LHCb-ANA-20XX-YYY
February 12, 2013

Search for the decays $B^0 \to K_S^0 \eta'$ and $\Lambda_b^0 \to \Lambda^0 \eta'$ with 8 TeV 2012 data.

J. McCarthy[1], N. Watson[1]. C. Lazzeroni[1]

[1] *University of Birmingham, Birmingham, United Kingdom*

Abstract

Will add abstract later

Contents

1 Introduction **1**

2 Dataset **4**

3 Monte Carlo **5**

4 Reconstructing the Decay **5**

5 Trigger **9**

6 Stripping **10**

7 Selection **10**

8 MultiVariate Selection **10**

9 Efficiencies **14**

10 Background Studies **14**

11 Results **16**

11.1 Mass Fits 16

12 Systematics **18**

References **23**

A Stripping Cuts **24**

B Variables and Output from BDT training **26**

1 Introduction

One of the exciting surprises to come from the study of light, neutral mesons is the concept of a mixing of singlet and octet states and the breaking of SU(3) flavour symmetry, most notably with the η and η' mesons. The physical η and η' particles are defined by a mixing of SU(3) singlet ($|\eta_0\rangle$) and octet ($|\eta_8\rangle$) states, defined by the mixing parameter θ_P:

$$\begin{pmatrix} |\eta\rangle \\ |\eta'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta_p & -\sin\theta_p \\ \sin\theta_p & \cos\theta_p \end{pmatrix} \begin{pmatrix} |\eta_8\rangle \\ |\eta_1\rangle \end{pmatrix} \tag{1}$$

where

$$|\eta_0\rangle = \frac{1}{\sqrt{3}}|u\bar{u} + d\bar{d} + s\bar{s}\rangle \tag{2}$$

and

$$|\eta_8\rangle = \frac{1}{\sqrt{6}}|u\bar{u} + d\bar{d} - 2s\bar{s}\rangle \tag{3}$$

Due to the symmetry of SU(3), it is possible, and often more convenient, to represent the particles using a different basis. Here we use the quark flavour basis, defined in [1], with the two flavour states:

$$|\eta_q\rangle = \frac{1}{\sqrt{2}}|u\bar{u} + d\bar{d}\rangle \tag{4}$$

$$|\eta_s\rangle = |s\bar{s}\rangle \tag{5}$$

The $\eta^{(\prime)}$ mesons are then a mixture of a light quark state and a strange quark state, defined with a different mixing angle ϕ_p, by the equation:

$$\begin{pmatrix} |\eta\rangle \\ |\eta\prime\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi_p & -\sin\phi_p \\ \sin\phi_p & \cos\phi_p \end{pmatrix} \begin{pmatrix} |\eta_q\rangle \\ |\eta_s\rangle \end{pmatrix} \tag{6}$$

This representation shows clearly how the different proportions of the strange quark state gives rise to the different masses of the two mesons, since the mass of the strange quark is much larger than the mass of the up or down quarks. Transformations between these two bases is simple due to a relation between the mixing angles [1]

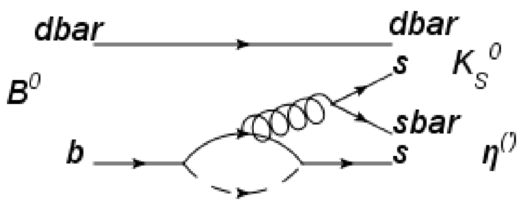
$$\theta_p = \phi_p - \tan^{-1}(\sqrt{2}) \tag{7}$$

Due to the non-perturbative nature of QCD, calculations can be difficult, and so the mixing angle, θ_p, has not been calculated exactly. However, Lattice QCD, has been used to estimate the value of the mixing angle to be $\theta_p \approx -14.1°$, $\phi_p \approx 40.6°$ [2]. This amount of mixing of the quark flavours is surprisingly large when compared to the mixing in equivalent particles in other SU(3) nonets. For example, in the nonet of vector mesons, the mixing of light and heavy flavours is small, with $|\omega\rangle \approx |\eta_q\rangle$ and $|\phi\rangle \approx |\eta_s\rangle$.

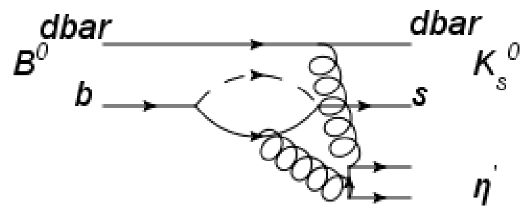
Using this quark flavour basis, one can also introduce a purely gluonic component to the wavefunction. Due to the much smaller mass of the η it is assumed that the gluonic contribution is negligible [1]. The gluonic component is then introduced into the η' wavefunction through a new mixing angle ϕ_G:

$$|\eta\rangle \approx \cos\phi_p|\eta_q\rangle - \sin\phi_p|\eta_s\rangle \tag{8}$$

