the selection efficiency, trigger efficiencies and background-to-signal ratios are calculated.

6.2.1 $B_s^0 \to D_s^\pm K^\mp$ signal selection

By using Eq. (6) an offline selection efficiency of $\varepsilon_{\mathrm{off}} = (0.910 \pm 0.005)\%$ is obtained. Equation (7) estimates that a total of (2.0±0.8)M events per year are produced. By combining these numbers, this results in a total of (18.18±7.3)k $B_s^0 \to D_s^\pm K^\mp$ selected events per year.

Due to the slightly different selection criteria, the L0 and L1 trigger efficiencies also differ. The L0 efficiency for $B_s^0 \to D_s^\pm K^\mp$ is a little lower: $\varepsilon_{\mathrm{L0}}$=(47.2±0.3)%, but the L1 efficiency is a little higher: $\varepsilon_{\mathrm{L1}}$=(62.6±0.4)%. The combined L0/L1 efficiency is (29.5±0.3)%. This results in a total of (5.4±2.2)k selected and triggered $B_s^0 \to D_s^\pm K^\mp$ events per year.

6.2.2 $B_s^0 \to D_s^- \pi^+$ background selection

Figure 18 shows the reconstructed mass of the true $B_s^0 \to D_s^\pm K^\mp$ events and the selected $B_s^0 \to D_s^- \pi^+$ background events. For this plot the 50 MeV/c^2 mass cut has not been applied.

	$N_{\rm prod}$	Factors (in %) contributing to $\varepsilon_{\rm tot}$					$N_{\rm phys}$
		$\varepsilon_{\rm det}$	$\varepsilon_{\rm rec/det}$	$\varepsilon_{\rm sel/rec}$	$\varepsilon_{\rm trg/sel}$	$\varepsilon_{\rm tot}$	
$B_s^0 \to D_s^{\pm} K^{\mp}$	(2.0±0.8)M	5.4	82.0	20.6	29.5	0.269	(5.4±2.2)k

Table 6: *Various efficiency contributions which lead to the total number of selected and triggered events per year.*

The 47 selected events in the 50 MeV/c^2 mass window give an offline selection efficiency of $\varepsilon_{\rm off} = (9.0 \pm 0.04) \times 10^{-5}$ and correspond to a background-to-signal ratio of:

$$\left(\frac{B}{S}\right)^{B_s \to D_s \pi}_{B_s \to D_s K} = 0.111 \pm 0.056 \quad . \tag{12}$$

Note that the background-to-signal can be improved, at the cost of event yield, by tightening the mass window.

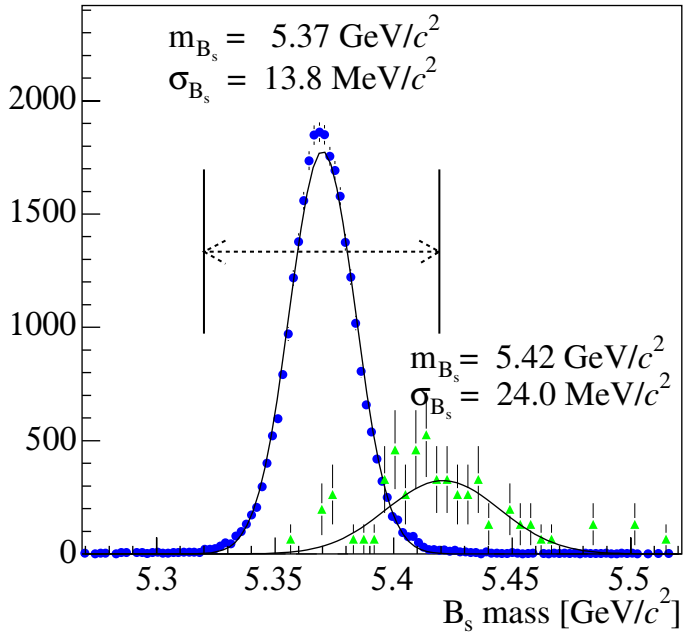

Figure 18: *The mass resolution for the selected $B_s^0 \to D_s^{\pm} K^{\mp}$ events together with the selected background $B_s^0 \to D_s^- \pi^+$ events. All selection criteria have been applied, apart from the 0 MeV/c^2 mass window cut. The figure shows the mass resolution and the bounds of the 50 MeV/c^2 mass window . The histograms are correctly normalized to compensate for the different branching ratios.*

