CERN-LHCb/2003-123

Selection of $B^0_{(s)} \to h^+h^-$ decays at LHCb

G. Balbi, M. Bargiotti, A. Bertin, M. Bruschi, A. Carbone, L. Fabbri, P. Faccioli, D. Galli, B. Giacobbe, D. Gregori, F. Grimaldi, U. Marconi*, I. Massa, M. Piccinini, N. Semprini Cesari, R. Spighi, V. Vagnoni*, S. Vecchi, M. Villa, A. Vitale and A. Zoccoli.
(LHCb Bologna Group)

Dipartimento di Fisica dell'Università di Bologna and Ist. Naz. di Fisica Nucleare, Sezione di Bologna
**Corresponding authors:* Umberto.Marconi@bo.infn.it, Vincenzo.Vagnoni@bo.infn.it

Abstract

The selection of $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+K^-$ decays, which are traditionally chosen as benchmark channels to demonstrate the physics potential of the detector for fully hadronic modes, is presented. The strategy followed to identify the best selection cuts is described. Estimated annual signal yields and background-to-signal ratios for specific and combinatorial backgrounds are quoted. Relevant quantities, such as mass resolution, proper time resolution and proper time acceptance of signal events are studied in detail. Mass and proper time distributions of combinatorial background events are also investigated.

1 Introduction

The LHCb experiment will start operating at the Large Hadron Collider (LHC) by the 2007, with the aim of performing precise measurements of CP-violating effects arising in the beauty-quark flavor-dynamics [1, 2]. LHCb will exploit the large beauty production cross-section of the 14 TeV proton-proton collisions of LHC, that is expected to be of the order of 500 μb. Thanks to the large statistics of signals that LHCb will collect, effects of New Physics, possibly arising as small deviations from the Standard Model predictions, can be revealed.

The LHCb detector is a single-arm forward spectrometer, extremely specialized in selecting B-meson decays. It has been recently re-optimized [2] with respect to the original Technical Proposal project [1]. It is characterized by excellent vertex, momentum and proper decay time resolutions, robust, flexible and efficient trigger system, as well as performant particle identification sub-detectors. A sketch of the LHCb re-optimized detector is shown in Figure 1.

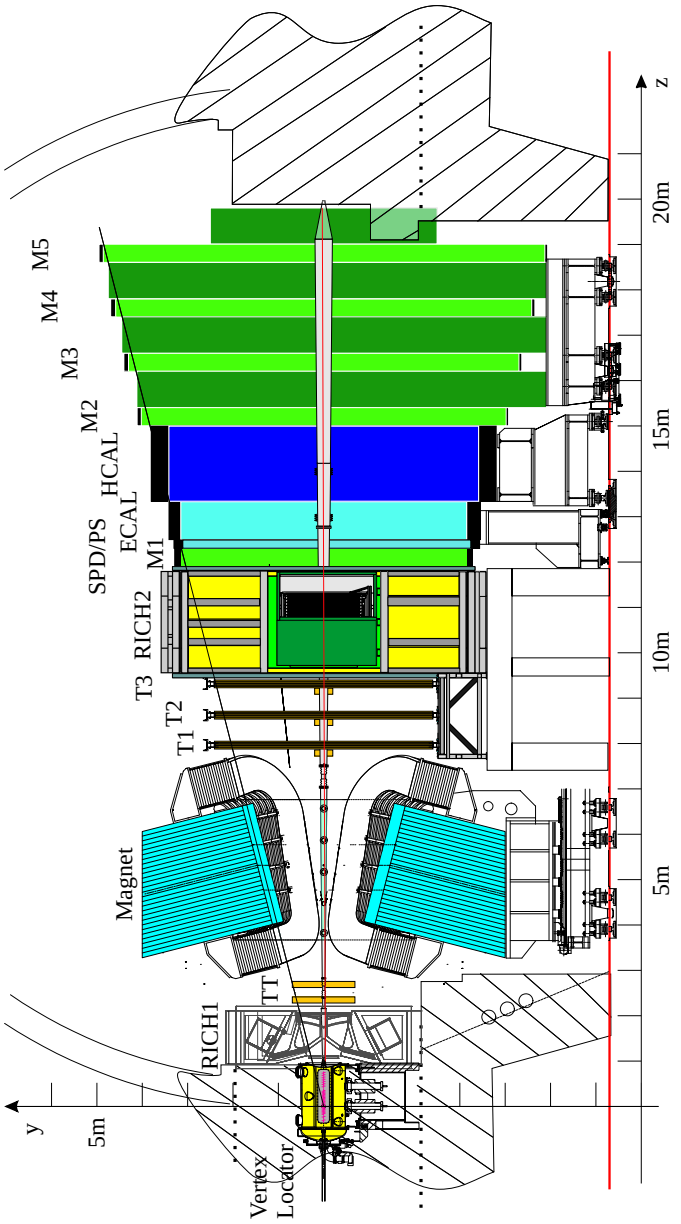

Figure 1: Schematic view of the LHCb re-optimized spectrometer in the y-z projection. Starting from the primary interaction region in the bottom part of the figure, the following components are visible: the VErtex LOcator (VELO), the first RICH detector (RICH1), the Trigger Tracking chamber (TT), the material-free magnetic field region, the Tracking chambers (T1, T2 and T3), the second RICH detector (RICH2), the first Muon chamber (M1), the Scintillating Pad and Preshower detectors (SPD/PS), the Electromagnetic CALorimeter (ECAL), the Hadronic CALorimeter (HCAL) and the remaining 4 Muon chambers (M2, M3, M4 and M5).

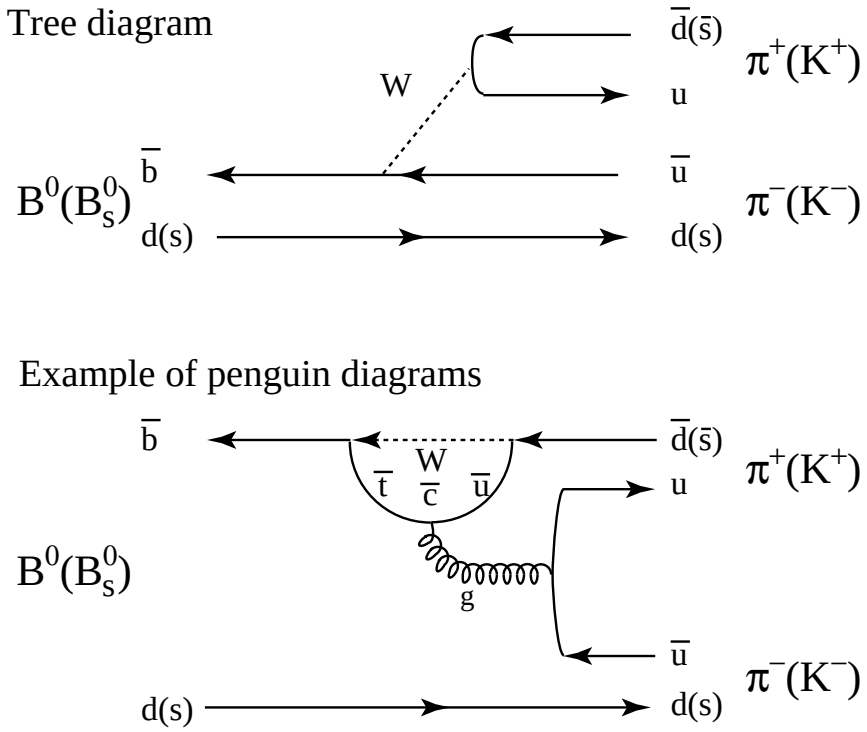

Figure 2: Tree and penguin diagrams generating $B^0 \to \pi^+\pi^-$ ($B_s^0 \to K^+K^-$) decays. By exchanging all the d ($\overline{d}$) quarks by the s ($\overline{s}$) quarks, the tree and penguin processes of the $B^0 \to \pi^+\pi^-$ generate those of the $B_s^0 \to K^+K^-$.

The violation of the CP symmetry in the neutral B-meson sector has been recently established at the asymmetric beauty factories by the BABAR and BELLE experiments, by direct measurements of the β angle of the unitarity triangle (UT), mainly through the "gold-plated" $B^0 \to J/\psi K_S^0$ decay [3, 4]. BABAR and BELLE have also provided first measurements of CP violation in the $B^0 \to \pi^+\pi^-$ decay [5, 6], although with limited statistics and large errors. The results of the two experiments are only weakly compatible at the moment, and it is not yet possible to draw any clear conclusion.

Since in addition to the $\overline{b} \to \overline{u} + W^+$ tree amplitude, large $\overline{b} \to \overline{d} + g\,(\gamma,\, Z^0)$ penguin amplitudes are expected to contribute to the $B^0 \to \pi^+\pi^-$ decay [7] (see Figure 2), it alone does not provide a clean strategy for measuring CKM phases. However, recent theoretical works [8, 9] have shown that the combined measurement of the $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ CP asymmetries, under the assumption of invariance of the strong interaction dynamics exchanging the $d \leftrightarrow s$ quarks (U-spin symmetry), provides a way to determine the γ angle. Moreover, alternative measurements of γ can be obtained from charged $B \to K\pi$ decays, although they have to be combined with $K\pi$ neutral modes [9]. The sensitivity that LHCb can achieve in measuring γ from the combined CP measurements of $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ is discussed in a separate LHCb note [10].

In the following sections, the LHCb performance in selecting $B^0_{(s)} \to h^+h^-$ decays is presented. In order to evaluate the overall signal efficiency and the level of residual backgrounds, signal and background events are first selected by means of the offline-selection algorithms, and then passed through the trigger simulation code.

2 Event generation, data samples and branching ratios

Proton-proton interactions are generated using PYTHIA 6.2 [11], considering hard QCD processes, single and double diffraction and elastic scattering. The occurrence of several proton-proton collisions in the same bunch-crossing is simulated (pile-up), and noise introduced by interactions happening in previous or subsequent bunch crossings (spill-over) is also taken into account. The decay of unstable particles is performed with the QQ program [12], including the B^0_s and b-baryon decays.

In order to increase the event generation efficiency, signal events are generated requiring both the b-hadrons to be inside the geometrical acceptance of the detector (forward direction and polar angle less than 400 $mrad$). For $b\bar{b}$ events, the same requirement is asked to be fulfilled at least by one b-hadron. This results in an efficiency $\epsilon^{signal}_{gen} = 0.3471 \pm 0.0003$ for signal events and $\epsilon^{b\bar{b}}_{gen} = 0.4321 \pm 0.0004$ for $b\bar{b}$ events, where the errors are statistical.

The detector response is simulated using the SICB FORTRAN-based program [13], which in turn makes use of the GEANT3 package [14]. The detector geometry and materials are described in detail, including with the active detector components and their front-end electronics, passive materials, such as the beam-pipe, infrastructural supports, etc.. Low-energy particles, mainly produced in secondary interactions inside the detector, are also traced applying a cut-off of 10 MeV for the hadrons and 1 MeV for electrons and photons. The event reconstruction is performed by the BRUNEL package [15], based on the GAUDI framework [16], written in C++.

The Monte Carlo mass-production was realized by means of the DIRAC distributed Monte Carlo production system [17], during the first LHCb Data Challenge, by using the computing facilities of 17 computing centres spread across Europe.

In this analysis two sources of background have been studied: the two-charged-body decays of b-hadrons, and combinatorial background coming from generic $b\bar{b}$ events, under the assumption that $b\bar{b}$ events are the dominant source of the combinatorial background, i.e., neglecting the contribute of $c\bar{c}$ and light-flavor events.

