Chapter 3

Selection of $B_s^0 \to \mu^+\mu^-$ events for the effective lifetime Measurement

The analysis described in Chapter X for the $B_s^0 \to \mu^+\mu^-$ effective lifetime requires $B_s^0 \to \mu^+\mu^-$ and $B_{(s)}^0 \to h^+h^-$ decays to be identified in the data sets recorded by the LHCb experiment. Although $B_s^0 \to \mu^+\mu^-$ decays leave a clear 2 muon signature in the detector, the selection of these decays is challenging because it is a very rare process and there are many other processes that can mimic a decay in the detector. The background processes are described in Section 3.1. To understand different aspects of the selection and analysis of decays, particle decays with a similar topology to are used. $B_{(s)}^0 \to h^+h^-$ decays, where $h = K, \pi$, are used because they have large branching fractions and are well understood from pervious LHCb analyses as well as a similar topology to $B_s^0 \to \mu^+\mu^-$ decays. The measurement of the $B_{(s)}^0 \to \mu^+\mu^-$ Branching Fractions, described in Chapter X, requires the use of $B^+ \to J/\psi K^+$ decays, as well as $B_{(s)}^0 \to \mu^+\mu^-$ and $B_{(s)}^0 \to h^+h^-$ decays, to be used as a noramlisation channel. This Chapter describes the selection of $B_{(s)}^0 \to \mu^+\mu^-$, $B_{(s)}^0 \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays for the effective lifetime and Branching fraction analyses. The analyses share many of the same selection requirements. The selection occurs in several stages, and the development of the selection relies on simulated events which are detailed in Section 3.2. The first step to select decays is choosing what requirements to place on the trigger which is followed by a set of loose selection requirements to remove obvious background events. These two steps are described in Sections 3.2.1 and 3.2.2. A tighter selection is applied to the output of the stripping as described in Section 3.2.4 and particle identification requirements are used in Section 3.2.5 to further reduced background events. Finally a multivariate classifier is described in Section **??** used as the final step in the selection to reduced the backgrounds to a level suitable for the analysis in

Chapter X to be preformed. Throughout this Chapter $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are selected in the same way.

The LHCb collaboration has published a number of papers studying the $B_s^0 \to \mu^+\mu^-$ decay, the selection described in this Chapter has been built up over a number of years by a range of different collaboration members. The studies detailed in sections X, X and Y was done for this thesis.

3.1 Backgrounds

A B_s^0 decaying into two muons leaves information in the LHCb detector with certain identifying characteristics. The two muons form a good vertex that is displaced from the primary vertex of the event because the Bs has a long lifetime and the combined momentum of the muons can be extrapolated backwards to the primary vertex because the muons are the only decay products of the B_s^0. There are other processes that occur in proton-proton decays that can leave information in the detector in a similar pattern to $B_s^0 \to \mu^+\mu^-$ decays. The reconstruction, described in Section X, produces many $B_s^0 \to \mu^+\mu^-$ candidates, the aim of the selection is to separate the real $B_s^0 \to \mu^+\mu^-$ decays from the background in the reconstructed candidates.

The main sources of background processes for $B_s^0 \to \mu^+\mu^-$ decays are;

- Elastic collisions of protons, $pp \to p\mu^+\mu^-p$, can produce a pair of muons whilst the protons travel down the beam pipe. The muons produced have low transverse momentum.

- Inelastic proton collisions can create two muons at the primary vertex. These muons can be combined to for a B_s^0 that decays instantaneously. This type of background is prompt combinatorial background.

- The $B_S^0 \to \mu^+\mu^-\gamma$ decays can mimic $B_s^0 \to \mu^+\mu^-$ when the photon is not reconstructed. The presence of the photon in the decay means that $B_S^0 \to \mu^+\mu^-\gamma$ is not helicity surpassed and could therefore be a sizeable background, however the photon gains a large transverse momentum therefore when reconstructed as $B_s^0 \to \mu^+\mu^-$ the B_s^0 mass is much lower than expected.

- $B_s^0 \to \mu^+\mu^-$ candidates can be formed when muons produced in separated semi-leptonic decays are combined. These are known as long lived combinatorial background because the reconstructed B_s^0 will not decay instantaneously.

- Semi-leptonic decays when on to the daughters is mis-identified as a muon and/or is not detected can mimic $B_s^0 \to \mu^+\mu^-$ decays. The resulting reconstructed mass of the B_s^0 is lower than expected due to the missing particle information. The semi-leptonic decays that contribute to $B_s^0 \to \mu^+\mu^-$ backgrounds in this way are $B^0 \to \pi^-\mu{+}\nu_\mu$, $B_s^0 \to K^-\mu{+}\nu_\mu$, $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$, $B^0 \to \pi^0\mu^+\mu^-$ and $B_c^+ \to J/\psi\mu^+\nu_\mu$ where $J/\psi \to \mu^+\mu^-$.

- $B_{(s)}^0 \to h^+h^-$ decays, where $h = K, \pi$, can form background when both hadrons are mis-identified as muons. This usually occurs when the hadrons decay in flight. The mis-identification of the hadrons leads to the reconstructed B_s^0 mass being lower than expected.

Separating the backgrounds from the $B_s^0 \to \mu^+\mu^-$ decays can be done relatively straightly forwardly for many of the background processes by taking advantage of the obvious differences between the background and $B_s^0 \to \mu^+\mu^-$ decays. However, distinguishing $B_s^0 \to \mu^+\mu^-$ decays from long lived combinatorial backgrounds, and mis-idetificed $B_{(s)}^0 \to h^+h^-$ and semi-leptonic decays is more challenging and $B_s^0 \to \mu^+\mu^-$ decays must be sacrificed in order to remove a sufficient about of the background processes for the analysis to be performed. For the effective lifetime analysis the $B^0 \to \mu^+\mu^-$ decay is not relevant and is therefore a background, however since the decays are extremely similar the $B_{(s)}^0$ masses are the only way to seperate the decays.