6.2.3 B-inclusive background selection

As already mentioned, the selection criteria have been tuned on a total of 5.4M B-inclusive events. For the $B_s^0 \to D_s^{\pm} K^{\mp}$ selection 1 of these events passed all cuts. This event is a random combinatorics event. Applying the same selection criteria to the other 5.2M events

also results in 1 selected event. However, this event turns out to be a true $B_s^0 \to D_s^- \pi^+$ event, which is a separately studied source of background. Therefore a total of $N_{\rm rec}^{\rm B\text{-}inclusive}$ =1 event is selected in the wide mass window from the 10.6M analyzed B-inclusive events.

The limited statistics for the selected background events only allows for a confidence level on the selection efficiency. At 90% confidence level, using the Feldman-Cousins method [5], the background selection efficiency is [0.04 - 1.78]$\times 10^{-8}$. For the background-to-signal ratio for the B-inclusive events a range of

$$\left(\frac{B}{S}\right)^{\rm B\text{-}inclusive}_{\rm B_s \to D_s K} = [0.024 - 0.978] \quad , \tag{13}$$

at 90% CL is expected. This corresponds to an upper limit of B/S<1.0.

6.3 Additional background studies

Since there only remains 1 background event, the accuracy of the background estimation is poor. A few additional studies have been performed to obtain a better estimate of the sources of background. The first is a study on a specific decay channel which can introduce additional background, $B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$. The second is a study of the B-inclusive background by opening some of the selection cuts.

6.3.1 $B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$ events

Section 6.2.2 showed that misidentification of the π from $B_s^0 \to D_s^- \pi^+$ results in the selection of background for $B_s^0 \to D_s^\pm K^\mp$. The mass distribution for these events is shifted to the higher mass side, with significant overlap. For the $B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$ decay also misidentification of the bachelor π can occur. Since the photon will not be considered in the reconstruction, the mass distribution will be shifted to the lower side. If the energy of the photon is low, there could be a significant overlap between the incorrectly reconstructed $B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$ decays and the $B_s^0 \to D_s^\pm K^\mp$ mass window.

The selection criteria are applied to 104k of simulated $B_s^0 \to D_s^{-*}\pi^+$ events. Figure 19 shows both the selected $B_s^0 \to D_s^\pm K^\mp$ and $B_s^0 \to D_s^{-*}\pi^+$ decays. Assuming that the branching fraction for $B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$ is equal to $B_s^0 \to D_s^- \pi^+$, the estimated background-to-signal for this specific decay channel is.

$$\left(\frac{B}{S}\right)^{\rm B_s \to D_s^* \pi}_{\rm B_s \to D_s K} = (0.022 \pm 0.011)\% \quad . \tag{14}$$

From this study it can be concluded that this type of background does not affect the purity of the selected $B_s^0 \to D_s^\pm K^\mp$ events. The energy of the photon is sufficient to shift the mass distribution to the lower side of the true B_s mass.

6.3.2 Opening of selection cuts

Since only 1 event remained from the event selection, no further cuts have been applied. The single selected B-inclusive background event originated from pure random combinatorics. Opening some selection criteria will introduce extra selected background events and with

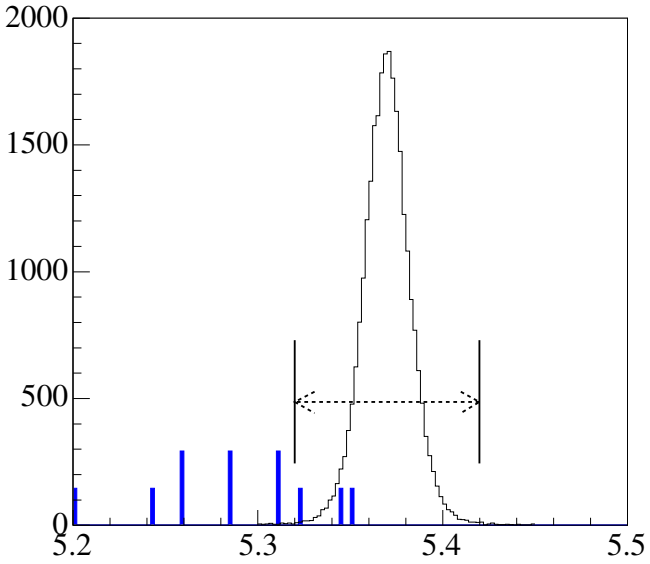

Figure 19: *Mass distribution of the* $B_s^0 \rightarrow D_s^{-*}(\rightarrow D_s^-\gamma)\pi^+$ *background and the signal* $B_s^0 \rightarrow D_s^{\pm}K^{\mp}$ *events. The distributions are correctly normalized.*

these additional background events a better understanding of the sources of background can be obtained.