The specific decay data samples used to assess the performance of reconstruction, offline-selection, trigger and tagging are shown in Table 1. In the same table, the assumed branching ratios are reported as well. Table 2 shows the size of the $b\bar{b}$ data samples used to estimate the combinatorial background levels of the various channels under study[1].

The assumed hadronization fractions, i.e. the probabilities that a b quark hadronizes into a given b-hadron of interest, are shown in Table 3.

[1] All the mentioned specific decay samples and a relevant fraction of the $b\bar{b}$ event sample, have been generated and reconstructed employing the CPU power of the LHCb-Bologna computing cluster [18], hosted at the INFN-CNAF computing centre, in Bologna.

Channel	#Events	Branching Ratio$\times 10^6$	Comment
$B^0 \to \pi^+\pi^-$	263×10^3	4.8±0.5	measured
$B_s^0 \to K^+K^-$	299×10^3	18.5±1.2	estimated as $B^0 \to K^+\pi^-$
$B^0 \to K^+\pi^-$	50×10^3	18.5±1.2	measured
$B_s^0 \to K^-\pi^+$	50×10^3	4.8±0.5	estimated as $B^0 \to \pi^+\pi^-$
$\Lambda_b \to p\pi^-$	50×10^3	4.8±0.5	estimated as $B^0 \to \pi^+\pi^-$
$\Lambda_b \to pK^-$	50×10^3	18.5±1.2	estimated as $B^0 \to K^+\pi^-$

Table 1: Size of the Monte Carlo data samples used to perform the $B^0_{(s)} \to h^+h^-$ analyses. The branching ratios assumed to compute the signal yields and background-to-signal ratios are also shown (measured branching ratios are taken from [20]).

Channel	$\#b\bar{b}$ events
$B^0 \to \pi^+\pi^-$	9962100
$B_s^0 \to K^+K^-$	9971100
$B^0 \to K^+\pi^-$	8972700
$B_s^0 \to K^-\pi^+$	8964200

Table 2: Size of the generic $b\bar{b}$ Monte Carlo data samples used to estimate the combinatorial background levels for each of the $B^0_{(s)} \to h^+h^-$ analyses. The number of generic $b\bar{b}$ events is not the same for all channels since an independent pre-selection with very loose cuts was performed for each channel in order to reduce the data sample size and thus speed up the analyses. Not all the available statistics (about 10.5×10^6 events) and in any case not the same statistics was used for each pre-selection.

b-hadron	Production Fraction
B^0	0.391 ± 0.013
B_s^0	0.100 ± 0.013
Λ_b	0.118± 0.021

Table 3: Hadronization probabilities for the b-quark to form the b-hadrons of interest assumed in this analysis. The not yet measured probability $f(b \to \Lambda_b)$ is conservatively assumed to be equal to the total probability $f(b \to b-\text{baryons})$. The fractions are taken from [20].

3 Event selection

The offline-selection program is based on the LHCb analysis package DAVINCI [19], in turn based on the GAUDI framework [16], and it is written in C++. All the analyses we are going to describe, except the $b\bar{b}$ pre-selection which was centrally realized at CERN, have been performed using the LHCb-Bologna computing cluster [18].

The selection starts with the identification of reconstructed tracks as charged pions or kaons. The particle identification algorithm combines the information from different sub-detectors: the RICH system, the ECAL and HCAL calorimeters, and the Muon system. For the $B^0_{(s)} \to h^+h^-$ channels, RICH information are particularly relevant for pion/kaon discrimination, while the calorimeters and the Muon system information provide discrimination between pions, electrons and muons.

The analysis is performed for single and multiple interaction events. In case more than one primary vertex is reconstructed, the $b\bar{b}$ production vertex is chosen to be the one with respect to which the B candidate has the smallest impact parameter significance. The primary vertex reconstruction is performed using the algorithm described in [2].

A crucial requirement for these selections consists in suppressing B decays with the same two-track topology, which can mimic the signal in case of particle mis-identification. In fact, these channels may exhibit their own CP asymmetries, thus biasing the CP measurement of the channel under consideration. In this purpose, LHCb can rely on the performance of the RICH particle identification system [2, 21], and on a rather good B-meson mass resolution, of about 17 MeV/c^2 for the $B^0_{(s)} \to h^+h^-$ channels.

3.1 Definition of selection cuts

The selection parameters used to isolate the signal decays are described in the following (see also Figure 3). Each charged track identified as a pion or a kaon is requested to have:

- a momentum p in the range $p_{min} < p < p_{max}$;
- a transverse momentum p_T larger than $(p_T)_{each}$;
- an impact parameter significance with respect to the primary vertex IP/σ_{IP} larger than $(IP/\sigma_{IP})_{each}$.

Pairs of tracks with opposite charges are formed requiring:

- at least one track with p_T larger than $(p_T)_{one}$;
- at least one track with IP/σ_{IP} larger than $(IP/\sigma_{IP})_{one}$.

Each pair surviving these conditions is fitted to a common vertex and used to form a $B^0_{(s)}$ candidate. The $B^0_{(s)}$ candidate must satisfy the following selection criteria:

- the vertex χ^2 must be smaller than χ^2_{max};

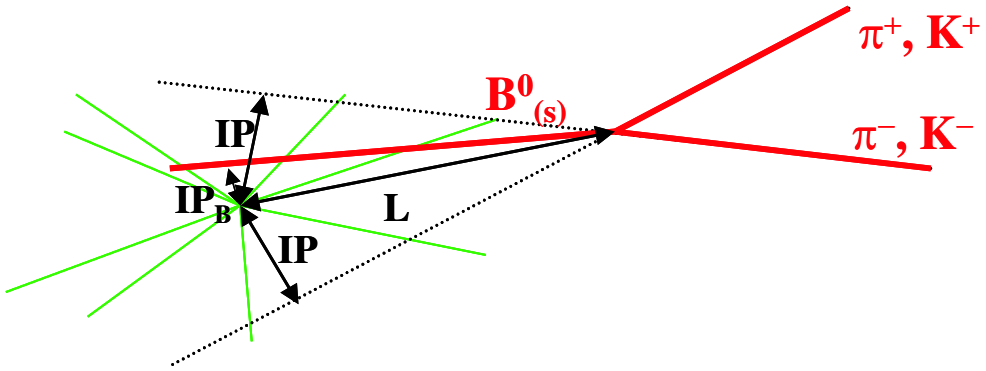

Figure 3: Schematic representation of the geometrical selection variables employed to isolate the $B^0_{(s)} \to h^+h^-$ decays from $b\bar{b}$ combinatorial background. Their meaning is described in the text.

- the p_T of the $B^0_{(s)}$ must be greater than $\left(p_T^B\right)_{min}$;
- the direction of flight must point to the primary vertex, with an impact parameter significance smaller than $(IP_B/\sigma_{IP_B})_{max}$;
- the distance of flight L, between the primary and the secondary vertex, must have a significance exceeding $(L/\sigma_L)_{min}$;
- the invariant mass of the candidate must lie within a $\pm\delta m$ of the nominal $B^0_{(s)}$ mass.

3.2 Cut value optimization

The 11 selection cuts have been optimized separately for each of the decays under study, in order to maximize the signal efficiency and minimize the backgrounds. In order to identify the best cut values, a multi-variate grid scan on the 9-dimension parameter space (i.e., excluding the momentum cuts) has been performed, by maximizing the function $S/\sqrt{S+B}$, where S is the signal event yield and B is the generic $b\bar{b}$ event yield. Due to the limited $b\bar{b}$ Monte Carlo statistics available, the mass cut on $b\bar{b}$ events has been relaxed to $\pm 0.6\,GeV/c^2$ ($\pm 1.2\,GeV/c^2$ for the $B^0_s \to K^-\pi^+$) with the aim of enlarging the effective $b\bar{b}$ statistics of about a factor $(0.6\,GeV/c^2)/\delta m$ ($(1.2\,GeV/c^2)/\delta m$ for the $B^0_s \to K^-\pi^+$), under the assumption of the independence of the mass value from the other selection variables. Nevertheless, only a handful of $b\bar{b}$ events survive the selections. Then, the optimal cut values obtained with this technique have been further en-tightened in order to remove possible modelization of the cuts on the actual $b\bar{b}$ data sample used, this way limiting the possibility of an underestimation of the background-to-signal ratios. In addition, a separate scan on the parameters p_{min}, p_{max} and δm has been performed, again maximizing the function $S/\sqrt{S+B}$, where S is the signal event yield and B is the specific backgrounds

Channel	Two body backgrounds
$B^0 \to \pi^+\pi^-$	$B_s^0 \to K^+K^-$, $B^0 \to K^+\pi^-$, $B_s^0 \to K^-\pi^+$, $\Lambda_b \to p\pi^-$, $\Lambda_b \to pK^-$
$B_s^0 \to K^+K^-$	$B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B_s^0 \to K^-\pi^+$, $\Lambda_b \to p\pi^-$, $\Lambda_b \to pK^-$
$B^0 \to K^+\pi^-$	$B^0 \to \pi^+\pi^-$, $B_s^0 \to K^+K^-$, $B_s^0 \to K^-\pi^+$, $\Lambda_b \to p\pi^-$, $\Lambda_b \to pK^-$
$B_s^0 \to K^-\pi^+$	$B^0 \to \pi^+\pi^-$, $B_s^0 \to K^+K^-$, $B^0 \to K^+\pi^-$, $\Lambda_b \to p\pi^-$, $\Lambda_b \to pK^-$

Table 4: Two body decays considered as backgrounds for each $B^0_{(s)} \to h^+h^-$ channel.

event yield. This is because specific two-body background decays are topologically similar to the signal and can survive the selection in case of particle mis-identification. The particle identification performance depends on the momentum of the particles themselves, and the mis-identification of one particle of the pair has the effect of shifting the invariant mass value. The two-body decays considered for each signal under study are shown in Table 4. The so determined optimal offline-selection cuts are shown in Table 5.

The distributions of all the selection variables (except momentum which is irrelevant for $b\bar{b}$ background), for signal and $b\bar{b}$ background events of all the channels under study, are shown in the Figures 4, 5, 6 and 7.

3.3 Trigger selection

Events surviving the offline-selection are then passed through the L0 and L1 trigger simulation algorithms. The L0 trigger reduces the event rate from a bunch-crossing frequency of 40 *MHz* to about 1 *MHz*. The L1 has to further reduce the rate to 40 *KHz* [2]. For a detailed description of the L0 and L1 working principles and tuning criteria see [22].

3.4 Results

A summary of the number of events surviving the offline-selection is shown in Table 6 (for the initial statistics see Tables 1 and 2). The same numbers, after offline-selection, L0 and L1 trigger filters, are reported in Table 7.

Figure 8 shows the invariant mass spectra of the four $B^0_{(s)} \to h^+h^-$ channels after trigger and offline-selection, together with the respective two-body backgrounds. A Gaussian

[2]The LHCb trigger project includes a third level, commonly referred to as High Level Trigger (HLT) [22]. The HLT algorithm is under development, but preliminary studies on $B^0_{(s)} \to h^+h^-$ channels, not discussed here, have shown that a further reduction of the event rate to about 12 *Hz* (i.e., the bandwidth allocated to $B^0_{(s)} \to h^+h^-$channels) is achievable, while maintaining an efficiency on $B^0_{(s)} \to h^+h^-$ signals of the order of 100%.