3.2 Simulated Particle Decays

Simulated events are needed for the development of the selection for $B_s^0 \to \mu^+\mu^-$ and $B_{(s)}^0 \to h^+h^-$ decays and to study aspects of the analysis strategy. Monte Carlo simulated events and their passage through the LHCb detector, as described in Section X, are used. The advantage of simulated events is that a clean sample a particular type of particle decay can be produced in much greater numbers compared to what is seen in data. A large range of different simulated decays types have been used over time for the development of the selection, the simulated decays used directly for studies documented in this thesis are listed in Table 3.1.

Simulated $B_s^0 \to \mu^+\mu^-$, $B^0 \to \mu^+\mu^-$, $B^0 \to K^+\pi^-$ and $B^+ \to J/\psi K^+$ decays for 2012 data taking condition are used for studying the stripping selection in Section X.

The training and testing of multivarite classifiers in Section X uses simulated $B_s^0 \to \mu^+\mu^-$ and $b\bar{b} \to \mu^+\mu^- X$ decays for for 2012 data taking conditions.

Simulated events for $B_s^0 \to \mu^+\mu^-$, $B_s^0 \to K^+K^-$ and $B^0 \to K^+\pi^-$ for data taken in 2011, 2012, 2015 and 2016 are used for developing the analysis method in Chapter X.

The production of simulated events is constantly being developed as understanding of the detector increases and to include changes made for each data is recorded at theLHC. Therefore there exists a number of different simulation versions that can be used to simulate events.

Each year data is collected at LHCb the conditions the experiment operates at and the proton collisions delivered by the LHC change. These changes include differences in the the selection used in the trigger for each year and increases in the centre of mass energy of proton collisions.

Therefore to understand data collected in different years simulated events from each year of data taking is needed, it is important to use similar simulation versions for each year so that the difference in the data taking conditions are not masked by differences in simulation versions. Simiarly for training multivarite classifiers consistent simulation versions are needed for the signal and background samples so that difference between signal and background distributions are not masked by differences in simulation versions.

In general the stripping selections are applied to simulated events, however events that do not pass the stripping selection are still saved and can be used after reprocessing the simulated events. However simulation conditions can be set up so that events that do not pass the stripping selection are discarded and can never be used and also cuts can be applied on particles when they are generated before the detector response is simulated and the events are reconstructed. This is used when a very large same of simulated events needs to be generated in order to have a suitably large same of events reconstructed and is the case for the samples of $b\bar{b} \to \mu^+\mu^- X$ simulated events.

Although in general simulated events accurately model what occurs in data there are several areas where this is not the case. The distributions of particle identification variables and properties of the underlying proton-proton event, such as the number of tracks in an event, are not well modelled in simulation. The mis-modelling of particle identification variables can be corrected for using the PIDCalib package but this is not used directly for this thesis. Comparisons of the number of tracks in an event between simulated decays and data can be used to re-weight simulated decays to that the under lying event is accurately modelled. This is described in Sections X.

3.2.1 Trigger

The trigger is the first step in the selection process and the structure of the trigger is described in Section X. Since $B_s^0 \to \mu^+\mu^-$ decays are very rare a broad set of trigger requirements is used in order to keep a high proportion of $B_s^0 \to \mu^+\mu^-$ decay at this step of the selection. Specific trigger lines are not used in the selection but rather the combined results of a large selection of trigger lines at each level of the trigger. The combinations of trigger lines used are the L0Global, Hlt1Phys and Hlt2Phys triggers. The L0Global tigger combines all trigger lines present in the L0 trigger, it selects an event provided at least one L0 selects it and rejects an event if no L0 trigger selects it. The Hlt1Phys and Hlt2Phys triggers are very similar to the L0Global trigger except that decisions are based only trigger lines related to physics processes and HLT trigger lines used for calibration are excluded.

Different trigger decisions on these lines are used to select decays for the Branching Fration and effective lifetime analyses. The Branching fraction selection imposed the loosest trigger requirements by requiring a event to pass the 'Dec' decision at each trigger level as illustrated in set 'A' of Table X. Trigger decisions are defined in Section X. The effective lifetime analysis has slightly more constrained trigger requirement, requiring an event passes either the 'TIS' or 'TOS' decision at each level of the trigger as illustrated in set 'B' of Table X. The trigger choice for the effective lifetime is motivated by the determination of the acceptance fuction in Section X.

Events are required to be either TIS, triggered independent of signal or TOS, triggered on signal, on the trigger lines used at each level of the trigger. Slightly different trigger requirements are used to select $B^0_{(s)} \to h^+h^-$ decays used to develop and validate the effective lifetime analysis, the same broad trigger lines are used but the requirement on the output varies depending on the use of the $B^0_{(s)} \to h^+h^-$ events. The are two sets of trigger requirments, set 'A' and 'C', in Table 3.2 are used to select $B^0_{(s)} \to h^+h^-$ decays, it will be made clear in later sections where $B^0_{(s)} \to h^+h^-$ decays are used which trigger requirements are imposed.

Decay	Year	Simulation Version	Generated Events
Stripping selection studies selection			
$B_s^0 \to \mu^+\mu^-$	2012	sim06b	2M
$B^0 \to \mu^+\mu^-$	2012	sim06b	
$B^0 \to K^+\pi^-$	2012	sim06b	
$B^+ \to J/\psi K^+$	2012	sim06b	
Analysis method development			
$B_s^0 \to \mu^+\mu^-$	2011	sim08	0.5M
	2012	sim08	
	2015	sim09	2M
	2016	sim09	
$B^0 \to K^+\pi^-$	2011	sim08	0.8M
	2012	sim08	8.5M
	2015	sim09	4M
	2016	sim09	
$B_s^0 \to K^+K^-$	2012	sim08	7M
	2015	sim09	
Multivarite classifer training			
$b\bar{b} \to \mu^+\mu^- X$, p_T	2012	sim06	
$b\bar{b} \to \mu^+\mu^- X$, p_T	2012	sim06	
$B_s^0 \to \mu^+\mu^-$	2012	sim06b	2M

Table 3.1 Simulated events used for developing the selection and the analysis method listed according to the studies the simulated events are used for. Requirements imposed on generated $b\bar{b} \to \mu^+\mu^- X$ decays are included alongside the decay type. The number of generated events includes even numbers for each magnet polarity.