By removing the requirement that the D_s vertex should be downstream of the B_s vertex (see Fig. 15), additional D_s candidates are introduced. By relaxing the cut on the likelihood of particle identification of the bachelor particle from $\Delta\mathcal{L}_{K\pi} > 5$ to $\Delta\mathcal{L}_{K\pi} > -5$ also extra B_s vertices are introduced. With these looser cuts a total of 68 B-inclusive events with a reconstructed mass above 4 GeV/c^2 are selected.

The background candidates are divided in two main groups:

- 29 decays where the D_s and the bachelor K are reconstructed from tracks originating from the same B decay;
- 39 decays where not all tracks originate from the same B decay.

Out of the 29 background events in the first group, 2 events have a random D_s vertex reconstructed, while 27 have three tracks originating from the same $D_s/D^{\pm}$ (true D_s reconstructed). In 10 cases the bachelor particle originates from a charm decay, while in 19 cases the bachelor does not originate from a charm decay.

Out of the 39 events in the second group, 38 contain a D_s reconstructed from random tracks, while only 1 contains a true D_s. For the bachelor particle 10 cases originate from a $b \rightarrow c \rightarrow s$ transition, 5 originate from other B-decay products and 24 cases where the bachelor does not originate from a B decay.

When cutting at a D_s proper time of < 4 and a p_t >2 GeV/c, the number of background events in the two groups reduce to respectively 29→22 and 39→6 events. This is a

reduction of approximately 60%. Assuming that these variables are completely independent of the already presented cuts, implementing these cuts will also reduce the background by approximately 60%. These two cuts reduce the signal selection efficiency only with 4%.

Based on these studies it is plausible that the background for the $B_s^0 \to D_s^{\pm} K^{\mp}$ selection can be reduced by a factor of two without too much loss in efficiency. This reduces the 90% CL upper limit for the B/S from 1.0 (as shown in Section 6.2.3) to B/S<0.5. In Table 7 the size of the different background sources are summarized.

	$B_s^0 \to D_s^- \pi^+$ B/S	$B_s^0 \to D_s^{-*}(\to D_s^- \gamma)\pi^+$ B/S	B-inclusive B/S
$B_s^0 \to D_s^{\pm} K^{\mp}$	0.111±0.056	0.022±0.011	<1.0 (0.5)

Table 7: *The background-over-signal ratio for the three types of studied background events. For the studied B-inclusive background both the "standard" 90% upper limit and the estimated 90% upper limit B/S when using the factorization method are given.*

The sensitivity studies [2] use the 90% upper limit of the expected B/S in the event simulations (B/S =0.5).

7 Proper time resolution

Since the selection criteria for the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\pm} K^{\mp}$ events are similar, the proper time dependencies for both decays are also almost identical.

Using the reconstructed B_s and primary vertex and the reconstructed B_s momentum, all with corresponding errors, the B_s life time can be fitted. This method is described in more detail in Ref. [7]. Figure 20 shows two proper-time dependent distributions of the selected B_s-mesons from the $B_s^0 \to D_s^- h^+$ events. The left plot in the figure shows the proper time distribution of the selected events. The right plot of the figure shows the time dependent combined reconstruction, selection and trigger efficiency of the $B_s^0 \to D_s^- h^+$ decay. The behavior can be fitted with a time dependent efficiency function $\varepsilon_t(t)$, which has the form

$$\varepsilon_t(t) = C \times \frac{(at)^3}{1 + (at)^3} \quad . \tag{15}$$

Here, a parametrizes the shape and C is a normalization factor. Fitting this distribution to the left plot of the figure results in a =1.29. Since the normalization is arbitrary, C is taken as an arbitrary number (explaining the vertical scale of the figure). In the trigger and selection criteria a detached B_s-vertex is explicitly required, thus rejecting B_s-mesons with small proper times. This is clearly visible in both plots.