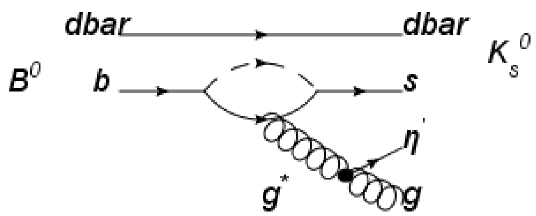
$$|\eta'\rangle \approx \cos\phi_G \sin\phi_p|\eta_q\rangle + \cos\phi_G \cos\phi_p|\eta_s\rangle + \sin\phi_G|gg\rangle \tag{9}$$



(a) Feynman diagram for the decay $B^0 \to K^0\eta^{(\prime)}$



(b) Fenyman diagram for non-spectator contribution for the decay $B^0 \to K^0\eta'$



(c) Feynman diagram for the anomalous coupling contribution to the decay $B^0 \to K^0\eta'$

Figure 1: Feyman diagrams for B decays into $\eta^{(\prime)}$ mesons, showing the enhanced branching fraction to decays to the η' due to the gluonic contribution to the wavefunction

The main consequence of this mixing is the difference in branching fractions for B decays to η and η'. The mixing leads to an enhanced branching ratio for the decays to η' compared with the equivalent decay to η. For example, the branching ratio for the decay $B^0 \to K^0\eta'$ has been measured to be $(6.6 \pm 0.4) \times 10^{-5}$ [3] compared with the branching ratio for $B^0 \to K^0\eta$, which has been measured to be $(1.23^{+0.27}_{-0.24}) \times 10^{-6}$ (54 times smaller!). The same trend has been seen in many other B decays to $\eta^{(\prime)}$. The reason for this is due to the gluonic contribution of the η' wavefunction. This is shown in Figure 1. Figure 2(a) showns the Feynman diagram for the $B^0 \to K^0\eta^{(\prime)}$ decay through the $b \to s$ loop transition. Figure 2(b) shows the non-spectator contribution, where a gluon is radiated from the spectator quark and forms the η' through the gluonic wavefunction. Figure 2(c) shows the η' begin produced via the so-called "anomalous" coupling between the η' and a gluon [4]. Since the last two diagrams are only available through the gluonic component of the wavefunction, it leads to a larger branching ratio for decays to η' over η.

By measuring the relative branching ratios of many different decays to η with respect to η', it is possible to make a measurement of the mixing angle θ_P. This analysis note describes the measurement of the branching ratio measurement for $B^0 \to K^0\eta'$ in pp collisions at $\sqrt{s} = 8$ TeV with the LHCb experiment. The current status of this measurement is presented in Table 1. An improved measurement will lead to a more precise measurement of the mixing angles, and a better understanding of the non-spectator and anomaly models. However, the main aim of this analysis is to use this channel as a control channel to search for the decay $\Lambda_b^0 \to \Lambda^0\eta^{(\prime)}$. Baryonic decays to $\eta^{(\prime)}$ have not yet been observed, and measurements of this type of decay will lead to a better understanding of these models.

Models of QCD have been used to estimate the branching ratio of the Λ_b^0 decay [5]. Depending on the model used, the branching ratio is expected to be between $\approx (4.0 - 19.0) \times 10^{-6}$. The branching ratio will be measured using the B^0 decay as the control channel. By measuring the ratio:

$$R = \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta')}{\mathcal{B}(B^0 \to K^0\eta')} \tag{10}$$

many of the systematic uncertainties and acceptance factors will cancel in the ratio. The branching ratios are calculated using the following formulae:

$$\mathcal{B}(B^0 \to K^0\eta^{'}) = \frac{N_S(B^0)}{2\mathcal{L}_{int}\sigma_{b\bar{b}}f_d\epsilon_{tot}(B^0) \times \mathcal{B}(\eta^{'} \to \rho^0\gamma) \times 0.5 \times \mathcal{B}(K_S^0 \to \pi^+\pi^-)} \tag{11}$$

where $N_S(B^0)$ is the number of signal events which have been observed in the data, and $\mathcal{L}_{int}$ is the total integrated luminosity. $\sigma_{b\bar{b}}$ is the cross section for producing b$\bar{b}$ quarks within the acceptance of the LHCb detector and has been measured to be $\sigma_{b\bar{b}} = (75.4 \pm 5.4 \pm 13.0)\mu$b [6]. f_d is the fraction of b hadrons produced which contain d quarks, i.e. the fraction which are B^0 mesons. The current world average measurement for this parameter is $f_d = 0.404 \pm 0.012$ [7]. $\mathcal{B}(\eta^{'} \to \rho^0\gamma) = 29.4 \pm 0.9\%$ is the branching fraction for the $\eta^{'} \to \rho^0\gamma$ decay and $\mathcal{B}(K_S^0 \to \pi^+\pi^-) = 69.2 \pm 0.05\%$ is the branching fraction for the K_S^0 decay [3]. The factor of 0.5 accounts for the fact that only half of the K^0 are classified as K_S^0. $\epsilon_{tot}(B^0)$ is the total efficiency for selecting signal events, which is determined by applying the selection to a sample of MC simulated signal events. The equation for calculating the Λ_b branching fraction is very similar:

$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta^{'}) = \frac{N_S(\Lambda_b^0)}{2\mathcal{L}_{int}\sigma_{b\bar{b}}f_{\Lambda_b}\epsilon_{tot}(\Lambda_b^0) \times \mathcal{B}(\eta' \to \rho^0\gamma) \times \mathcal{B}(\Lambda^0 \to p^+\pi^-)} \tag{12}$$

Here, $\mathcal{B}(\Lambda^0 \to p\pi^-) = 63.9 \pm 0.5\%$ is the branching fraction for the Λ^0 decay. f_{Λ_b} is the fraction of Λ_b^0 baryons produced from b or $\bar{b}$ quarks. It has been calculated by the CDF collaboration as $f_{\Lambda_b} = 0.227 \pm 0.067$ [8].

Table 1: Summary of measurements of B^0 decays to $\eta^{(\prime)}$, along with PDG average.

	Branching Ratio ($\times 10^{-6}$)			
Decay	**Babar** [9]	**Belle** [10]	**Cleo2** [11]	**PDG Average** [3]
$B^0 \to K_S^0\eta'$	$68.5 \pm 2.2 \pm 3.1$	$58.9^{+3.6}_{-3.5} \pm 4.3$	$89^{+18}_{-16} \pm 9$	66 ± 4
$B^0 \to K_S^0\eta$	$1.15^{+0.43}_{-0.38} \pm 0.9$	$1.27^{+0.33}_{-0.29} \pm 0.8$	N/A	$1.23^{+0.27}_{-0.24}$

Table 2: Summary of datasets used for analysis

Polarity	**Luminosity** (pb^{-1})
Mag-Up	959 ± 34
Mag-Down	1002 ± 35
Total	**1961 ± 69**

The ratio can therefore be calculated using:

$$R = \frac{N_S(\Lambda_b^0)}{N_S(B^0)} \times \frac{\epsilon_{tot}(B^0)}{\epsilon_{tot}(\Lambda_b^0)} \times \frac{f_d}{f_{\Lambda_b}} \times \frac{0.5 \times \mathcal{B}(K_S^0 \to \pi^+\pi^-)}{\mathcal{B}(\Lambda^0 \to p\pi^-)} \tag{13}$$

The ratio of branching fractions can be calculated from the PDG values along with the uncertainty on those measurements:

$$\frac{\mathcal{B}(K_S^0 \to \pi^+\pi^-)}{\mathcal{B}(\Lambda^0 \to p\pi^-)} = 1.083 \pm 0.009 \tag{14}$$

The selection will be optimised using the control channel, and the results from the Λ_b^0 selection will be kept blind (i.e. no results in the mass window of the Λ_b^0 will be shown) until the selection has been reviewed and approved.

2 Dataset

The dataset used corresponds to an integrated luminosity of 1.96 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV recorded by the LHCb experiment [12] during 2012 and processed with Stripping20. The database tags used for the processing of this data are `dddb-20120831` for the detector description database, and `cond-20120831` for the conditions database.

The data sets are summarised in Table 2. The error on the luminosity is taken as 3.5% from [13]

3 Monte Carlo

Monte Carlo has been produced with the same conditions as data in 2012 (MC2012). The tags used for the simulations were `sim-20121025-vc-mu100` for the conditions database and `dddb-20120831` for the detector description database.

The $B^0 \to K_S^0\eta'$, $\Lambda_b^0 \to \Lambda^0\eta'$, $K_S^0 \to \pi^+\pi^-$, $\Lambda^0 \to p\pi^-$ decays are simulated using the phase space (PHSP) model.

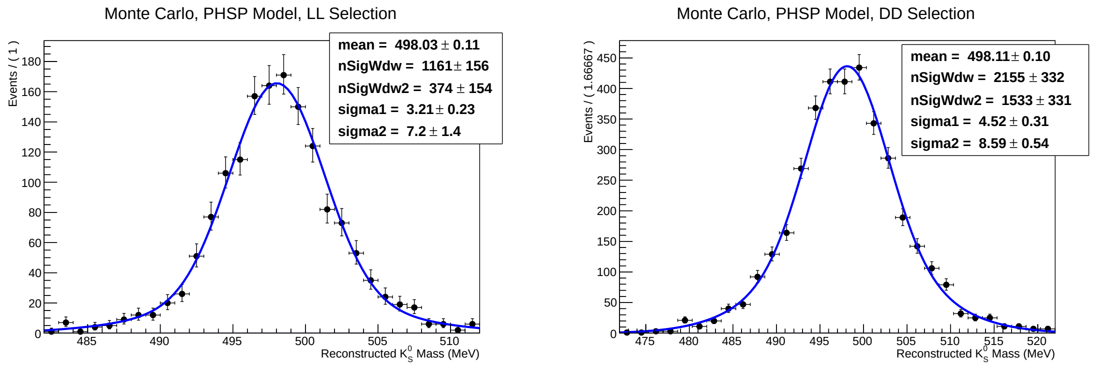
For the η' decay, two models are under investigation. The first is the SVP model, which is designed for decays of a scalar particle into a vector particle and a photon. It is therefore perfect for the $\eta' \to \rho^0\gamma$ decay. The $\rho^0 \to \pi^+\pi^-$ is subsequently modelled with the PHSP model. The worry with this model is that won't correctly model the non-resonant $\pi^+\pi^-\gamma$ contribution. This combination of SVP and PHSP models is not expected to model perfectly the non-resonant $\pi^+\pi^-\gamma$ contribution, which is accepted by the stripping line, another sample is produced which uses the PHSP model for the $\eta' \to \pi^+\pi^-\gamma$ decay. These two different samples will then be compared with data.