Trigger Line	Trigger decision
set A	
L0Global	Dec
Hlt1Phys	Dec
Hlt2Phys	Dec
set B	
L0Global	TIS or TOS
Hlt1Phys	TIS or TOS
Hlt2Phys	TIS or TOS
set C	
L0Global	TIS
Hlt1Phys	TIS
Hlt2Phys	TIS

Table 3.2 Trigger lines used to select $B_s^0 \to \mu^+\mu^-$ and $B_{(s)}^0 \to h^+h^-$ decays. Set 'A' is used to select decays for the Branching Fraction analysis. Set 'B' is used to select $B_s^0 \to \mu^+\mu^-$ decays for the effective lifetime anaylsis. Sets 'A' and 'C' are used to select $B_{(s)}^0 \to h^+h^-$ decays used to develop the $B_s^0 \to \mu^+\mu^-$ effective lifetime analysis.

Trigger decision				
Trigger Line	Branching Fraction analysis	$B_s^0 \to \mu^+\mu^-$ effective lifetime	$B_{(s)}^0 \to h^+h^-$ effective lifetime	
			set A	set B
L0Global	Dec	TIS or TOS	TIS	Dec
Hlt1Phys	Dec	TIS or TOS	TIS	Dec
Hlt2Phys	Dec	TIS or TOS	TIS	Dec

Table 3.3 Trigger lines used to select $B_s^0 \to \mu^+\mu^-$ and $B_{(s)}^0 \to h^+h^-$ decays. Set 'A' is used to select decays for the Branching Fraction analysis. Set 'B' is used to select $B_s^0 \to \mu^+\mu^-$ decays for the effective lifetime anaylsis. Sets 'A' and 'C' are used to select $B_{(s)}^0 \to h^+h^-$ decays used to develop the $B_s^0 \to \mu^+\mu^-$ effective lifetime analysis.

There was a problem with the implementation of the Hlt2Phys Dec decision in 2016 simulated events.This only affect the selection of $B^0_{(s)} \to h^+h^-$ decays. In order to emulate this trigger a combination of Hlt2 lines that select $B^0_{(s)} \to h^+h^-$ events, listed in Table 3.4, is used instead of HLT2Phys when the Dec decision is required.

$B^0_{(s)} \to h^+h^-$ trigger lines
Hlt2Topo2BodyDecision Dec
Hlt2B2HH Lb2PPiDecision Dec
Hlt2B2HH Lb2PKDecision Dec
Hlt2B2HH B2PiPiDecision Dec
Hlt2B2HH B2PiKDecision Dec
Hlt2B2HH B2KKDecision Dec
Hlt2B2HH B2HHDecision Dec

Table 3.4 Trigger lines used to emulate the Hlt2Phys$_{D}ecdecisionfor\mathrm{B}^0_{(s)} \to h^+h^-$ data and simulated events.

3.2.2 Loose Selection

The events that pass the selection requirements in the trigger are reconstructed and the number of events is reduced by a set of loose selection cuts. These selection cuts are aimed at removing obvious background decays by exploiting the differences between real $B^0_{(s)} \to \mu^+\mu^-$ decays and backgrounds that mimic them. The loose selection is composed to two parts; the stripping selection, outlined in Section 1.2.4, and a set pre-selection cuts.

The Branching Fraction and effective lifetime analyses share the same loose selection, afterwards tighter cuts are applied to remove specific backgrounds many of which are tuned for each analysis, hence the loose selection cuts are called the pre-selection cuts.

The stripping lines used in the last published Branching Fraction analysis were designed at the start of Run 1 by studying the efficiencies of different selection cuts from simulated events []. However since then improvements have been made to the simulation of particle decays at LHCb, therefore it is prudent to check the accuracy of the selection efficiencies with updated simulated events and investegate where improvements can be made to the efficiency of the stripping selection used to select $B^0_s \to \mu^+\mu^-$ events. Furthermore, before the start of Run 2 during the long shut down

the stripping selection was re-run on the data collect in Run 1 allowing changes to be made to existing stripping lines and new lines to be added. The stripping selection and efficiency studies are described in Sections 3.2.2.1 and 3.2.2.2. In general the stripping selection imposes only loose requirements so that as much information as possible is still avaliable to develop the analysis and understand background events after the stripping selection. Therefore after the stripping selection further loose selection requirements are applied to remove different background decays before the tighter selection requirements detailed in Section 3.2.3.

3.2.2.1 Run 1 Stripping Selection

The stripping selection cuts and cuts applied during the reconstruction of particle decays for the Run 1 $B^0_{(s)} \to \mu^+\mu^-$ Branching Fraction analysis [] to select $B^0_{(s)} \to \mu^+\mu^-$, $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ are shown in Table X. The selection of $B^+ \to J/\psi K^+$ and $B^0_{(s)} \to h^+h^-$ decays is kept as similar as possible to the selection of $B^0_s \to \mu^+\mu^-$ decays to avoid introducing systematic errors when $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays are used in the normalisation for the Branching Fraction measurement. The selection of $B^+ \to J/\psi K^+$ event must diverge from the $B^0_s \to \mu^+\mu^-$ selection due to additoinal particles in the final state of the decay. The stripping selection imposes more cuts to select $B^0_{(s)} \to h^+h^-$ decays compared to $B^0_s \to \mu^+\mu^-$ because $B^0_{(s)} \to h^+h^-$ decays are much more abundant therefore extra cuts are needed to reduce the number of events passing the stripping to an acceptable level. The cuts applied to $B^0_{(s)} \to h^+h^-$ in the stripping are the later applied to $B^0_s \to \mu^+\mu^-$ events after the stripping selection.