The proper time resolution for the triggered and selected B_s-mesons is shown in Fig. 21. The distribution is fitted with a double Gaussian, where the first Gaussian has a width of 33±1 fs and describes 69% of the entries, while the rest of the entries is described by a Gaussian with a width of 67±3 fs.

The calculated errors and corresponding pull distribution are shown in Fig. 22. These errors are also used in the simulation of events for the sensitivity studies.

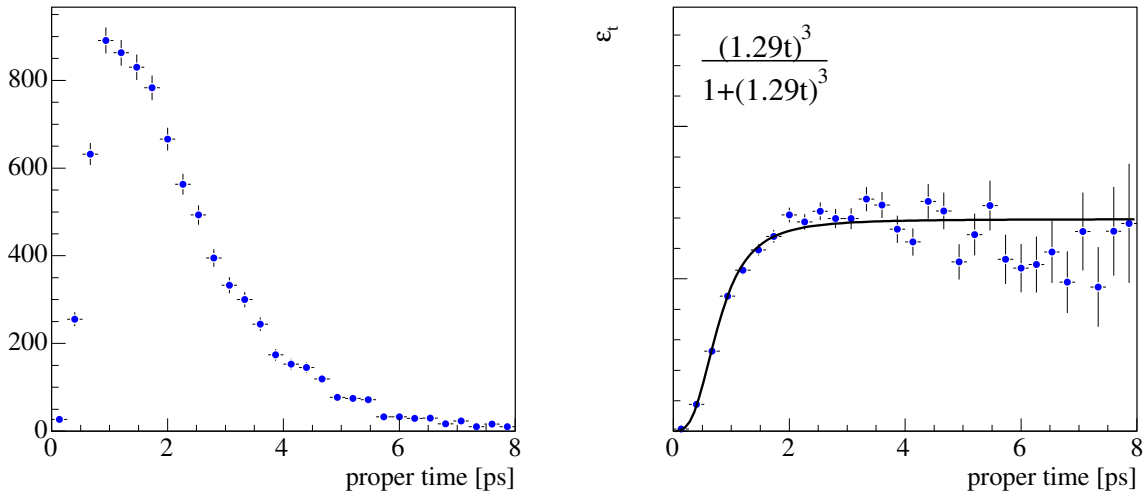

Figure 20: *The left plot shows the proper time of the selected $B_s^0 \to D_s^- h^+$ events. The right plot shows the time dependent efficiency of selected and triggered events. The function $\varepsilon_t(t)$ (Eq. 15) is fitted through the data points.*

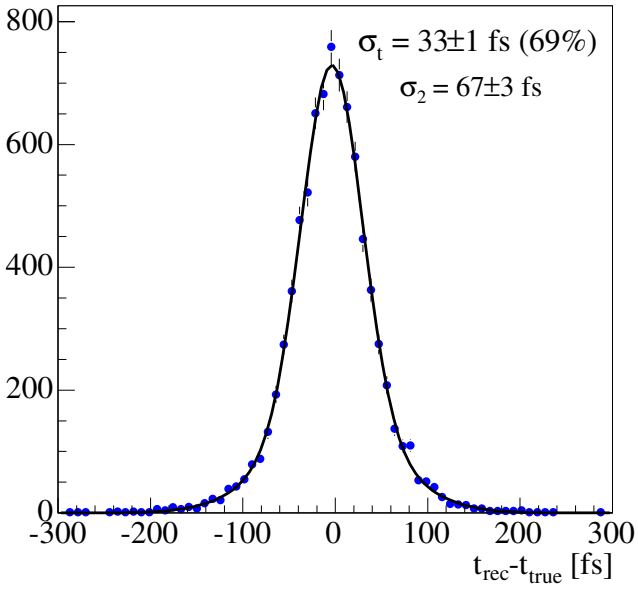

Figure 21: *Time resolution distribution for the reconstructed B_s-mesons.*

8 Conclusion

The presented selection criteria are common for both $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\pm K^\mp$, with the exception of a tighter particle ID cut on the bachelor K. The simulation studies shows that