Channel	$B^0 \to h^+h^-$		$B_s^0 \to h^+h^-$	
	$\pi\pi$	$K\pi$	KK	πK
$p_{min}\,[GeV/c]$	2.50	2.75	2.75	2.75
$p_{max}\,[GeV/c]$	100	200	125	100
$(p_T)_{each}\,[GeV/c]$	1.2	1.2	0.8	1.4
$(p_T)_{one}\,[GeV/c]$	3.2	3.0	2.6	3.4
$(IP/\sigma_{IP})_{each}$	6	6	5	7
$(IP/\sigma_{IP})_{one}$	12	11	9	14
χ^2_{max}	4	5	5	4
$\left(p_T^B\right)_{min}\,[GeV/c]$	1.6	1.4	1.0	1.6
$(IP_B/\sigma_{IP_B})_{max}$	2.25	2.50	2.75	2.25
$(L/\sigma_L)_{min}$	19	17	14	20
$\delta m\,[MeV/c^2]$	50	50	50	40

Table 5: Summary of the selection cuts for the different $B^0_{(s)} \to h^+h^-$ offline analyses. The meaning of the cuts is described in the text.

Event type	Offline selection algorithm			
	$B^0 \to \pi^+\pi^-$	$B_s^0 \to K^+K^-$	$B^0 \to K^+\pi^-$	$B_s^0 \to K^-\pi^+$
$B^0 \to \pi^+\pi^-$	**15514**	14	1343	110
$B_s^0 \to K^+K^-$	82	**27318**	1588	684
$B^0 \to K^+\pi^-$	89	35	**4061**	22
$B_s^0 \to K^-\pi^+$	50	79	94	**2154**
$\Lambda_b \to p\pi^-$	7	0	17	19
$\Lambda_b \to pK^-$	0	26	4	2
$b\bar{b} \to inclusive$	9	10	16	6

Table 6: Number of events surviving the offline-selection. The sizes of the original data samples are shown in Tables 1 and 2.

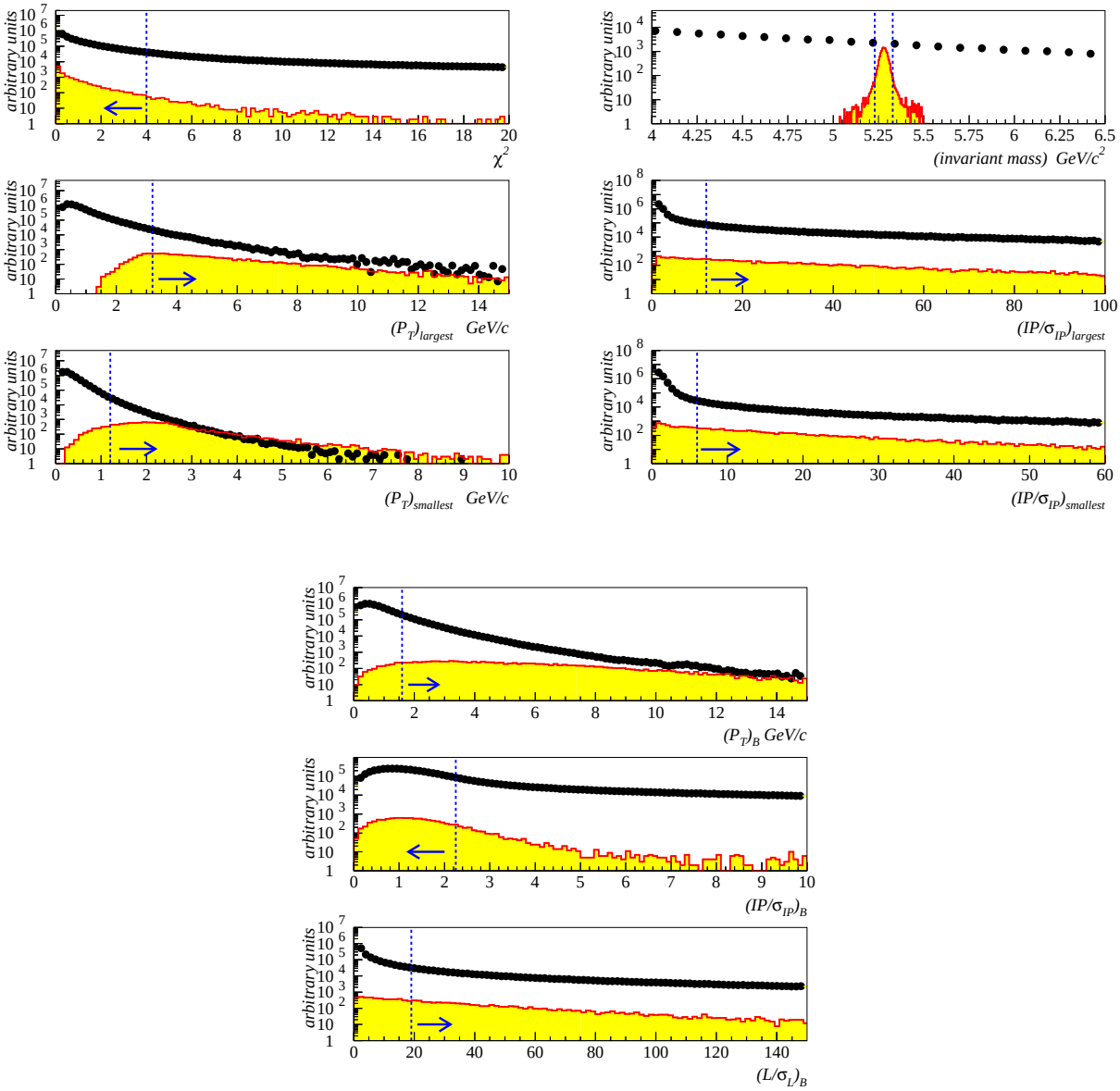

Figure 4: Distributions of all the selection variables for true $B^0 \to \pi^+\pi^-$ events (light histograms) and $b\bar{b}$ background combinatorics (filled circles). The plots are obtained after track reconstruction and particle identification, but before trigger and offline-selection. Vertical dashed lines indicate the cut value, while the arrows indicate the accepted region. The meaning of each cut is explained in the text. Plots of the "smallest" and "largest" variables show the distributions of the smallest and largest respective variables of the track pair. In this case, the shown cut values correspond respectively to the ones labeled "each" and "one" in Table 5.

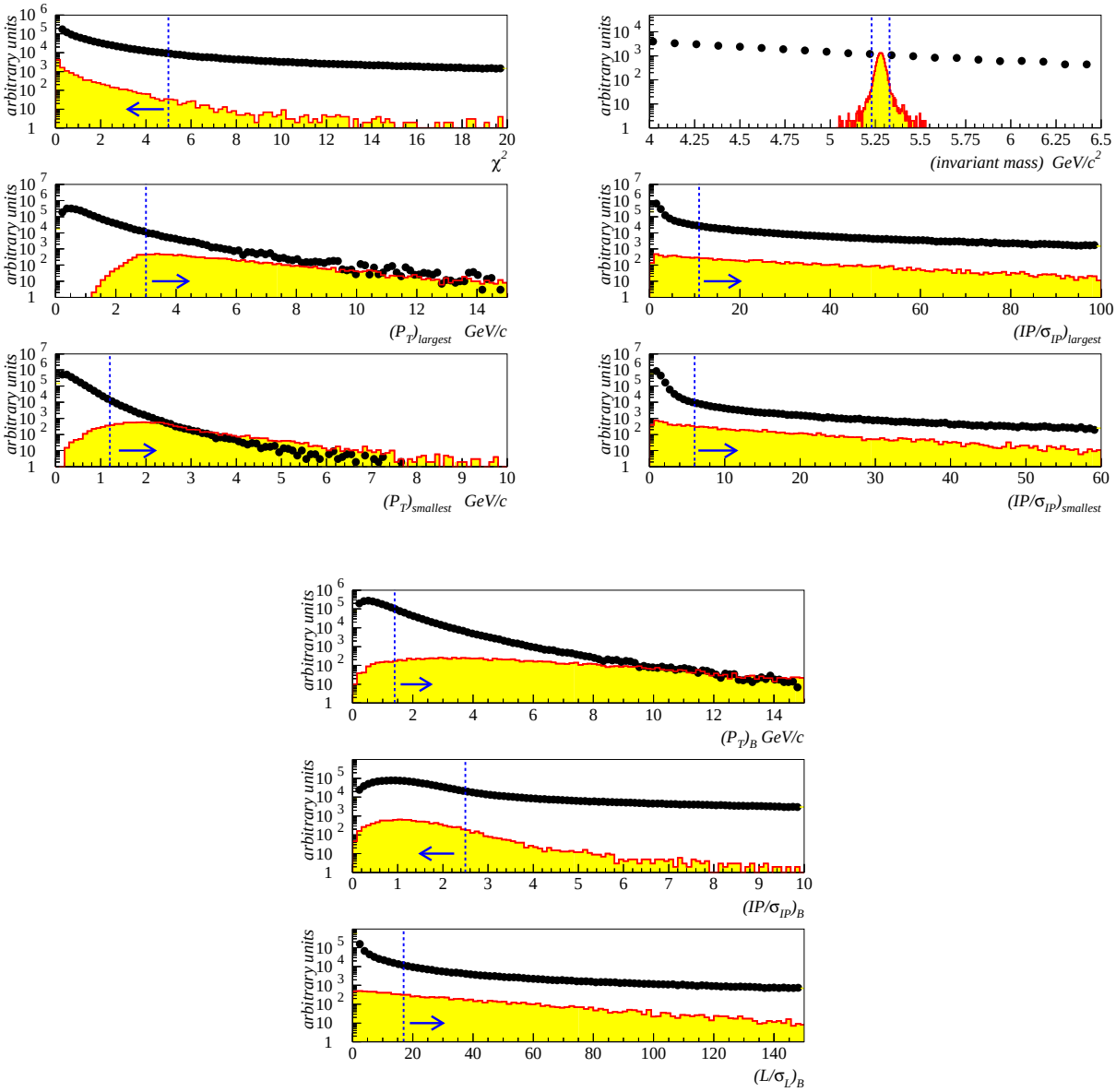

Figure 5: Distributions of all the selection variables for true $B^0 \to K^+\pi^-$ events (light histograms) and $b\bar{b}$ background combinatorics (filled circles). The plots are obtained after track reconstruction and particle identification, but before trigger and offline-selection. Vertical dashed lines indicate the cut value, while the arrows indicate the accepted region. The meaning of each cut is explained in the text. Plots of the "smallest" and "largest" variables show the distributions of the smallest and largest respective variables of the track pair. In this case, the shown cut values correspond respectively to the ones labeled "each" and "one" in Table 5.