Particle	$B_s^0 \to \mu^+\mu^-$	$B_{(s)}^0 \to h^+h^-$	$B^+ \to J/\psi K^+$
B_s^0 or B^+	|M - M_{PDG}| < 1200 MeV/c^2	|M - M_{PDG}| < 500 MeV/c^2	|M - M_{PDG}| < 500 MeV/c^2
	DIRA > 0	DIRA > 0	Vertex χ^2/ndof < 45
	FD χ^2 > 121	FD χ^2 > 121	IP χ^2 < 25
	IP χ^2 < 25	IP χ^2 < 25	
	Vertex χ^2/ndof < 9	Vertex χ^2/ndof < 9	
	DOCA < 0.3 mm	DOCA < 0.3 mm	
		τ < 13.248 ps	
		p_T > 500 MeV/c	
J/ψ			|M - M_{PDG}| < 60 MeV/c^2
			DIRA > 0
			FD χ^2 > 169 (225)
			Vertex χ^2/ndof < 9
			DOCA < 0.3 mm
Daugther μ or h	Track χ^2/ndof < 3	Track χ^2/ndof < 3	Track χ^2/ndof < 3
	isMuon = True		isMuon = True
	Minimum IP χ^2 > 9	Minimum IP χ^2 > 9	Minimum IP χ^2 > 25
		0.25 GeV/c < p_T < 40 GeV/c	
		p < 350 GeV/c	
		ghost probability < 0.3 (0.4)	
K^+			Track χ^2/ndof < 3
			p_T > 0.25 GeV/c
			Minimum IP χ^2 > 25

Table 3.5 Selection requirements applied during the stripping selection for Run 1 data used in the $B_{(s)}^0 \to \mu^+\mu^-$ Branching Fraciton analysis [] to select $B_{(s)}^0 \to \mu^+\mu^-$, $B_{(s)}^0 \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays. The track χ^2/ndof and isMuon cut are applied during the reconstuction.

The variables used in the stripping selection are:

- the reconstructed mass, M, - the mass and momenta of the decay products of the B meson (or J/ψ) are combined to provide its reconstructed mass. Cuts on the mass remove events with a reconstructed mass far from the expected particle mass and are therefore clearly to be background events. Loose mass requirements are made on for the $B_s^0 \to \mu^+\mu^-$ selection to allow for the study of semi-leptonic backgrounds that have a mass less than the B_s^0 mass when mis-identified as $B_s^0 \to \mu^+\mu^-$ decays;

- the "direction cosine", DIRA, - this is the cosine of the angle between the mometum vector of the B meson and the vector connecting the primary vertex to the secondary vertex where the B meson decays. For correctly reconstructed events the direction cosine should be very close to one, requiring events to have positive value ensuring events are travelling in the incorrect direction form the SV to the PV are removed;

- the flight distance (FD) χ^2- this is computed by performing the fit for the production vertex of a particle but including the tracks from its decay products that originate from the decay vertex in the fit as well. For a B meson the FD χ^2 is likely to be large because B mesons have long lifetimes therefore the tracks of its decays products will not point back to the production vertex. Alternatively a J/ψ will have a small χ^2 because it decays instantaneously;

- track fit $\chi^2/ndof$ - provides a measure of the quality of a fitted track, placing an upper limit on this parameter removes poor quality tracks and therefore backgrounds composed of poorly reconstructed decays;

- vertex fit $\chi^2/ndof$ - provides a measure of how well tracks can be combined to form a vertex, placing an upper limit on this parameter removes poorly constrained verticies and therefore backgrounds composed of poorly reconstructed decays;

- "distance of closest approach" (DOCA) - this is the distance of closest appoach of two partilcles computed from the straight tracks in the VELO. For the decay products of a particle, for example the muons from $B_s^0 \to \mu^+\mu^-$, this distance would ideally be zero because they originate from the same vertex;

- decay time τ- this is the length of time a particle lives for as it travels from its production vertex to its decay vertex. Applying an upper decay time cut removes unphysical background decays;

- isMuon - particle identification variable defined in Section **??** that returns True for muons and False for other particles;
- transverse mometum, p_T- the component of a particle's momentum perpendicular to the beam axis. Decay products of B mesons are expected to have relatively high p_T values due to the heavy B meson masses however an upper limit removes unphysical backgrounds
- momentum, p- an upper limit on the mometum of a particle removes unphysical backgrounds
- ghost probability - defined in Section **??** provides the probability of a tracking being composed on random hits in the detector.
- impact parameter (IP) χ^2- this is the change in the fit for a primary vertex caused by removing one track in the fit. In a $B_s^0 \to \mu^+\mu^-$ decay, the B_s^0 is produced at the PV therefore it should have a small IP χ^2 value whereas the muons will be displace from the PV because of the realtively long lifetime of the B_s^0 and therefore will have a large IP χ^2;
- Minimum muon impact parameter (IP) χ^2- this is the IP χ^2 of the muons with respect to all PVs in the event, this is to remove prompt muons created at any PV in the event and therefore reduce the prompt combinatorial background.

3.2.2.2 Improving Stripping Selection Efficiency for $B_{(s)}^0 \to \mu^+\mu^-$ decays

	Efficiency		
Requirement	$B_s^0 \to \mu^+\mu^-$	$B_{(s)}^0 \to h^+h^-$	$B^+ \to J/\psi K^+$
B |M - M_{PDG}|	(100.00 ± 0.00)%	(±)%	(±)%
$B_{(s)}^0$ or J/ψ DIRA	(99.43 ± 0.01) %	(±)%	(±)%
$B_{(s)}^0$ or J/ψ FD χ^2	(83.89 ± 0.06) %	(±)%	(±)%
$B_{(s)}^0$ or J/ψ IP χ^2	(96.88 ± 0.03) %	(±)%	(±)%
$B_{(s)}^0$ or J/ψ vertex χ^2/ndof	(97.36 ± 0.03) %	(±)%	(±)%
$B_{(s)}^0$ or J/ψ DOCA	(99.86 /pm 0.01) %	(±)%	(±)%
μ or h minimum IP χ^2	(80.47 ± 0.06) %	(±)%	(±)%
Efficiency after above cuts	(73.75 ± 0.07) %	(±)%	(±)%
Efficiency after all striping line cuts	(73.75 ± 0.07) %	(±)%	(±)%

Table 3.6 Stripping line efficiencies for $B_s^0 \to \mu^+\mu^-$, $B_{(s)}^0 \to h^+h^-$ and $B^+ \to J/\psi K^+$ 2012 simulated decays after broad trigger requirements. Selection cuts applied are listed in Table **??**.