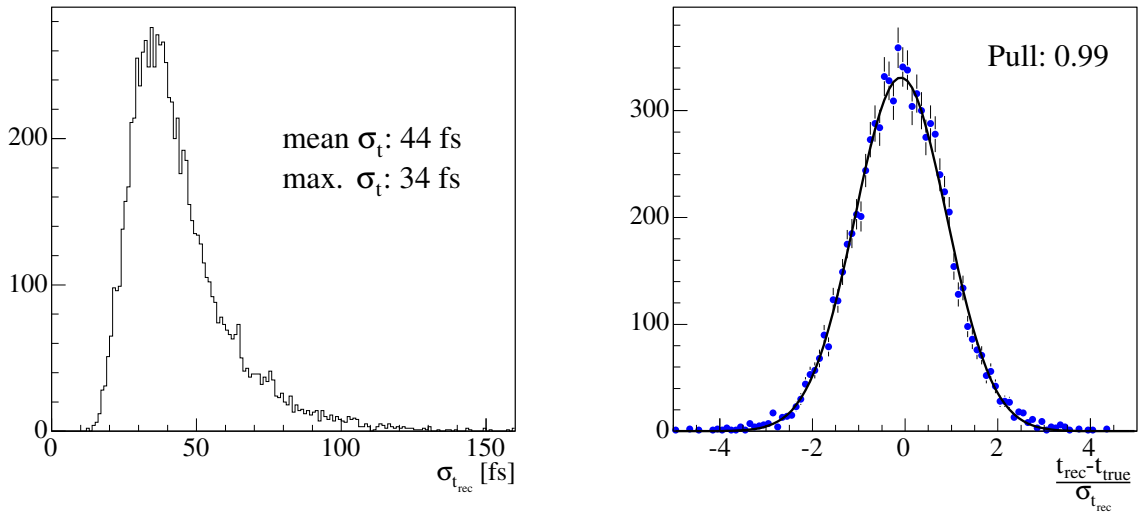

Figure 22: *Calculated error on the proper time (left) and the proper time pull distribution (right). In the left figure the average error and the position of the maximum of the distribution are given.*

LHCb is able to trigger and select (82±26)k of $B_s^0 \to D_s^- \pi^+$ events per year. The selection gives a B/S of 0.32±0.10.

The lower branching ratio of the $B_s^0 \to D_s^\pm K^\mp$ decay requires additional studies of the background. It is expected that one year of data provides a total of (5.4±2.2)k $B_s^0 \to D_s^\pm K^\mp$ events. The contamination from $B_s^0 \to D_s^- \pi^+$ events gives a B/S = 0.111±0.056. The contamination due to $B_s^0 \to D_s^{-*} \pi^+$ events is very small, B/S=0.022±0.011. The limited statistics for B-inclusive background only allows for an upper limit of the background. At 90% CL the B/S<1.0. The loosening of some selection cuts provided a better understanding of the background sources. By applying some new cuts more than 50% of the background could be reduced, with small loss of signal efficiency (4%). This conclusion has been used in the sensitivity studies, where B/S=0.5 is assumed [2].

References

[1] R. Hierck, "Optimisation of the LHCb detector", PhD. thesis, CERN-LHCb/2003-114

[2] R. Hierck, J. van Hunen, M. Merk, "The sensitivity for $\Delta m_{\rm s}$ and $\gamma+\phi_s$ from $\rm B_s^0 \to D_s^- \pi^+$ and $\rm B_s^0 \to D_s^{\pm} K^{\mp}$ decays", CERN-LHCb/2003-103.

[3] K. Hagiwara *et al.* [Particle Data Group], Phys. Rev. D **66** (2002) 010001, and 2003 off-year partial update for the 2004 edition available on the PDG web site: `http://pdg.lbl.gov/`.

[4] R. Antunes Nobrega *et al.* [LHCb Collaboration], "LHCb Reoptimization Technical Design Report", CERN-LHCC-2003-031.

[5] G. J. Feldman and R. D.Cousins, "Unified approach to the classical statistical analysis of small signals", Phys. Rev. D**57** (1998) 3873-3889

[6] R. Antunes Nobrega *et al.* [LHCb Collaboration], "Trigger Technical Design Report", CERN-LHCC/2003-030.

[7] G. Raven , "Selection of $\rm B_s^0 \to J/\psi\phi$ and $\rm B^{\pm} \to J/\psi K^{\pm}$", CERN-LHCb/2003-118.