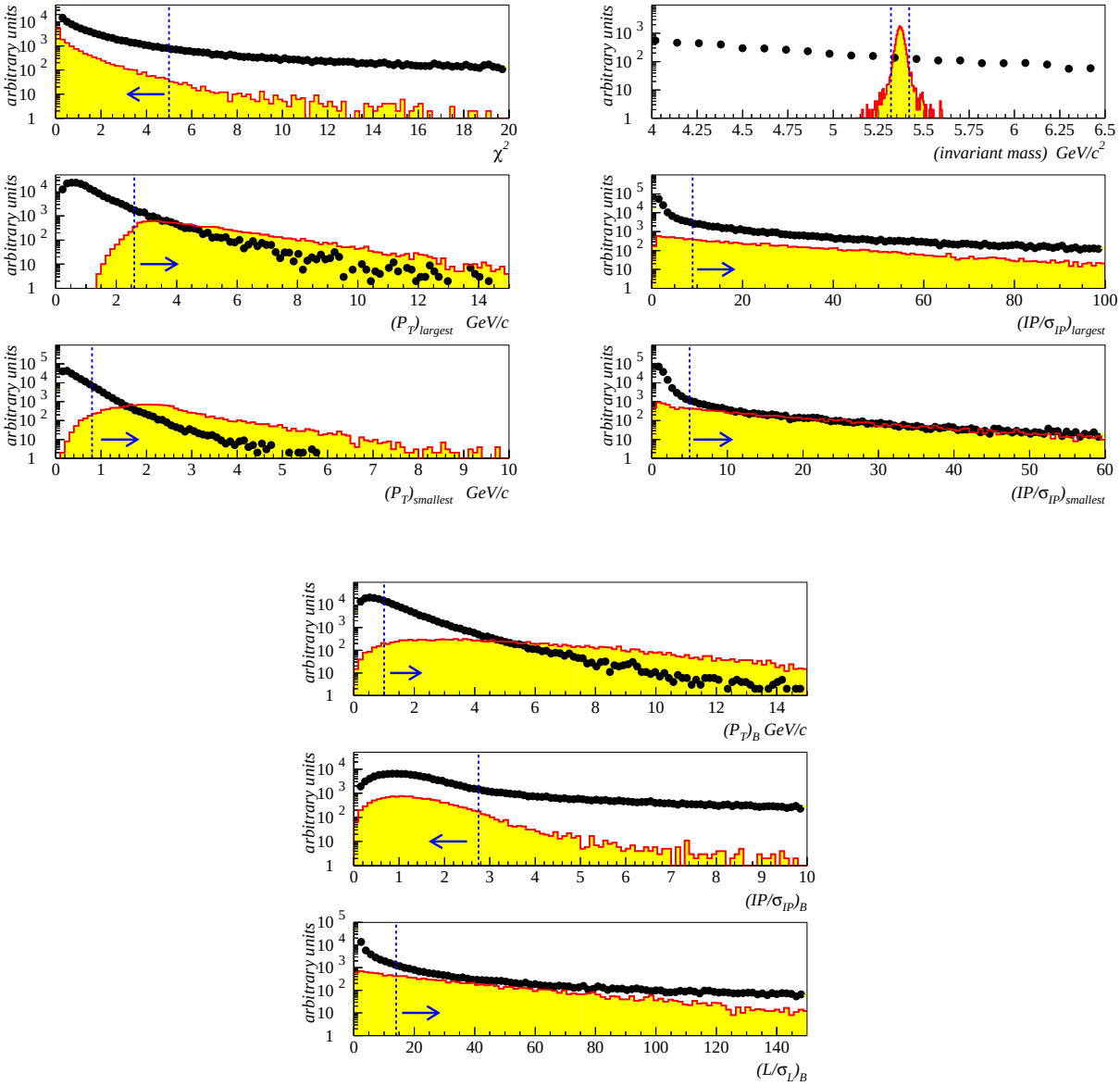

Figure 6: Distributions of all the selection variables for true $B_s^0 \to K^+K^-$ events (light histograms) and $b\bar{b}$ background combinatorics (filled circles). The plots are obtained after track reconstruction and particle identification, but before trigger and offline-selection. Vertical dashed lines indicate the cut value, while the arrows indicate the accepted region. The meaning of each cut is explained in the text. Plots of the "smallest" and "largest" variables show the distributions of the smallest and largest respective variables of the track pair. In this case, the shown cut values correspond respectively to the ones labeled "each" and "one" in Table 5.

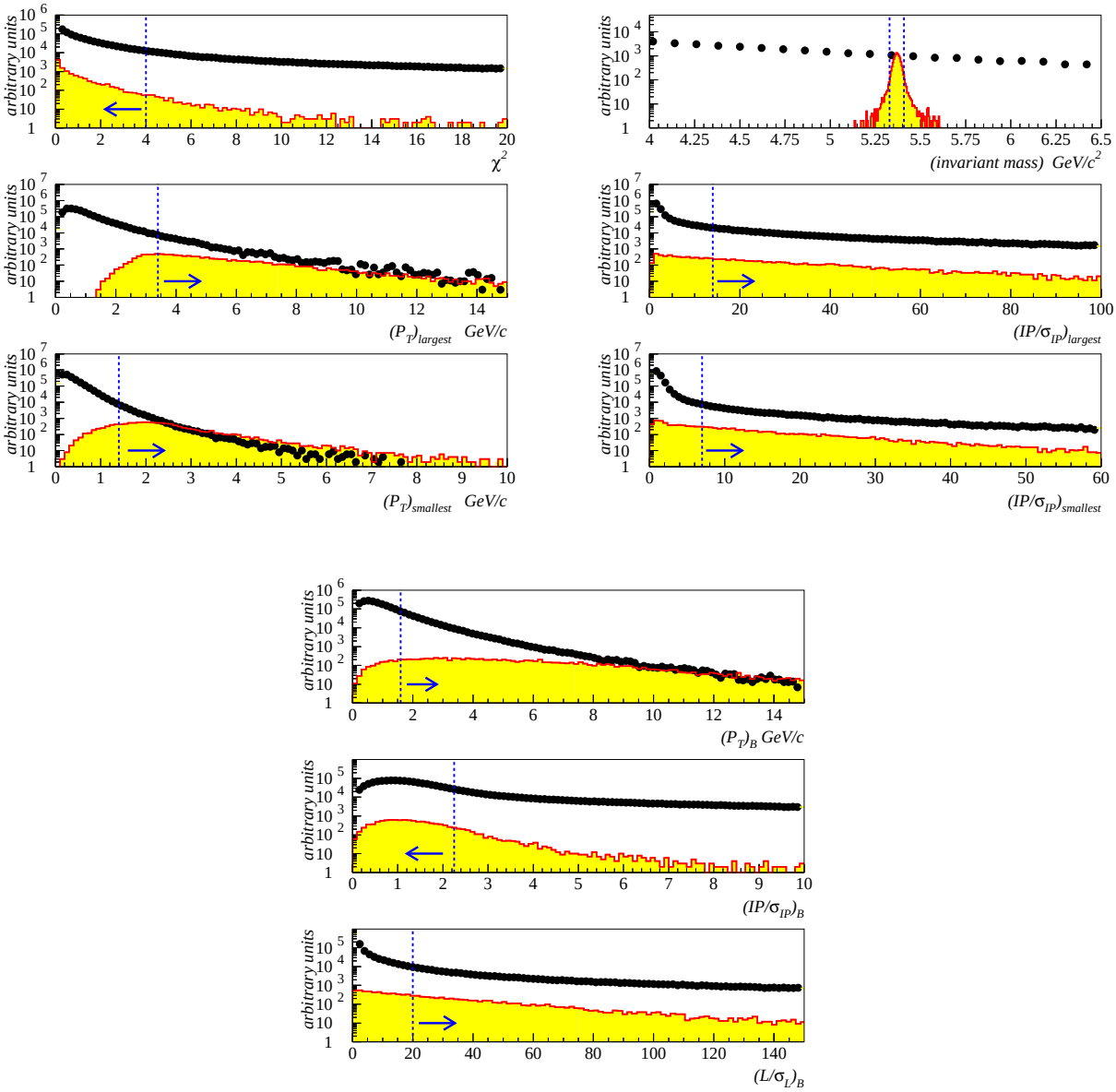

Figure 7: Distributions of all the selection variables for true $B_s^0 \to \pi^+ K^-$ events (light histograms) and $b\bar{b}$ background combinatorics (filled circles). The plots are obtained after track reconstruction and particle identification, but before trigger and offline-selection. Vertical dashed lines indicate the cut value, while the arrows indicate the accepted region. The meaning of each cut is explained in the text. Plots of the "smallest" and "largest" variables show the distributions of the smallest and largest respective variables of the track pair. In this case, the shown cut values correspond respectively to the ones labeled "each" and "one" in Table 5.

Event type	Offline selection algorithm			
	$B^0 \to \pi^+\pi^-$	$B_s^0 \to K^+K^-$	$B^0 \to K^+\pi^-$	$B_s^0 \to K^-\pi^+$
$B^0 \to \pi^+\pi^-$	**8311 (5211)**	9 (9)	820 (495)	61 (44)
$B_s^0 \to K^+K^-$	26 (20)	**14156 (8499)**	962 (628)	452 (270)
$B^0 \to K^+\pi^-$	47 (28)	19 (9)	**2195 (1349)**	7 (5)
$B_s^0 \to K^-\pi^+$	32 (19)	46 (24)	56 (25)	**1216 (790)**
$\Lambda_b \to p\pi^-$	3 (1)	0 (0)	9 (4)	7 (2)
$\Lambda_b \to pK^-$	0 (0)	15 (6)	1 (0)	1 (1)
$b\bar{b} \to inclusive$	4 (2)	3 (0)	8 (5)	3 (1)

Table 7: Number of events surviving the offline-selection and the L0 (and L1) trigger. The sizes of the original data samples are shown in Tables 1 and 2.

fit to the signal spectra gives a resolution of about $17.5\,MeV/c^2$ for all the four channels. The background spectra have been weighted according to assumed branching ratios, hadronization fractions and Monte Carlo statistics. As it can be seen from the plots, by suitable choice of the δm cut, this source of background can be kept relatively small. The corresponding background-to-signal ratios are reported in the next section.

4 Selection efficiencies, event yields and B/S estimates

The total signal efficiency is calculated as the fraction of signal events that are triggered (by L0 and L1), reconstructed and selected with offline-selection cuts. It can be factorized as:

$$\epsilon_{tot} = \epsilon_{det} \times \epsilon_{rec/det} \times \epsilon_{sel/rec} \times \epsilon_{trig/sel} \tag{4.1}$$

where ϵ_{det} is the detection efficiency (including geometrical acceptance in 4π solid angle and material effects in the detector), $\epsilon_{rec/det}$ is the reconstruction efficiency with respect to detected events, $\epsilon_{sel/rec}$ is the efficiency of the offline-selection with respect to reconstructed events, and $\epsilon_{trig/sel}$ is the (L0 and L1) trigger efficiency with respect to offline-selected events.

The annual signal and specific background yields are then computed as:

$$S = L_{int} \times \sigma_{b\bar{b}} \times 2 \times f_B \times BR \times \epsilon_{tot} \tag{4.2}$$

where $L_{int} = 2\,fb^{-1}$ is the annual integrated luminosity (one year of data taking is assumed to be $10^7\,s$, and the instantaneous luminosity is assumed to have the nominal value $\mathcal{L} = 2 \times 10^{32}\,cm^{-2}s^{-1}$), $\sigma_{b\bar{b}} = 500\,\mu b$ is the $b\bar{b}$ production cross-section expected at $14\,TeV$, f_B is the production fraction of the B-meson of interest (see Table 3), the factor 2 takes into account the production of both b-hadrons and $\bar{b}$-hadrons and BR is the branching ratio of