The efficiencies of the stripping lines for selecting $B^0_{(s)} \to \mu^+\mu^-$, $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays are shown in Table 3.6. The efficiencies are evalutated using 2012 sim06 simulated events and a loose set of trigger requirements (set A from Table 3.2) have been imposed. Only cuts that are applied to select $B^0_{(s)} \to \mu^+\mu^-$ decays are evaluated, the efficiencies for the isMuon and track $\chi^2/ndof$ are not included because decays that do not pass these requirements are not included in the samples of simulated events.

The selection efficiencies are similar accross the different decays for the selection cuts that are shared with the $B^0_{(s)} \to \mu^+\mu^-$ selection, and the total efficicency is around 70 % for all decays.

The similarity of selection efficiencies across the different decays is futher illustrated in Figure 3.1 which shows the ratio of the efficiencies $B^0_{(s)} \to \mu^+\mu^-$ and $B^+ \to J/\psi K^+$ where each cut has been applied independantly. With the exception of the IP χ^2 cuts on the daughter particles, the ratio of efficiencies is well within 2% of 1 for the range of cuts values shown. The ratio of the $B_s^0 \to \mu^+\mu^-$ and $B^+ \to J/\psi K^+$ efficiencies for the daughter particle IP χ^2 markedly deviates from unity, showing that the IP χ^2 distribution of the muons and kaon are very different as seen previous in [**?**]. If the other selection cuts are applied to the simularted events before the daughter IP χ^2 requirement the ratio of $B^0_{(s)} \to \mu^+\mu^-$ and $B^+ \to J/\psi K^+$ efficiencies is much closer to 1.

The efficiencies for most of the stripping cuts is ~ 97 %, $however, the efficienices of the cuts on the FD\chi$ of the $B^0_{(s)}$ or J/ψ and the daugher IP χ^2 of the muon or hadron pair are lower at 83% and 80%, respectively. Therefore improvements to the stripping selection efficiencies could be acheived by altering these two selection requirements.

The set of events removed by each cut in a stripping selection is not independant. Therefore the effect of changing on cut on the total efficiencyu of a stripping selection must be considered. Figure 3.2 shows the total efficiency of the $B_s^0 \to \mu^+\mu^-$ stripping line on simulated $B_s^0 \to \mu^+\mu^-$ events that have passes the trigger requirement for a range of cut values for the FD χ^2 and daugher IP χ^2 requirements. As expected the lower the cut values are the more efficient the stripping line becomes. However one of the main purposes of the stripping selection is to reduce the size of the data set by removing obvious background events, therefore the cuts cannot be set as loose as possible. The curve on Figure 3.2 is used as a guide to study a set of FD χ^2 and daughter IP χ^2 cut values in more depth, taking in to account both the signal efficiency and the amount of data retained by the selection, the set of chosen cuts aims to keep both cuts as tight as possible for a certain efficiency. Any changed applied to the

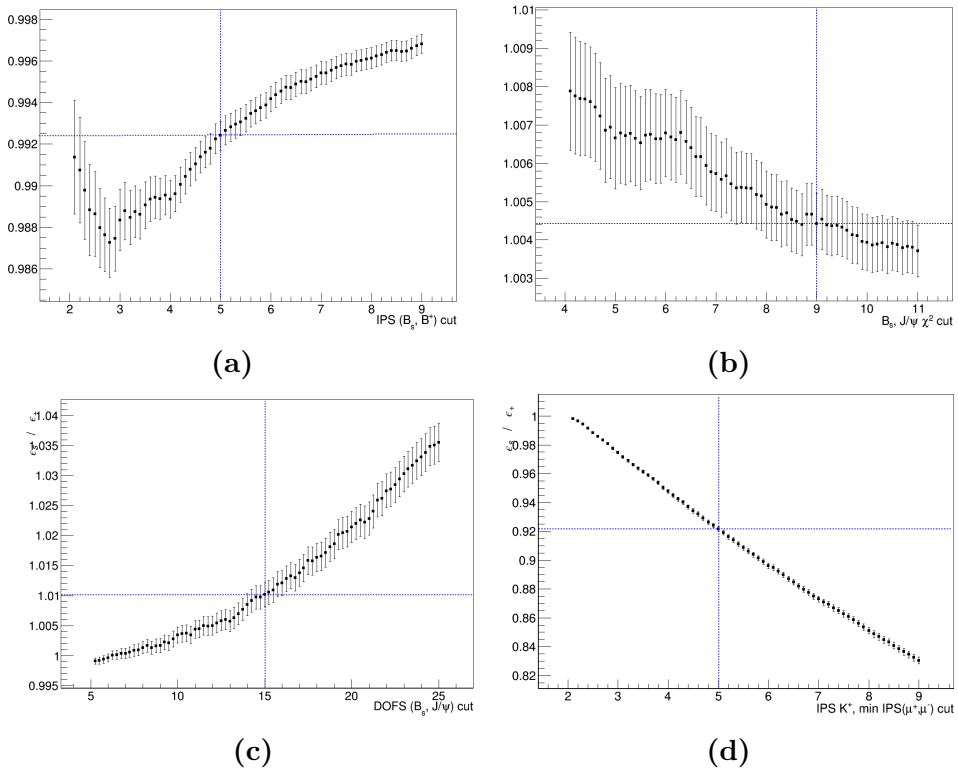
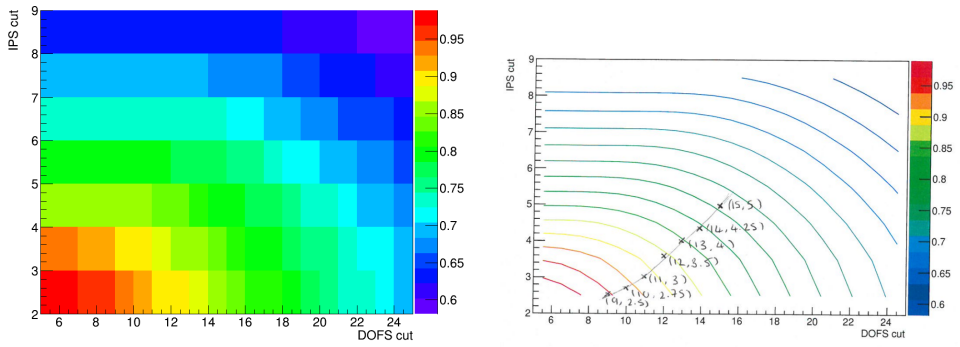


Fig. 3.1 The ratio of $B^0_{(s)} \to \mu^+\mu^-$ to $B^+ \to J/\psi K^+$ stripping efficiencies on MC events when each cut has been applied independently of all other cuts. The current cut values are marked by the blue lines.