Monte Carlo samples are also produced with the same conditions in order to investigate various possible background samples. Each sample is produced with approximately equal numbers of magnet-up and magnet-down conditions. The samples are also split according to the trigger conditions used. Half are simulated with the trigger conditions used before June 2012 (MayJune) and half with the post June trigger (JulySept) Overall approximately 1 million events of each of the signal samples are produced, and 500 000 events of the background samples are produced. Each sample is then processed with the stripping lines of interest.

4 Reconstructing the Decay

The K^0 is reconstructed through its decay $K_S^0 \to \pi^+\pi^-$, which has a branching ratio of 69.2%. The K_S^0 can also decay through $K_S^0 \to \pi^0\pi^0$ with a branching fraction of 30.69%, however, neutral particles like π^0 are difficult to reconstruct as they leave no tracks in the detector. The K_S^0 will then not be reconstructed accurately, and so decays to neutral particles are not considered in this analysis. Also, to my knowledge, K_L^0s have not yet been successfully resonstructed in LHCb.

The K_S^0 is classified according to where it decays in the LHCb detector. If it decays before the VELO then the tracks will be long tracks, so the K_S^0 will be reconstructed as Long-Long (LL). If it decays after the VELO, then it will be reconstructed with Downstream-Downstream (DD) tracks. The stripping and selection optimisation is therefore split into a LL and DD selection.

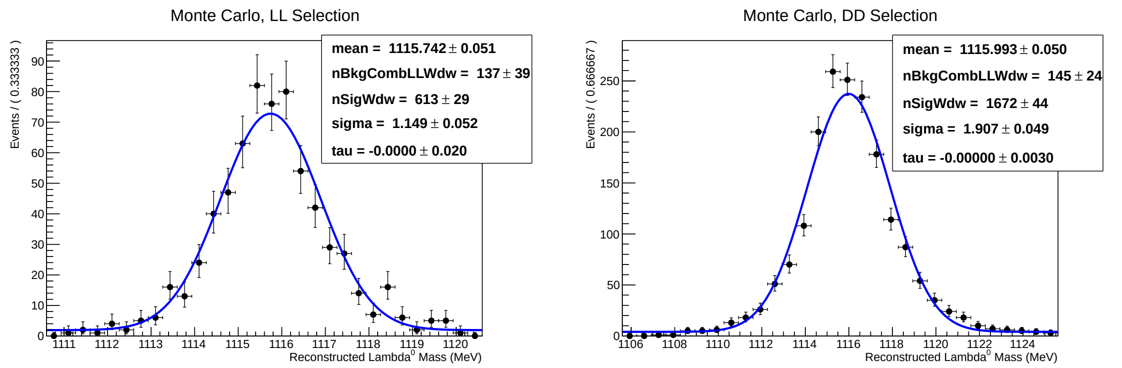


(a) K^0_S reconstructed as Long-Long (LL)

(b) K^0_S reconstructed as Downstream-Downstream (DD)

Figure 2: Reconstructed K^0_S mass from the decay $K^0_S \to \pi^+\pi^-$ in a Monte Carlo sample.

Figure 2 shows the reconstructed mass of K^0_S particles using the track information of the $\pi^\pm$ from a sample of Monte Carlo simulated events. The shape of the mass is fit with 2 gaussians. It is clear that the mass resolution of the K^0_S reconstructed from long tracks is better than that of the downstream tracks. This is because information from the VELO improves the resolution on the momenta of the pions, and that leads to a more precise measurement of the reconstructed mass. The resolution measured is used to choose a mass window around the nominal K^0 mass. For the LL selection, the resolution is ≈ 3.2 MeV, and so a mass window of $M(K^0_S) \pm 10$ MeV is used, while for the DD selection the mass window will be $M(K^0_S) \pm 15$ MeV based on a mass resolution of ≈ 4.5 MeV.