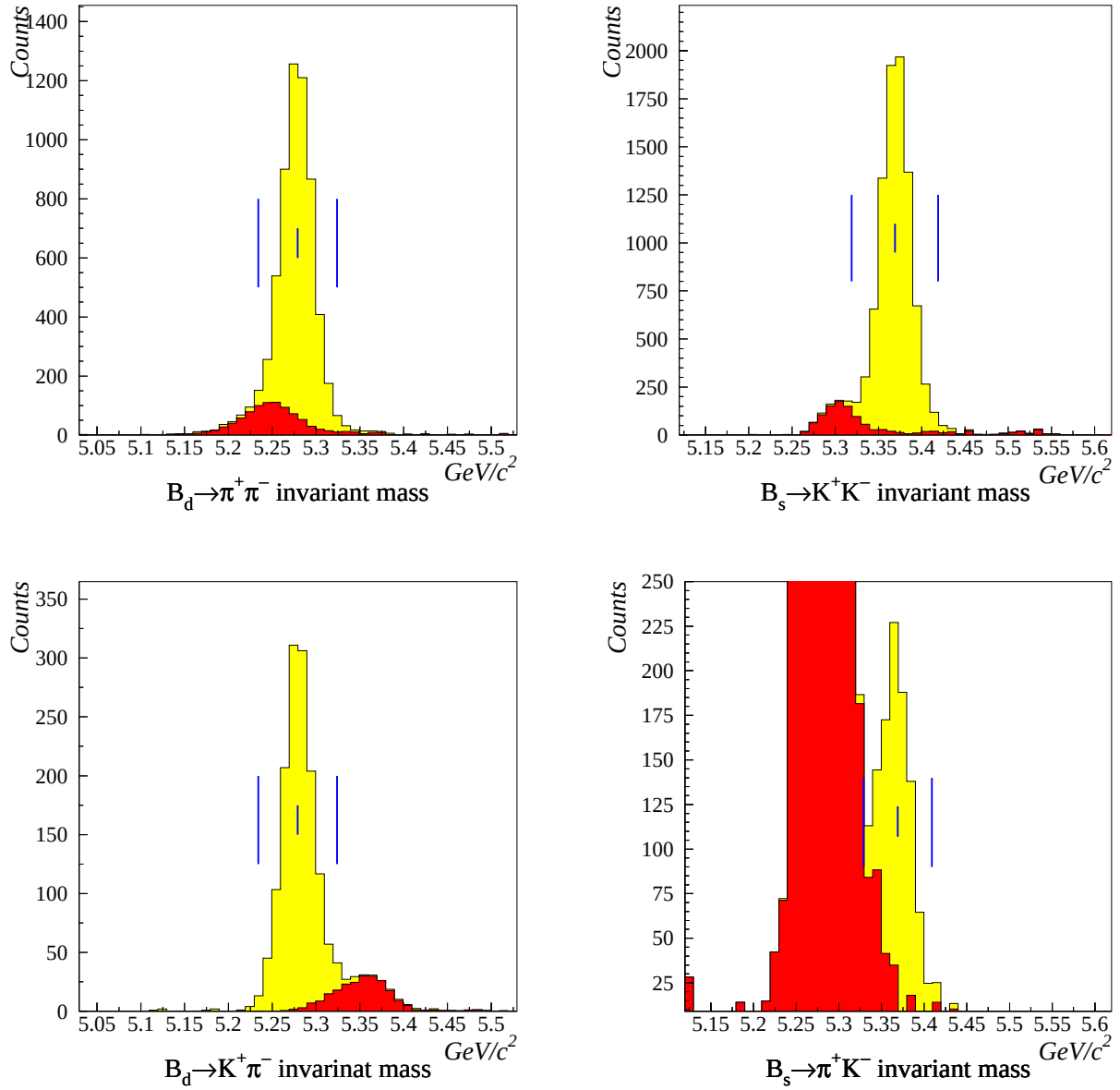

Figure 8: Cumulative invariant mass spectra of the four $B^0_{(s)} \to h^+h^-$ channels after trigger and offline-selection (light histograms), together with the respective two-body backgrounds (dark histograms). The accepted mass region is indicated by vertical bars. The huge background in the left sideband of the $B^0_s \to K^-\pi^+$ plot is largely due to the $B^0 \to K^+\pi^-$ decay, which has about a factor 15 of larger production probability and same $K\pi$ signature.

Channel	Factors (in %) forming ϵ_{tot} (in %) $\epsilon_{det} \times \epsilon_{rec/det} \times \epsilon_{sel/rec} \times \epsilon_{trig/sel} = \epsilon_{tot}$					Annual signal yield	B/S from spec. back.	B/S from incl. $b\bar{b}$ back.
	ϵ_{det}	$\epsilon_{rec/det}$	$\epsilon_{sel/rec}$	$\epsilon_{trig/sel}$	ϵ_{tot}			
$B^0 \to \pi^+\pi^-$	12.2	91.6	18.3	33.6	0.688	26 k	0.13	< 0.72
$B_s^0 \to K^+K^-$	12.0	92.5	28.6	31.1	0.988	37 k	0.04	< 0.51
$B^0 \to K^+\pi^-$	12.2	92.0	25.2	33.2	0.94	135 k	0.04	< 0.22
$B_s^0 \to K^-\pi^+$	12.0	92.1	13.5	36.7	0.548	5.3 k	0.41	< 1.28

Table 8: Summary of signal efficiencies, untagged annual signal yields and background-to-signal ratios (B/S) from specific and generic $b\bar{b}$ backgrounds. B/S ratios are computed without applying the trigger in order to increase the effective statistics. B/S from $b\bar{b}$ background are quoted as 90% C.L. upper limits.

the B-meson decay under study. In order to increase the effective background statistics, the B/S ratios are estimated without applying the trigger.

All the efficiencies, signal yields and B/S ratios for specific and combinatorial $b\bar{b}$ backgrounds for each $B^0_{(s)} \to h^+h^-$ channel are summarized in Table 8. It can be noted that:

$$\epsilon_{tot}(B_s^0 \to K^+K^-) > \epsilon_{tot}(B^0 \to K^+\pi^-) > \epsilon_{tot}(B^0 \to \pi^+\pi^-) > \epsilon_{tot}(B_s^0 \to \pi^+K^-). \quad (4.3)$$

This reflects the fact that increasing tighter cuts are used in the $B_s^0 \to K^+K^-$, $B^0 \to K^+\pi^-$, $B^0 \to \pi^+\pi^-$ and $B_s^0 \to \pi^+K^-$ offline-selections (see Table 5), in order to isolate signal events from combinatorial $b\bar{b}$ background. The difference in the selection cuts are due to the combined effects of a larger presence of pion tracks in the background, with respect to kaons, and of different branching ratios of the various decays (see Table 1), as well as different hadronization fractions of the respective B-mesons (see Table 3).

5 Tagging efficiencies

Flavor tagging is necessary in order to study the CP asymmetries and flavor oscillations of the channels under study. For a detailed description of the flavor tagging algorithm, see [23]. Here we only give a summary of the results for the $B^0_{(s)} \to h^+h^-$ decay channels.

The statistical uncertainty on the measured CP asymmetries is related to the effective tagging power ϵ_{eff}, which is defined as:

$$\epsilon_{eff} = \epsilon_{tag}(1 - 2\omega)^2 \quad (5.1)$$

where ϵ_{tag} is the probability that the tagging procedure gives an answer (tagging efficiency), while ω is the probability that this answer is wrong (wrong tagging fraction or mistag probability).

Table 9 summarizes the performance of the tagging procedure, combining the results of the different tagging categories, for the signal decays under study.

Channel	ϵ_{tag} (%)	ω (%)	ϵ_{eff} (%)
$B^0 \to \pi^+\pi^-$	41.8 ± 0.7	34.9 ± 1.1	3.8 ± 0.5
$B_s^0 \to K^+K^-$	49.8 ± 0.5	33.0 ± 0.8	5.8 ± 0.5
$B^0 \to K^+\pi^-$	43.2 ± 1.4	33.3 ± 2.1	4.8 ± 1.0
$B_s^0 \to K^-\pi^+$	49.5 ± 1.8	30.4 ± 2.6	7.6 ± 1.7

Table 9: Performance of the combined tag for the $B^0_{(s)} \to h^+h^-$ offline-selected and triggered decays. Shown uncertainties are statistical.

6 Miscellaneous distributions

In this section we show and discuss some distributions obtained as sub-products of this analysis, which are also relevant to study the LHCb sensitivity on the CP-violating observables of $B^0_{(s)} \to h^+h^-$ decays [10].

6.1 Signal distributions

Figure 9 shows the resolution on the three coordinates of the $b\bar{b}$ production vertex in signal event samples. Fits of double Gaussians to the distributions give a core resolution of about $8\,\mu m$ for x and y, and $45\,\mu m$ for z (75% of events in core Gaussian, second Gaussian is 2 to 3 times wider). The resolution on z is slightly asymmetric and biased. This is due to decay products of b-hadrons that are included in the primary vertex finding [2]. Such a bias is not present in minimum bias events. In principle, it can be eliminated by removing the B decay products and re-fitting the primary vertex, once the B final states have been reconstructed.

Figure 10 shows the resolution on the three coordinates of the $B^0_{(s)}$ decay vertex. Fits of double Gaussians to the distributions give a core resolution of about $11\,\mu m$ for x and y, and $90\,\mu m$ for z (85% of events in core Gaussian, second Gaussian is 2 to 2.5 times wider). Differently from the primary vertex case, no bias on the secondary vertex position is observed.

The proper time resolution is a very important quantity, in particular for resolving the fast oscillations of the $B_s^0 - \bar{B}_s^0$ system, where a bad resolution can dilute the sensitivity on the CP measurements. It is defined as the difference between the reconstructed proper decay time and the true one:

$$\Delta t = t_{rec} - t_{true} \tag{6.1}$$

The Δt distributions for triggered and offline-selected events for each $B^0_{(s)} \to h^+h^-$ channel are shown in Figure 11. The distributions are well fitted by double Gaussians. Results of the fits are summarized in Table 10. The parameters in Table 10 correspond to the following parametrization of the proper time resolution:

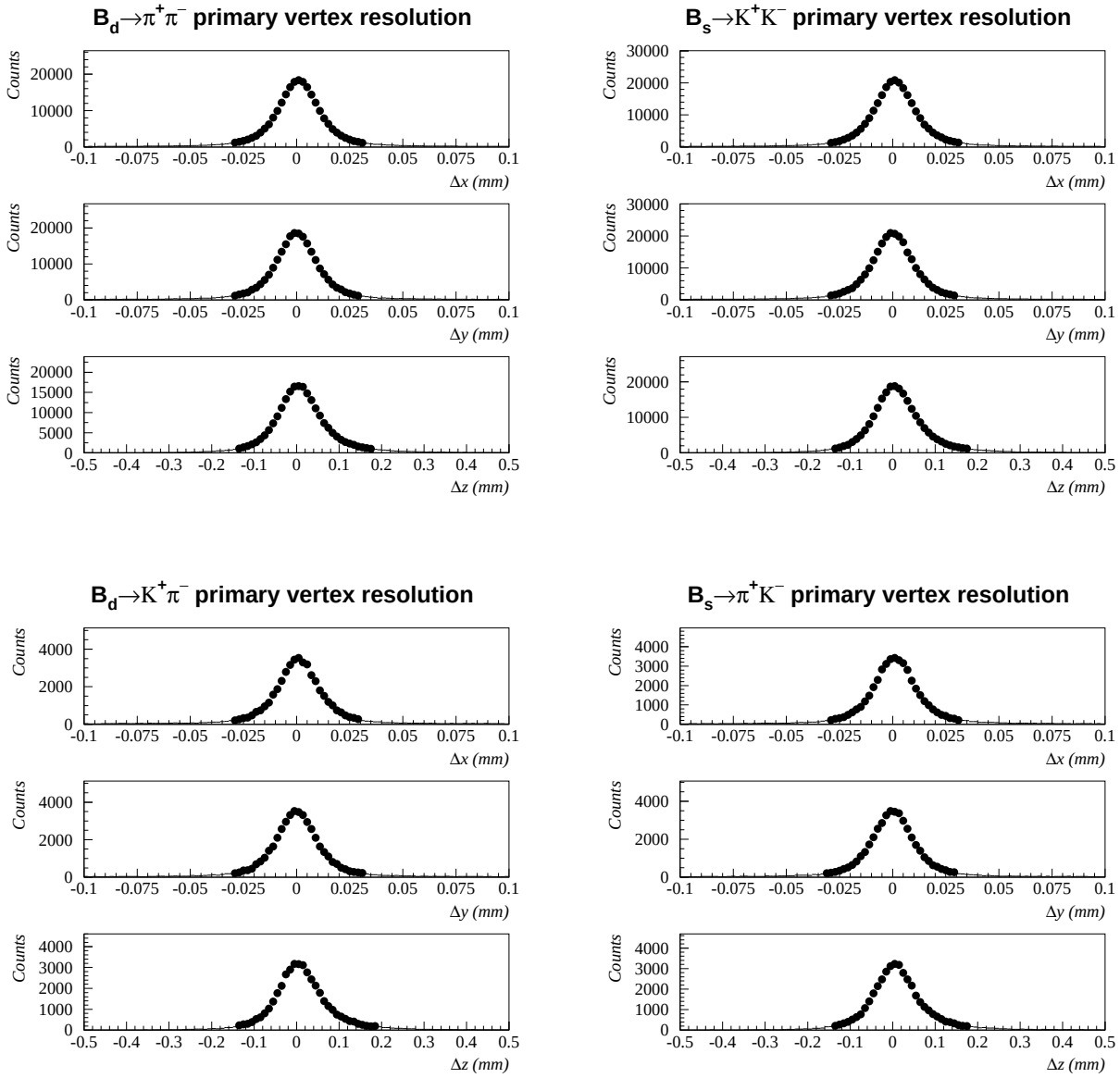

Figure 9: Resolutions on the x, y and z coordinates of the $b\bar{b}$ production vertex in each $B^0_{(s)} \to h^+h^-$ event sample.