(a) **(b)**

Fig. 3.2 Efficiency figures.

$B^0_{(s)} \to \mu^+\mu^-$ stripping line must be propagated through into the stripping lines for $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays in order to keep the selection as similar as possible for across all the decays. Changes to the $B^0_{(s)}$ FD χ^2 in $B^0_s \to \mu^+\mu^-$ correspond to a change in the $B^0_{(s)}$ FD χ^2 in $B^0_{(s)} \to h^+h^-$ and the J/ψ FD χ^2 for $B^+ \to J/\psi K^+$, similarly changes to the muon IP χ^2 in $B^0_s \to \mu^+\mu^-$ correspond to changes in the hadron IP χ^2 in $B^0_{(s)} \to h^+h^-$ and the muon IP χ^2 in $B^+ \to J/\psi K^+$. Table **??** showns the total efficiencies of the $B^0_{(s)} \to \mu^+\mu^-$, $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ stripping lines along side the amount of data retained for the set of cuts illustrated in Figure 3.2b. The data retention is computed by applying the stripping selection to a sub-set of 2012 data, with the trigger requirements applied, to find the number of events that pass the stripping lines for each pair of FD χ^2 and daughter IP χ^2 cuts. The number of events for each set of cuts is normalised to the number of events passing the original Run 1 stripping line requirements in order to show the fractional increase caused by loosening the cut values.

(Not explained that some cuts are in the reconstruction very well, so are applied before the stripping and I cannot change them and I've not made the trigger requirements clear either or probably exactly what cuts are applied to get the efficiencies, I've not included the B2JpsiK plots - do I want to? I don't think they add much to the dicussion, they show a comparison of efficiencies and that it's flat - I can ask Harry if it would add much to the discussion).

$B^0_{(s)}$, J/ψ FD χ^2	μ,h $IP\chi^2$	Stripping line efficiency			Stipping line retention		
		$B^0_s \to \mu^+\mu^-$	$B^0 \to K^+\pi^-$	$B^+ \to J/\psi K^+$	$B^0_s \to \mu^+\mu^-$	$B^0 \to K^+\pi^-$	$B^+ \to J/\psi K^+$
15	5.00	XX %	XX %	XX %	XX	XX	XX
14	4.25	XX %	XX %	XX %	XX	XX	XX
13	4.00	XX %	XX %	XX %	XX	XX	XX
12	3.50	XX %	XX %	XX %	XX	XX	XX
11	3.00	XX %	XX %	XX %	XX	XX	XX
10	2.75	XX %	XX %	XX %	XX	XX	XX
9	2.50	XX %	XX %	XX %	XX	XX	XX

Table 3.7 Retention of data and stripping line efficiencies. Efficiencies are $B^0 \to K^+\pi^-$ but retention is $B^0_{(s)} \to h^+h^-$ because the stripping line selects all $B^0_{(s)} \to h^+h^-$ decays, h is K or π.

An increase of 16% can be gained in the stripping selection efficiencies by using the loosest cuts in Table **??** however the loosest cuts increase the amount of data passing the $B^0_{(s)} \to \mu^+\mu^-$ stripping selection by a factor of 7. The final set of cuts used in the stripping selection must be a compromise between the selection efficiency and the amount of data that passes the selection. The selection cuts of B_s^0 FD $\chi^2 > 121$ and minumum muon IP $\chi^2 > 9$ would increase the $B^0_{(s)} \to \mu^+\mu^-$ selection efficiency by from 71 % to 82 % and the amount of data retained would be doubles. The increase of the data retained by the $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ lines is smaller and the efficiencies are similar to the $B^0_{(s)} \to \mu^+\mu^-$ selection efficiencies. Therefore these cuts are applied in the stripping selection for this analysis. However the increase in efficiency after the stripping selection will not necessarily be propagated through the whole analysis.

3.2.3 Final Selection and Pre-selection

The stripping selection requirements along with a number of other loose selection cuts to select $B^0_{(s)} \to \mu^+\mu^-$, $B^0_{(s)} \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays in 2011, 2012, 2015 and 2016 are shown in Table 3.8. The selection includes the looser stripping cuts on the $B^0_{(s)}$ and J/ψ FD χ^2 and muon and hadron IP χ^2. Additionally the selection of $B^0_{(s)} \to \mu^+\mu^-$ decays includes the momentum, ghost track probabilty and decay time cuts made in the $B^0_{(s)} \to h^+h^-$ stripping line, but were absent in the $B^0_{(s)} \to \mu^+\mu^-$ stripping line. Several cuts are included to remove specific background events, the mass requirement on $B^0_{(s)} \to \mu^+\mu^-$ decays is tightened to ensure $B^0_S \to \mu^+\mu^-\gamma$ decays are not within the mass window and a lower bound is placed on the B meson transverse mometum to remove pairs of muons originating from $pp \to p\mu\mu p$ decays. The decays that pass the selection cuts in Tables 3.8 in both data and simulated events are used to develop tigher selection requirements.