(a) Λ^0 reconstructed as long-long (LL)

(b) Λ^0 reconstructed as downstream-downstream (DD)

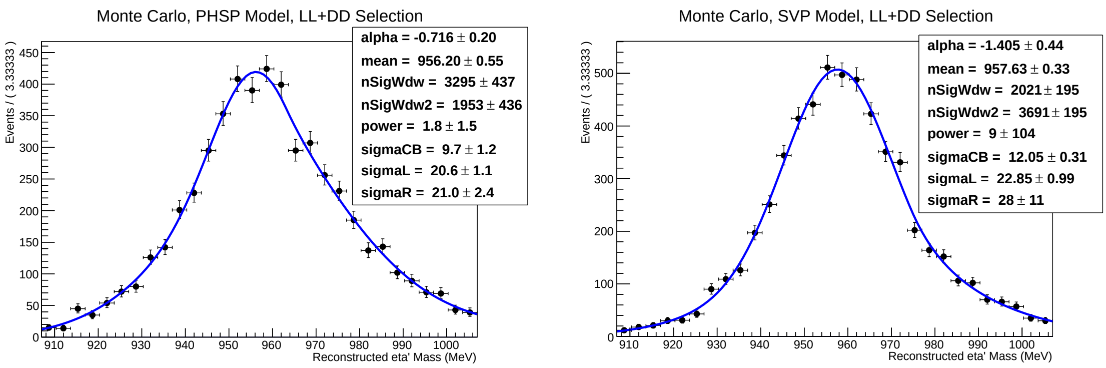
Figure 3: Reconstructed Λ^0 mass from the decay $\Lambda^0 \to p\pi^-$ in a Monte Carlo sample.

The Λ^0 is reconstructed through the decay $\Lambda^0 \to p\pi^-$, which happens with a branching fraction of 63.9%. The Λ^0 also decays through $\Lambda^0 \rightarrow n\pi^0$, with a branching fraction of 35.8%, however, only the decay with charged particles is used.

As with the K_S^0, the Λ^0 is a long lived particle, and so can be reconstructed with both long and downstream tracks. The selection is therefore split into LL and DD selections. The resolution of the reconstructed mass is shown in Figure 3. This resolution shows a factor 3 improvement compared with the resolution of the reconstructed K_S^0. Once again the resoution of the LL selection is better than for the DD selection. The mass window for selecting Λ^0 particles will be $M(\Lambda^0) \pm 4.5$ MeV for the LL selections and $M(\Lambda^0) \pm 6$ MeV for the DD selection.

Table 3: Decays of η' meson [3]

Decay	Branching Fraction
$\eta' \rightarrow \rho^0\gamma$	29.4%
$\eta' \rightarrow \pi^+\pi^-\eta$	44.6%
$\eta' \rightarrow \pi^0\pi^0\eta$	20.7%



(a) η' generated with the PHSP decay

(b) η' generated with the SVP decay

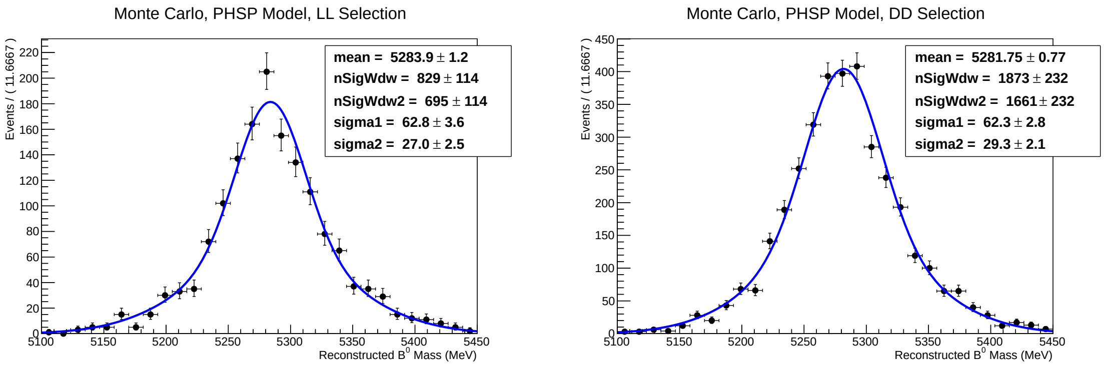
Figure 4: Reconstructed η' mass from $\eta' \rightarrow \pi^+\pi^-\gamma$ in Monte Carlo

There are three main decays which can be used to reconstruct the η', which are summarised in Table 3. Since the decay to $\pi^0\pi^0\eta$ contains only neutral particles, it will not be used at all in this analysis. Initially, only the decay $\eta' \rightarrow \rho^0\gamma$, including the non-resonant $\eta' \rightarrow \pi^+\pi^-\gamma$ will be considered. This is because the η is more difficult to reconstruct, and the overall branching fraction for that decay is lower. However, when higher statistics are required in subsequent analyses, the $\eta' \rightarrow \pi^+\pi^-\eta$ will also be used.

The reconstructed η' in Monte Carlo is shown in Figure 4. Figure 5(a) shows the reconstructed η' from the PHSP decay, while Figure 5(b) shows the SVP sample. The mass is fit with a bifurcated gaussian to model the photon resolution, and a Crystal ball into account the radiative tail of the ρ^0 mass. Since the reconstruction of the η' is not dependant on the K^0_S reconstruction, the samples with LL and DD K^0_S are added together to improve the statistics.