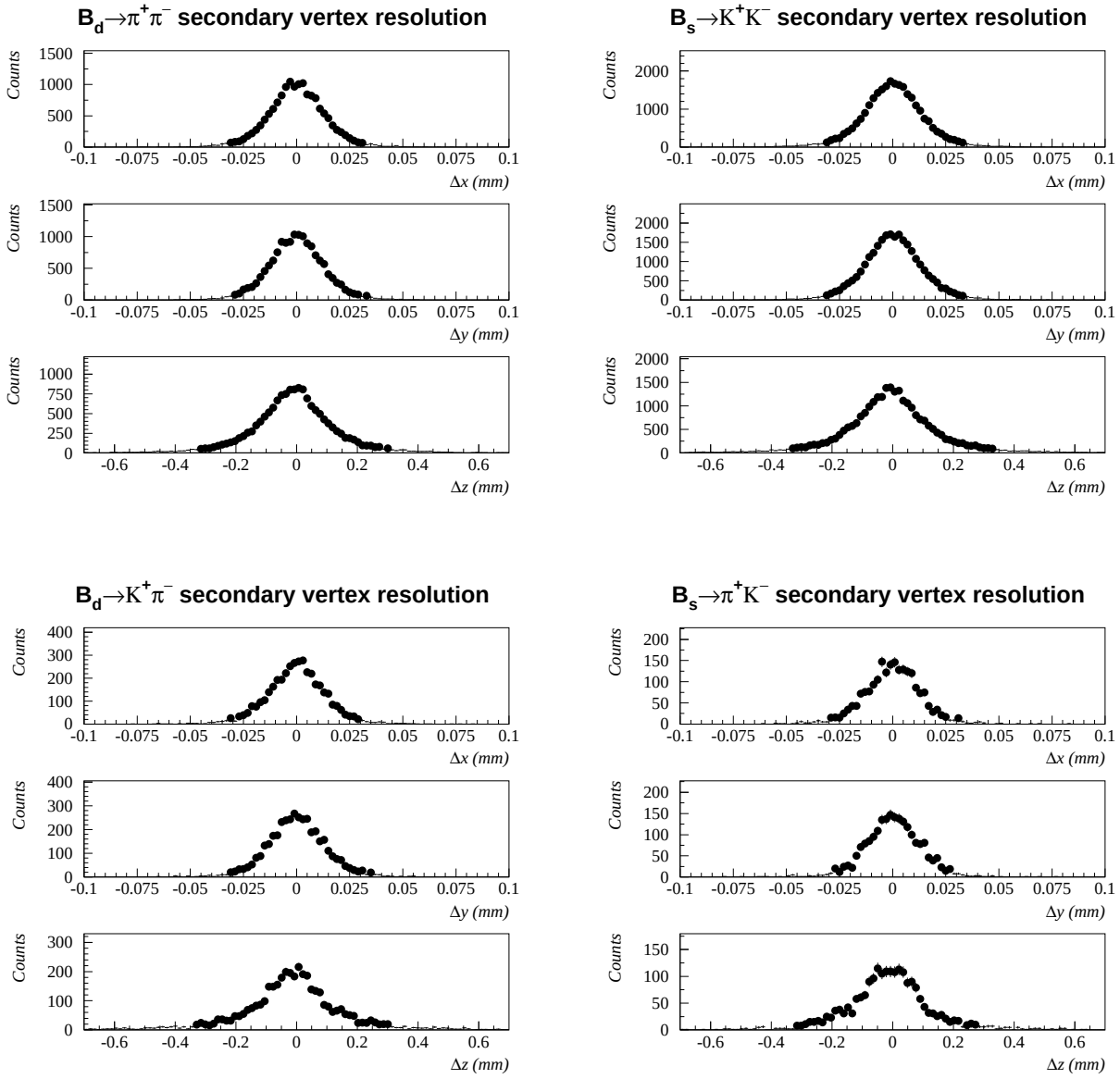

Figure 10: Resolutions on the x, y and z coordinates of the $B^0_{(s)}$ decay vertex for offline-selected events in each $B^0_{(s)} \to h^+h^-$ event sample.

Channel	f (in %)	$\overline{\Delta t}$ (fs)	$\sigma_{\Delta t_1}$ (fs)	$\sigma_{\Delta t_2}$ (fs)
$B^0 \to \pi^+\pi^-$	86 ± 5	-4.9 ± 0.5	32.6 ± 1.3	59.9 ± 4.2
$B_s^0 \to K^+K^-$	85 ± 3	-4.6 ± 0.5	34.4 ± 1.0	67.3 ± 3.1
$B^0 \to K^+\pi^-$	88 ± 9	-5.3 ± 1.0	32.8 ± 2.3	60.6 ± 9.0
$B_s^0 \to K^-\pi^+$	80 ± 12	-5.5 ± 1.3	29.8 ± 3.0	54.1 ± 6.9

Table 10: Results of double Gaussian fit to the proper time resolution for triggered and offline-selected events of each $B^0_{(s)} \to h^+h^-$ channel. The meaning of each parameter is described in the text. Shown uncertainties are statistical.

Channel	α	β
$B^0 \to \pi^+\pi^-$	$(8.05 \pm 0.11) \times 10^{-2}$	-2.35 ± 0.08
$B_s^0 \to K^+K^-$	$(12.07 \pm 0.14) \times 10^{-2}$	-1.76 ± 0.03
$B^0 \to K^+\pi^-$	$(10.84 \pm 0.31) \times 10^{-2}$	-2.28 ± 0.12
$B_s^0 \to K^-\pi^+$	$(6.24 \pm 0.23) \times 10^{-2}$	-2.52 ± 0.18

Table 11: Results of fit to the proper time acceptance for triggered and offline-selected events of each $B^0_{(s)} \to h^+h^-$ channel. The meaning of each parameter is described in the text. Shown uncertainties are statistical.

$$G(\Delta t) = \frac{f}{\sqrt{2\pi}\sigma_{\Delta t_1}} \exp\left(-\frac{(\Delta t - \overline{\Delta t})^2}{2\sigma^2_{\Delta t_1}}\right) + \frac{(1-f)}{\sqrt{2\pi}\sigma_{\Delta t_2}} \exp\left(-\frac{(\Delta t - \overline{\Delta t})^2}{2\sigma^2_{\Delta t_2}}\right) \tag{6.2}$$

As it can be seen, the results for each channel under study are each other compatible within statistical errors. Fits with single Gaussians give a resolution $\sigma \simeq 40\, fs$. The fits show the presence of a small bias, reflecting the bias in the primary vertex position mentioned above.

The dependence of Δt on the true proper time has been studied in detail. Figure 12 shows the dependence of the average mean, average standard deviation and average estimated error of Δt, as a function of the true proper time, for triggered and offline-selected events of each $B^0_{(s)} \to h^+h^-$ channel. As it can be seen, only a moderate dependence on the true proper time is observed.

Finally, a relevant quantity that has been studied is the signal acceptance for triggered and offline-selected events as a function of the proper decay time, defined as the untagged measured proper time distribution divided by the generated one. By studying the shape of this function for each $B^0_{(s)} \to h^+h^-$ decay, shown in Figure 13, it comes out that it can be well parametrized as:

$$\epsilon(t) = \frac{\alpha}{1 + \exp[(t \times ps^{-1})^\beta]} \tag{6.3}$$

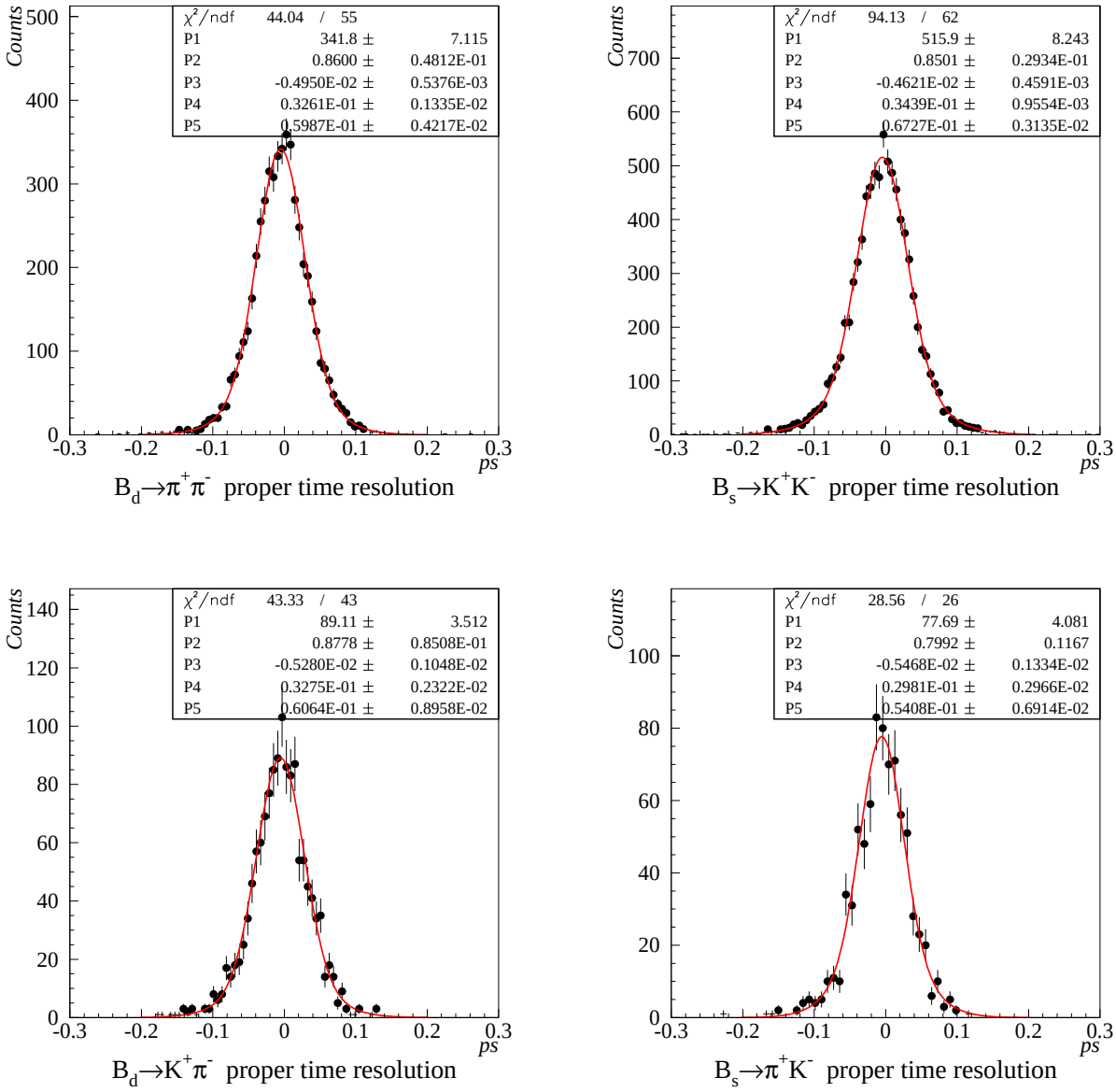

Figure 11: $\Delta t = t_{rec} - t_{true}$ distributions for triggered and offline-selected events. The solid curves are double Gaussian functions with common mean fitted to the data (see text).