Particle	$B_s^0 \to \mu^+\mu^-$	$B_{(s)}^0 \to h^+h^-$	$B^+ \to J/\psi K^+$
B_s^0 or B^+	4900 MeV/c^2 < M < 6000 MeV/c^2	|M - M_{PDG}| < 500 MeV/c^2	|M - M_{PDG}| < 500 MeV/c^2
	DIRA > 0	DIRA > 0	Vertex χ^2/ndof < 45
	FD χ^2 > 121	FD χ^2 > 121	IP χ^2 < 25
	IP χ^2 < 25	IP χ^2 < 25	
	Vertex χ^2/ndof < 9	Vertex χ^2/ndof < 9	
	DOCA < 0.3 mm	DOCA < 0.3 mm	
	τ < 13.248 ps	τ < 13.248 ps	
	p_T > 500 MeV/c	p_T > 500 MeV/c	
J/ψ			|M - M_{PDG}| < 60 MeV/c^2
			DIRA > 0
			FD χ^2 > 121
			Vertex χ^2/ndof < 9
			DOCA < 0.3 mm
Daugther μ or h	Track χ^2/ndof < 3 (4)	Track χ^2/ndof < 3 (4)	Track χ^2/ndof < 3
	isMuon = True		isMuon = True
	Minimum IP χ^2 > 9	Minimum IP χ^2 > 9	Minimum IP χ^2 > 25
	0.25 GeV/c < p_T < 40 GeV/c	0.25 GeV/c < p_T < 40 GeV/c	
p < 500 GeV/c	p < 500 GeV/c		ghost probability < 0.3 (0.4)
ghost probability < 0.3 (0.4)	heightK^+		
Track χ^2/ndof < 3			
			p_T > 0.25 GeV/c
			Minimum IP χ^2 > 25
			isLong = True

Table 3.8 Loose selection cuts applied to select $B_s^0 \to \mu^+\mu^-$, $B_{(s)}^0 \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays, where selection is different between Run 1 and Run 2 the Run 2 values are shown in parenthesis next to the Run 1 values.

3.2.4 Pre-selection/Offline selection?

The output of the stripping selection still includes many background decays, further cuts shown in Table **??** reduce the background decays. Some selection cuts are designed to remove specific background decays and the selection for $B_s^0 \to \mu^+\mu^-$ decays used in the Branching Fraciton and effective lifetime analyses starts to diverge slightly.

The BDTS is a multivariate classifer that is designed to reduce the number of combinatorial background events. It is a Boosted Decision Tree (BDT) (see Section **??** for a detailed description) that is trained on $B_s^0 \to \mu^+\mu^-$ and $b\bar{b} \to \mu^+\mu^- X$ simulated decays that have passed the $B_{(s)}^0 \to \mu^+\mu^-$ selection requirements in Table **??**. The BDTS uses variables similar to those in the stripping selection to classify events;

- impace parameter χ^2 of the
- vertex χ^2 of the $B_{(s)}^0$
- direction cosine of
- distance of closest apporach of the muons
- minimum impact paramter χ^2 of the muons with respect to all primary verticies in the event
- impact paramter of the $B_{(s)}^0$, this is the distance of closest approch of the B to the primary vertex

The BDTS is applied to all candidates passing the $B_{(s)}^0 \to \mu^+\mu^-$, $B_{(s)}^0 \to h^+h^-$ and $B^+ \to J/\psi K^+$ stripping lines, and candidates are required to have a BDTS value above 0.05. The chosen cut value has a efficiency of X % on $B_s^0 \to \mu^+\mu^-$ decays and reject X % of $b\bar{b} \to \mu^+\mu^- X$ decays.

The semi-leptonic $B_c^+ \to J/\psi\mu^+\nu_\mu$ decays when $J/\psi \to \mu^+\mu^-$ contribute to the background of $B_{(s)}^0 \to \mu^+\mu^-$ decays when a muon from the J/ψ forms a good vertex with the muon from the B_c^+ decay. Due to the high mass of the B_c^+ this could place mis-reconstructed candidates within the B_s^0 mass window. A 'jpsi veto' can be used to remove background events from $B_c^+ \to J/\psi\mu^+\nu_\mu$ decays. The veto works by removing events where one muon from the $B_{(s)}^0 \to \mu^+\mu^-$ candidate combined with any other opositely charged muon in the event has $_{\mu\mu} - m\mathrm{J}/\psi| < 30\ MeV/\mathrm{c}^2$. The veto has a rejection power of X % on $B_c^+ \to J/\psi\mu^+\nu_\mu$ events that have passed $B_{(s)}^0 \to \mu^+\mu^-$ selection cuts in Table **??** and rejects only % of $B_{(s)}^0 \to \mu^+\mu^-$ signal events. The

expected number of $B_c^+ \to J/\psi\mu^+\nu_\mu$ events after the full selection can be found in Section X.

The simiulated $b\bar{b} \to \mu^+\mu^- X$ decays are used to train the multivarite classifier (Sect. **??**) used in the Branching Fraction and effective lifetime analyses and variables included within the classifier (Sect. **??**) as well as the BDTS. The simulated $b\bar{b} \to \mu^+\mu^- X$ decays had tighter cuts on the FD χ^2 of the $B^0_{(s)}$ and minimum IP χ^2 of the muons than those listed Table **??** applied when it was created. Events that did not pass those cuts are not avaliable for use in the classifier training, therefore these must be applied to all data and simulated decays to ensure the most effective and reliable performance of the multivarite classifers used in the analyses.

The effective lifetime analysis requires that $B_s^0 \to \mu^+\mu^-$ candidates have a dimuon invariant mass greater than 5320 MeV/c^2. The motivation comes from mass fit studies that are detailed in Section X. The consequence of this cut is to remove $B^0 \to \mu^+\mu^-$ events and most background from mis-identified semi-leptonic and $B^0_{(s)} \to h^+h^-$ decays. The expect number of $B^0 \to \mu^+\mu^-$ and mis-identified decays after the full selection can be found in Section X.