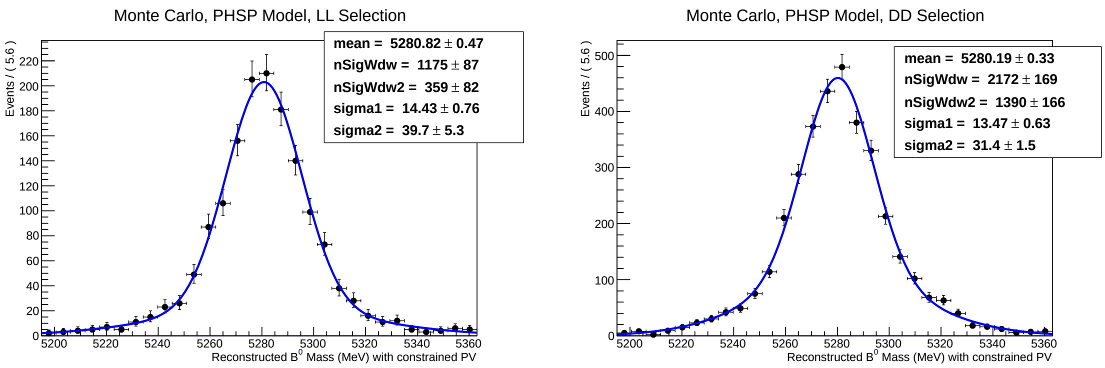
There is a small difference between the two models, with the SVP model giving a slightly narrower shape.

More than 95% of events are reconstructed within a mass window of $M(\eta') \pm 50$ MeV, and so this is the mass window used.



(a) B^0 reconstructed with LL selection (b) B^0 reconstructed with DD selection

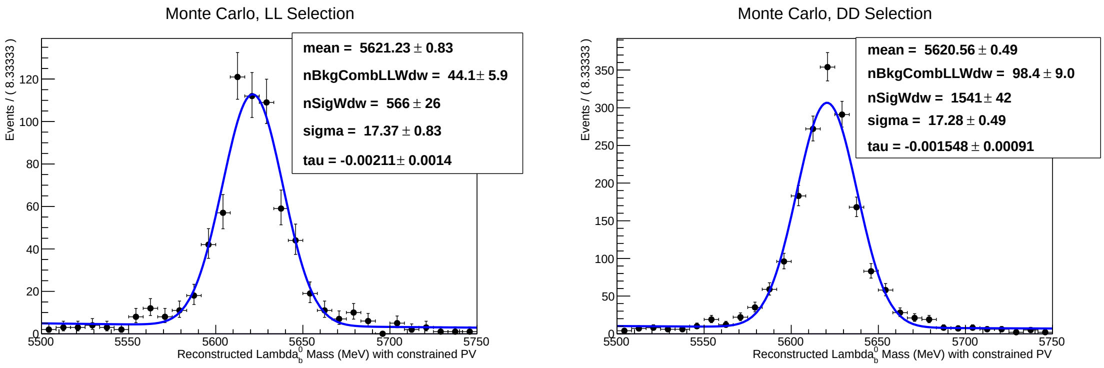
Figure 5: Reconstructed B^0 mass from $B^0 \to \eta' K^0_S$ in Monte Carlo



(a) B^0 reconstructed with LL selection (b) B^0 reconstructed with DD selection

Figure 6: Reconstructed B^0 mass from $B^0 \to \eta' K^0_S$ in Monte Carlo using DecayTreeFitter to constrain the daughter masses

The B^0 mass reconstructed from Monte Carlo samples, is shown in Figure 5, from which the mass resolution is expected to be 40 MeV.



(a) Λ_b^0 reconstructed with LL selection

(b) Λ_b^0 reconstructed with DD selection

Figure 7: Reconstructed Λ_b^0 mass from $\Lambda_b^0 \to \eta'\Lambda^0$ in Monte Carlo

Another method of reconstructing the B^0 is to use the DecayTreeFitter tool to refit the decay constraining the mass of the daughter particles and the primary vertex. By supplying the daughters with information about the B^0, a better fit quality is obtained with an improved resolution. The results of this fit are shown in Figure 6. The mass resolution obtained from this fit is 16 MeV and the separation between signal and background is expected to be better. The Λ_b^0 is also reconstructed using DecayTreeFitter. The mass resolutions are comparable to the B^0 resolution, as shown in Figure 7.

5 Trigger

The analysis makes use of the hadron hardware trigger, and multibody software trigger. The LHCb trigger is described in [14]. The requirements at level 0 are for L0Hadron_TOS or L0Global_TIS to trigger. For the High Level Trigger, the Hlt1TrackAllL0Decision line is required to trigger as TOS, and the topological Hlt2Topo2,3,4BodyBBDTDecision_TOS lines are required to be triggered. Due to the presence of neutral daughters in the decay, the performance of the trigger lines is not as high as other analyses. The efficiencies will be presented in Section 9.

The dataset is split into two distinct periods. Prior to the Technical Stop in June 2012, a bug was present in the Hlt trigger, such that the trigger did not perform as well as expected for K_S^0 reconstructed as DD. This bug was removed during the June TS, and so the efficiencies will be calculated separately for the two data taking periods.