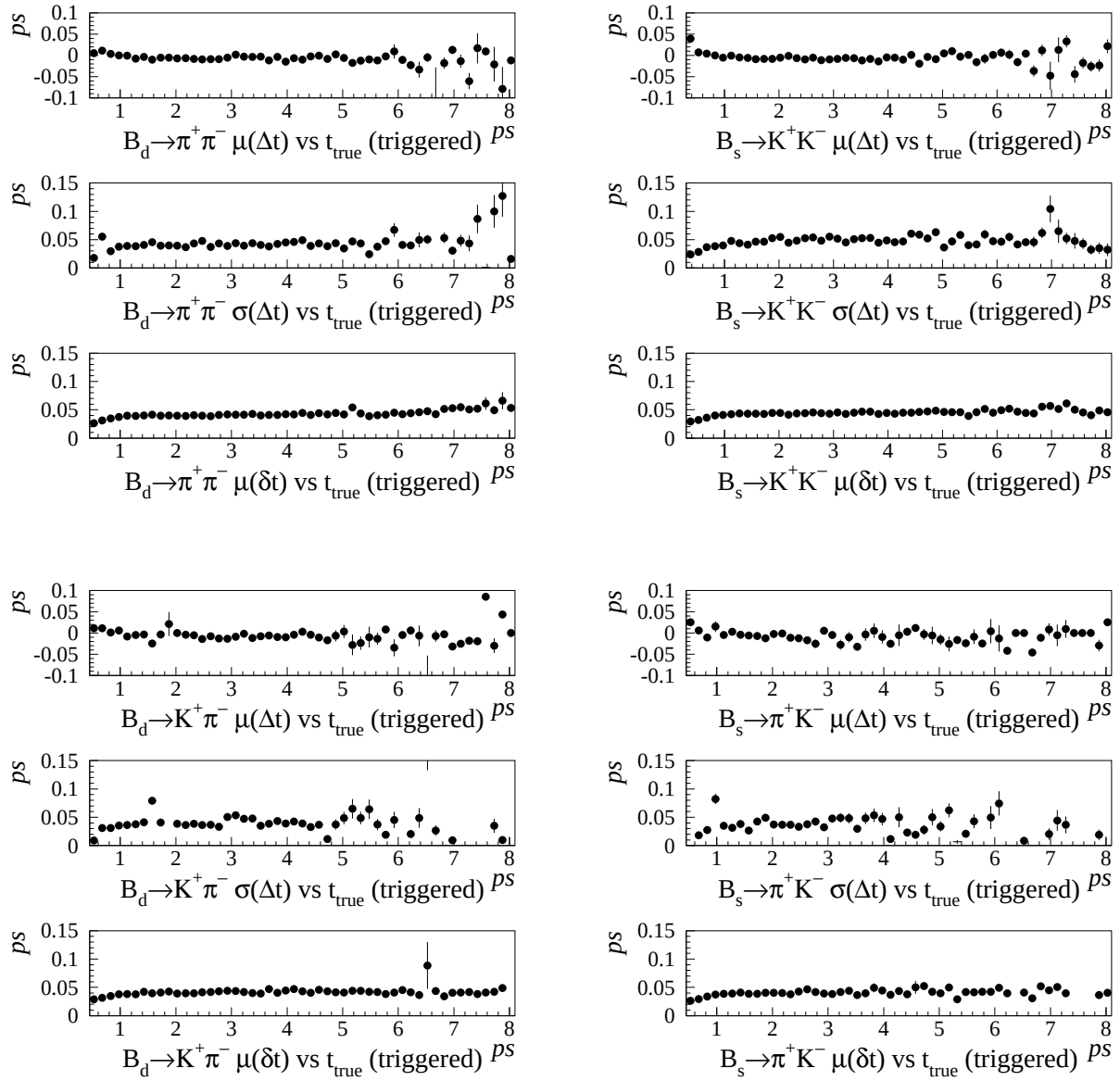

Figure 12: Average mean, average standard deviation and average estimated value of $\Delta t = t_{rec} - t_{true}$, as functions of the true proper time, for triggered and offline-selected events of each $B^0_{(s)} \to h^+h^-$ channel. Single points not following the average dependence are due to smaller statistics in the bins or to outliers not removed from the plots.

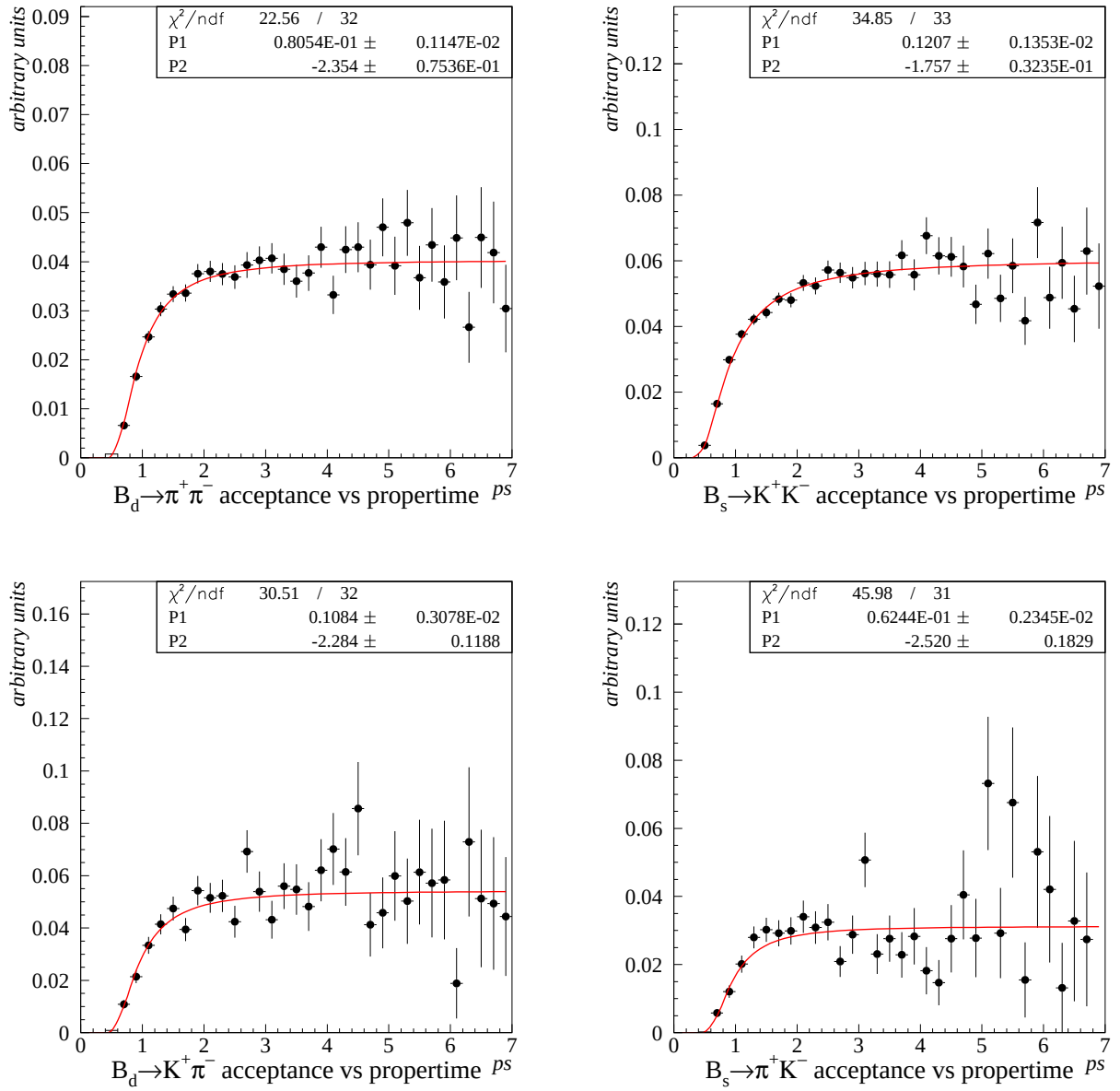

Figure 13: Signal acceptance as function of the proper time for triggered and offline-selected events of each $B^0_{(s)} \to h^+h^-$ channel (in arbitrary units). The solid curves represent a suitable parametrization fitted to the data (see text).

Channel	$\mu\,[(GeV/c^2)^{-1}]$	$\eta\,[ps^{-1}]$	δ
$B^0 \to \pi^+\pi^-$	1.71 ± 0.27	1.02 ± 0.12	-3.09 ± 0.54
$B_s^0 \to K^+K^-$	1.21 ± 0.32	0.89 ± 0.13	-1.03 ± 0.28
$B^0 \to K^+\pi^-$	0.54 ± 0.24	1.15 ± 0.15	-2.16 ± 0.35
$B_s^0 \to K^-\pi^+$	0.44 ± 0.28	0.87 ± 0.12	-2.19 ± 0.48

Table 12: Results of fit to the mass and proper time distributions of $b\bar{b}$ events, after trigger and offline-selection with slightly reduced selection cuts. The meaning of each parameter is described in the text. Shown uncertainties are statistical.

where the term (ps^{-1}) ensures the correct dimensionality. Results of the fits with this function to the proper time acceptance distributions are summarized in Table 11.

6.2 Generic $b\bar{b}$ distributions

Due to the limited $b\bar{b}$ Monte Carlo statistics available, only a handful of $b\bar{b}$ events survive the offline-selection, and in case the trigger is also applied, the number of surviving events is of the order of few units for all channels under study (see Tables 6 and 7). Then, a study on $b\bar{b}$ selected events has been performed by slightly relaxing some of the offline-selection cuts, in order to end up with a sizable statistics. We show here a study of the mass and proper time distributions of triggered and offline-selected events with reduced cuts.

The $b\bar{b}$ mass distributions are shown in Figure 14. The distributions have been fitted with an exponential function:

$$g_B(m) \propto e^{-\mu m}. \tag{6.4}$$

Figure 15 shows the proper time distributions of $b\bar{b}$ events. The functional form of the $b\bar{b}$ proper time distribution has been guessed to be the same as the untagged signal one, which can be studied in more detail due to the larger statistics of selected signal samples. The adopted function is:

$$b(t) \propto \frac{e^{-\eta t}}{1 + \exp[(t \times ps^{-1})^{\delta}]}. \tag{6.5}$$

The results of the mass and proper time fits are summarized in Table 12. The "lifetime" of $b\bar{b}$ selected events is compatible with a value of $1\,ps$, i.e. about 2/3 of the $B_{(s)}^0$ lifetime, for all the $B_{(s)}^0 \to h^+h^-$ channels.