$B_s^0 \to \mu^+\mu^-$	$B_{(s)}^0 \to h^+h^-$	$B^+ \to J/\psi K^+$
BDTS > 0.05	BDTS > 0.05	BDTS > 0.05
J/ψ veto $_{\mu\mu} - m\mathrm{J}/\psi \mid < 30\ MeV/\mathrm{c}^2$	J/ψ veto $_{\mu\mu} - m\mathrm{J}/\psi \mid < 30\ MeV/\mathrm{c}^2$	-
$B_{(s)}^0$ FD $\chi^2 > 225$	$B_{(s)}^0$ FD $\chi^2 > 225$	J/ψ FD $\chi^2 > 225$
μ minimum IP $\chi^2 > 25$	h minimum IP $\chi^2 > 25$	μ minimum IP $\chi^2 > 25$
M > 5320 MeV/c^2	-	-

Table 3.9 Selection cuts applied to select $B_s^0 \to \mu^+\mu^-$, $B_{(s)}^0 \to h^+h^-$ and $B^+ \to J/\psi K^+$ decays.

3.2.5 Particle Identification

Particle identification (PID) variables are used to refine the selection of $B^0_{(s)} \to \mu^+\mu^-$ candidates and to seperate different $B^0_{(s)} \to h^+h^-$ decays.

In the selection of $B^0_{(s)} \to \mu^+\mu^-$ decays PID variables are particularly useful to reduce the backgrounds coming from mis-identified semi-leptonic decays and $B^0_{(s)} \to h^+h^-$ decays and also help to reduce the number of combinatorial backgroudn events.

The PID requirements to select $B^0_{(s)} \to \mu^+\mu^-$ decays in the Branching Fraction and $B^0_s \to \mu^+\mu^-$ in the lifetime analysis are in Table **??** alongside requirements to seperate $B^0_{(s)} \to h^+h^-$ decays. Two types of PID variables, defined in Section 1.2.2.4, are used; DLL variables and ProbNN variables. DLL variables are useful to seperate $B^0_{(s)} \to h^+h^-$ decays where h is either a pion or kaon because the varaibles compare different particle hypotheses with the pion hypotheses.

Requirements on ProbNN variables vary with the year of data taking because the classifers used in ProbNN variables are tuned to give the best performance depending on the different data taking conditions in the detector for each year.

Tighter PID requirements are used to select $B^0_{(s)} \to \mu^+\mu^-$ decays for the Branching Fraction measurements compared to those used for the effective lifetime measurement. This is because more mis-identidied decays extend in to the mass window than the B^0_s mass window therefore tigher requirements are necessary to reduce them to an acceptable level.

Prehaps a better way for the PID is to put the requirements in the text because the is could be clearer or maybe have two tables one for bsmumu and one for bhh? Also another consideration prehaps it would be clearere to seperate the effective lifetime and the BF after the stripping selection. I could clearly explain that the stripping and loose selection is used for both and therefore the BF decays must be considered in the stripping selection but after that the selection will focus on the effective lifetime.

Bibliography

[1] C. member states. http://home.cern/about/member-states.

[2] S. Amato *et al.*, "LHCb technical proposal," 1998.

[3] A. A. Alves, Jr. *et al.*, "The LHCb Detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[4] "LHCb technical design report: Reoptimized detector design and performance," 2003.

[5] R. Aaij *et al.*, "LHCb Detector Performance," *Int. J. Mod. Phys.*, vol. A30, no. 07, p. 1530022, 2015.

[6] R. Aaij *et al.*, "Performance of the LHCb Vertex Locator," *JINST*, vol. 9, p. 09007, 2014.

[7] R. Aaij *et al.*, "Measurement of the track reconstruction efficiency at LHCb," *JINST*, vol. 10, no. 02, p. P02007, 2015.

[8] M. Adinolfi *et al.*, "Performance of the LHCb RICH detector at the LHC," *Eur. Phys. J.*, vol. C73, p. 2431, 2013.

[9] F. Archilli *et al.*, "Performance of the Muon Identification at LHCb," *JINST*, vol. 8, p. P10020, 2013.

[10] R. Aaij *et al.*, "Absolute luminosity measurements with the LHCb detector at the LHC," *JINST*, vol. 7, p. P01010, 2012.

[11] O. Lupton and G. Wilkinson, *Studies of* $\mathrm{D}^0 \to \mathrm{K}^0_\mathrm{S}\mathrm{h}^+\mathrm{h}'^-$ *decays at the LHCb experiment*. PhD thesis, Oxford U., Jul 2016. Presented 14 Sep 2016.

[12] P. Mato, "GAUDI-Architecture design document," 1998.

[13] R. Antunes-Nobrega *et al.*, *LHCb computing: Technical Design Report*. Technical Design Report LHCb, Geneva: CERN, 2005. Submitted on 11 May 2005.

[14] F. Stagni *et al.*, "LHCbDirac: Distributed computing in LHCb," *J. Phys. Conf. Ser.*, vol. 396, p. 032104, 2012.

[15] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," *Nucl. Instrum. Meth.*, vol. A389, pp. 81–86, 1997.

[16] I. Belyaev, T. Brambach, N. H. Brook, N. Gauvin, G. Corti, K. Harrison, P. F. Harrison, J. He, C. R. Jones, M. Lieng, G. Manca, S. Miglioranzi, P. Robbe, V. Vagnoni, M. Whitehead, J. Wishahi, and the LHCb Collaboration, "Handling of the generation of primary events in Gauss, the LHCb simulation framework," *Journal of Physics: Conference Series*, vol. 331, p. 032047, 2011.

[17] M. Clemencic, G. Corti, S. Easo, C. R. Jones, S. Miglioranzi, M. Pappagallo, and P. Robbe, "The LHCb simulation application, Gauss: Design, evolution and experience," *J. Phys. Conf. Ser.*, vol. 331, p. 032023, 2011.

[18] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 05, p. 026, 2006.

[19] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[20] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth.*, vol. A462, pp. 152–155, 2001.

[21] P. Golonka and Z. Was, "PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays," *Eur. Phys. J.*, vol. C45, pp. 97–107, 2006.

[22] S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003.

[23] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Trans. Nucl. Sci.*, vol. 53, p. 270, 2006.

[24] I. Bird, "Computing for the Large Hadron Collider," *Ann. Rev. Nucl. Part. Sci.*, vol. 61, pp. 99–118, 2011.

[25] W. L. C. Grid. http://www.cern.ch/LHCgrid.

[26] S. Paterson and A. Tsaregorodtsev, "DIRAC Infrastructure for Distributed Analysis," Feb 2006.