6 Stripping

The stripping line used are the B2XEta lines in Stripping 20. For the $B^0 \to K_S^0\eta'$ the StrippingB2XEtaB2etapKSLL(DD)Line lines are used for the K_S^0 reconstructed with LL (DD) tracks. For the $\Lambda_b \to \Lambda^0\eta'$, the StrippingB2XEtaLb2etapLLL(DD)Line for Λ^0 reconstructed with LL (DD) tracks was used.

The selection applied by the B2XEta stripping lines is summarised in Table 15 in Appendix A.

7 Selection

A further selection is applied on top of the stripping in order to reduce the number of background events while retaining good efficiency for signal events. The number of combinatoric candidates is reduced, primarily by placing a cut on the ghost probability of the p and $\pi^{\pm}$ from the K_S^0, Λ^0 and η', such that the track $\chi^2 < 0.4$. Also a cut is placed on the decay vertex of the B^0, so that the B^0 ENDVERTEX$\chi^2 < 20$. Finally, kinematic cuts are placed on some of the daughters: $B^0 p_T > 2$ GeV, $B_p^0 > 20$ GeV, $\eta' p_T > 2.5$ GeV and $\gamma p_T > 300$ MeV. To reduce the number of partially reconstructed backgrounds, a cut on the PID of the photon is applied: $-3 < \gamma_{PID} < 3$.

8 MultiVariate Selection

A boosted decision tree (BDT) [15] [16] is used to improve the separation between signal and background. The TMVA tool is used to train and apply the BDT to the data.

In the training stage, a sample of pure signal events and a sample of pure background events are supplied to the TMVA. The Monte Carlo are produced with 2012 conditions, and processed with Stripping20, and each particle is matched with MC truth information. For the background sample, a random 10% of the 2012 data is used, using the upper mass sideband only (i.e. reconstructed mass M(B^0)> 5400 MeV). The samples are randomly divided, with one half being used for training the BDT, and one half being used for testing. The BDT is trained with a maximum depth of 2 for each tree, and the LL and DD selections are trained separately, with 400 trees for the DD selection and 200 trees for the LL selection which suffers from lower MC statistics.

In order to ensure the samples used for training match the data as closely as possible, the selection is applied prior to the training. Ideally, the trigger line cuts would also be applied, but this would not leave enough statistics in the Monte Carlo sample for training. The number of events used for training is summarised in Table 4.

Table 4: The number of events used in training the BDT

Selection	Number of Signal	Number of Background
B^0 LL	1471	6913
B^0 DD	2753	20998

Table 5: Variables used in training the BDT

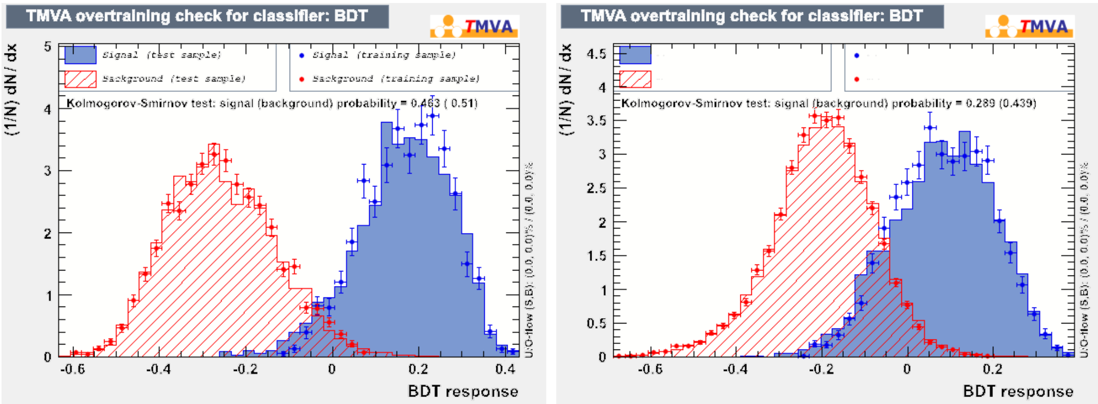
Particle	**Variables**
B^0 (Λ_b^0)	p_T, log(FD χ^2), log($\tau\chi^2$), log(1-DIRA Angle), End Vertex χ^2
K_S^0 (Λ^0)	P, log(IP χ^2), log(FD χ^2)
$\pi^\pm$ (p)	log(IP χ^2)
η'	p_T, log(IP χ^2
γ	log(p_T)

The variables used to train the BDT are given in Table 5. The K_S^0 impact parameter and flight distance, and the impact parameter of the $\pi^\pm$ from the K_S^0 are not used in the training of the DD BDT, since the VELO information are not available, and so the measurements are less precise. The same variables are used for the Λ_b^0 selection.

The distribution in each variable for signal and background are shown in Appendix B, for the B^0 selection (Figure 15 for the LL selection and Figure 16 for the DD selection) and for the Λ_b^0 selection (Figure 17 for the LL selection and Figure 18 for the DD selection).