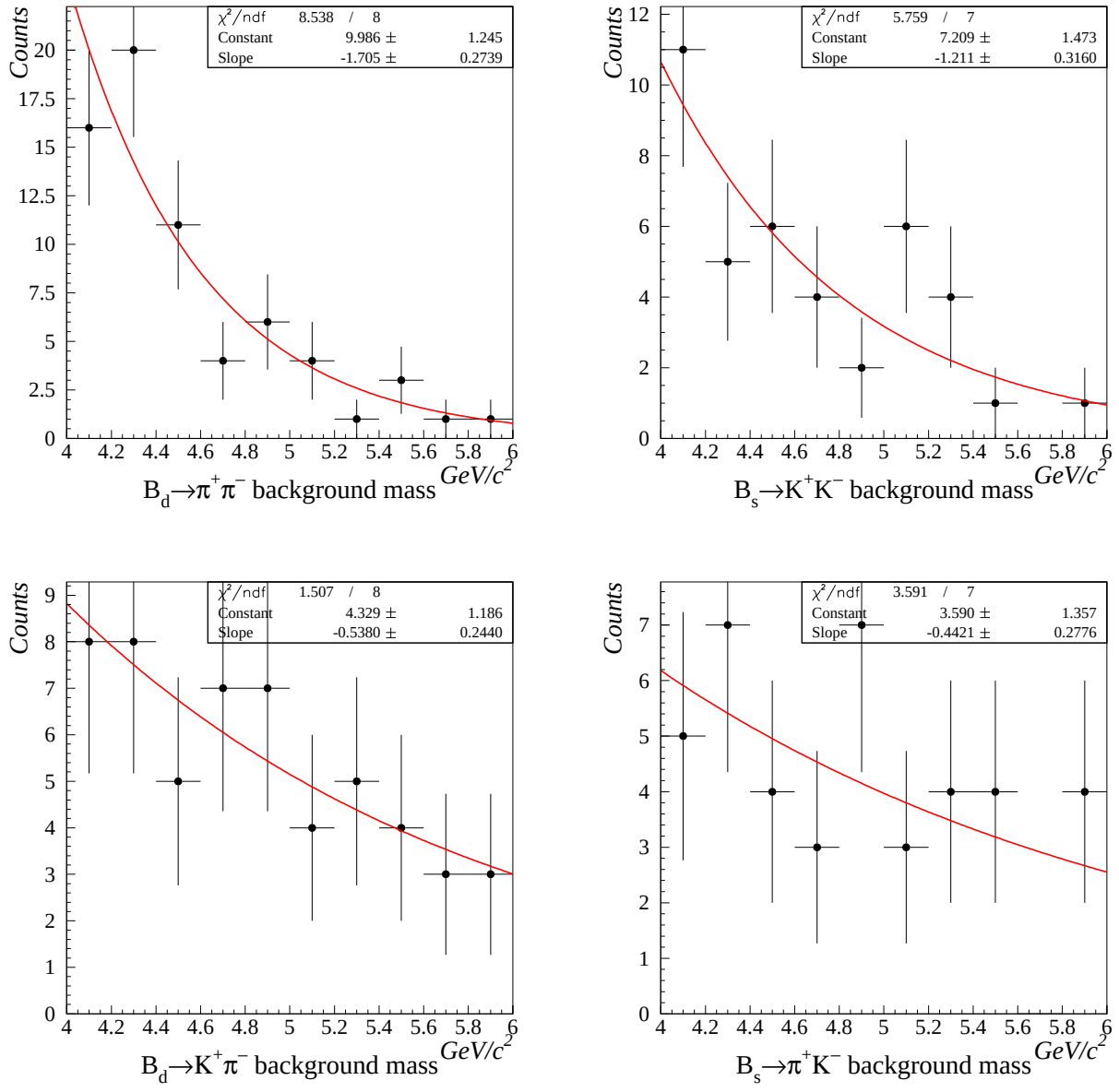

Figure 14: Mass distributions for $b\bar{b}$ events, after trigger and offline-selection with slightly reduced selection cuts. The solid curves are exponential functions fitted to the data.

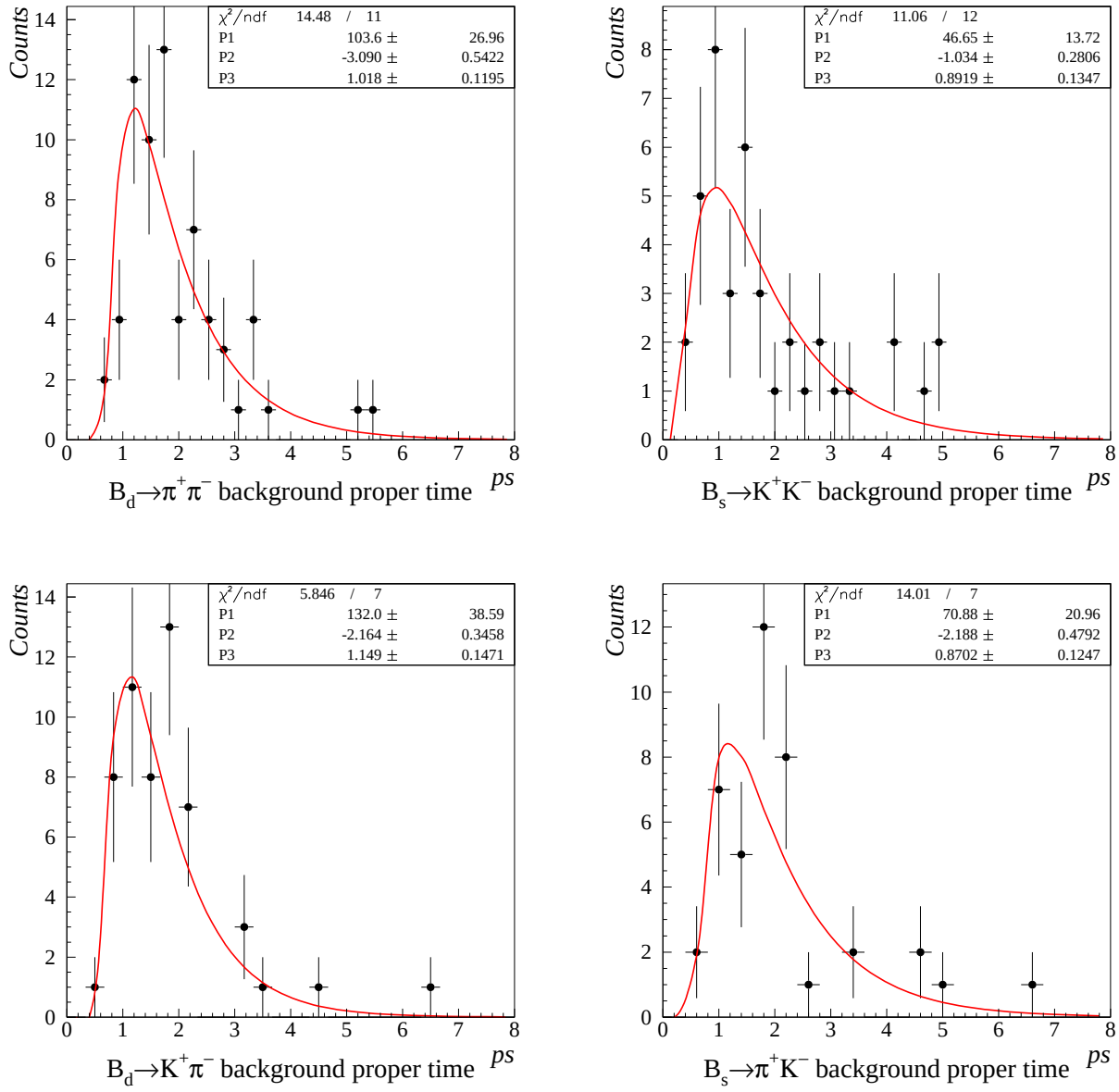

Figure 15: Proper time distributions for $b\bar{b}$ events, after trigger and offline-selection with slightly reduced selection cuts. The solid curves represent a suitable parametrization fitted to the data (see text).

7 Conclusions

The selection of $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$ decays at LHCb has been presented. Annual signal yields and background-to-signal ratios for specific and combinatorial backgrounds have been estimated. The studies described in this note indicate that large signal samples with quite a good purity can be collected by the experiment. These large statistics can be used to measure the respective CP-violating observables with outstanding precision, already after one year of data taking. Results of a study on the CP measurement for these channels, as well as on the extraction of the CKM angle γ that can be performed by combining the $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ CP-violating observables, are discussed in a separate LHCb note [10].

Acknowledgements

Special thanks go to T. Nakada, O. Schneider and G. Wilkinson, for useful discussions and suggestions, as well as for their constant attention during the development of this work. H. Dijkstra, E. Rodrigues and F. Teubert for their help in interpreting the trigger results, M. Calvi, O. Dormond, C. Matteuzzi and M. Musy for their help in interpreting the tagging results, are also kindly acknowledged. Grateful thanks are expressed to A. Golutvin for having reviewed the manuscript and for useful suggestions on improvements. The authors would like to thank the whole LHCb Computing Group for the effort undertaken in generating the huge statistics of Monte Carlo events necessary for this analysis, as well as for the other analyses described in the LHCb re-optimization TDR. Finally, many thanks go to F. Ruggieri and the INFN-CNAF staff for their technical support in keeping the LHCb-Bologna computing cluster, massively employed for event generation, reconstruction and analysis, in perfect shape and efficiency.

References

[1] LHCb Collaboration, LHCb Technical Proposal, CERN-LHCC/98-004.

[2] LHCb Collaboration, LHCb Technical Design Report, CERN-LHCC/2003-030.

[3] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89** (2002) 201802.

[4] BELLE Collaboration, K. Abe *et al.*, Phys. Rev. D **66** (2002) 071102.

[5] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89** (2002) 281802.

[6] BELLE Collaboration, K. Abe *et al.*, Phys. Rev. D **68** (2003) 012001.

[7] R. Fleischer and T. Mannel, Phys. Lett. B **397** (1997) 269.

[8] R. Fleischer, Phys. Lett. B **459** (1999) 306.

[9] R. Fleischer and J. Matias, Phys. Rev. D **66** (2002) 054009.

[10] V. Vagnoni *et al.*, CP sensitivity with $B^0_{(s)} \to h^+h^-$ decays at LHCb, CERN-LHCb/2003-124.

[11] T. Sjonstrand *et al.*, Comput. Phys. Commun. **135** (2001) 238.

[12] QQ program, `http://www.lns.cornell.edu/public/CLEO/soft/QQ`.

[13] LHCb Collaboration, SICB - The LHCb GEANT3 based Simulation Program, `http://lhcb-comp.web.cern.ch/lhcb-comp/SICB/`.

[14] R. Brun *et al.*, GEANT3, Internal Report, CERN DD/EE/84-1 (1987).

[15] LHCb Collaboration, BRUNEL - The LHCb Reconstruction Program, `http://lhcb-comp.web.cern.ch/lhcb-comp/Reconstruction/`.

[16] LHCb Collaboration, GAUDI Project, `http://lhcb-comp.web.cern.ch/lhcb-comp/Frameworks/Gaudi/`.

[17] LHCb Collaboration, DIRAC - The LHCb Production Tool, `http://lhcb-comp.web.cern.ch/lhcb-comp/Production/`.

[18] G. Avoni *et al.*, A Beowulf-class computing cluster for the Monte Carlo production of the LHCb experiment, Comput. Phys. Commun. **150** (2003) 129.

[19] LHCb Collaboration, DAVINCI - The LHCb Analysys Program, `http://lhcb-comp.web.cern.ch/lhcb-comp/Analysis/`.

[20] Particle Data Group, K. Hagiwara *et al.*, 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/`.

[21] LHCb Collaboration, RICH Technical Design Report, CERN-LHCC/2000-37.

[22] LHCb Collaboration, Trigger Technical Design Report, CERN-LHCC/2003-31.

[23] M. Calvi, O. Dormond and M. Musy, LHCb flavor tagging performance, CERN-LHCb/2003-115.