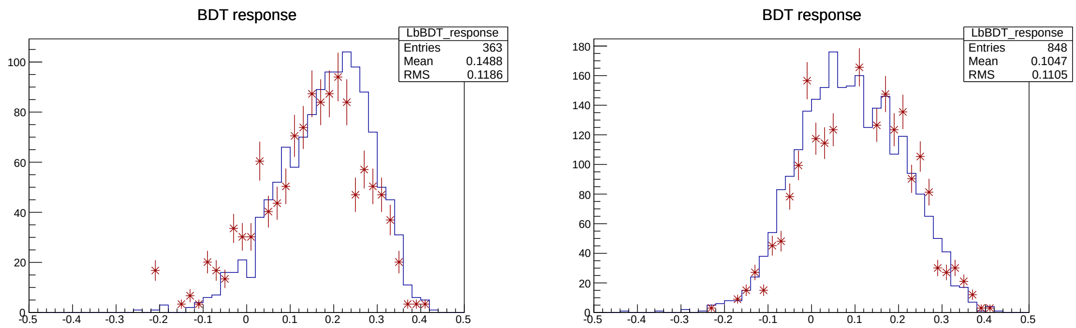
The variables are combined by the BDT into one powerful variable, called the BDT response. The BDT response for the LL and DD selections are shown in Figure 8. The response of background events are shown in red, and the response of signal events are shown in blue. The response for the training samples are superimposed as points to show there is no overtraining of the BDT.

The statistics available from the Λ_b^0 Monte Carlo are very low, due to the reconstruction and stripping efficiency of for the Λ_b^0 selection. Only 646 events are available for training the LL selection, which is insuficient to train a BDT. Therefore the BDT which was trained for the B^0 is applied for the Λ_b^0 selection. To check this is valid, the variables used in the BDT need to be compared for both the B^0 signal and Λ_b^0. These variables are shown plot together in Appendix B for both the LL in Figure 17 and DD selection in Figure 18. The BDT output for the B^0 and Λ_b^0 are shown in Figure 9. The response is very similar for both selections, confirming that we can use the same BDT for both selections.



(a) B^0 LL BDT Response (b) B^0 DD BDT Response

Figure 8: The BDT response of signal events (blue) and background events (red). The test sample is superimposed.



(a) LL Selection (b) DD Selection

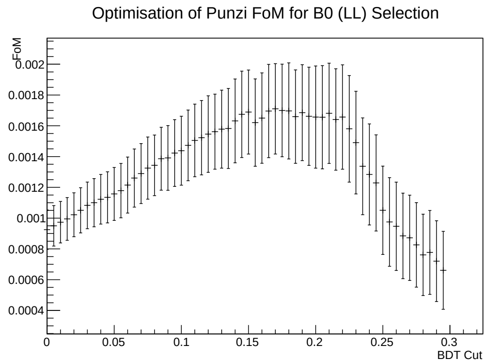
Figure 9: Output from the BDT trained on the B^0 selection, for B^0 (histogram) and for the Λ_b^0 (points) overlaid. The output of the two is close enough to be confident using the BDT for the Λ_b^0 selection.

The BDT is then applied to the data, and each event is assigned a BDT response based on how “signal-like” the event is. A cut can then be placed on the BDT response to reduce the number of background events and improve the purity of the signal observed.

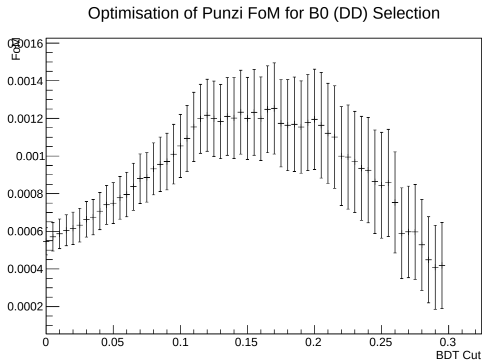
The optimal BDT cut is chosen by optimisation the traditional Punzi Figure of Merit, defined as:

$$\text{FoM} = \frac{\epsilon_{\text{MVA}}}{\frac{a}{2} + \sqrt{B}} \tag{15}$$

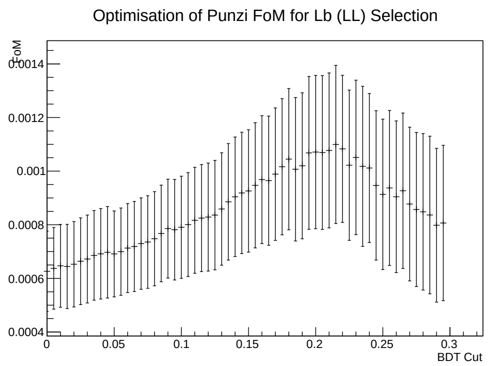
where ϵ_{MVA} is the selection efficiency for a particular BDT cut. B is the number of combinatoric bckground events passing the BDT cut. This is evaluated by extrapolating the exponential shape from the sidebands into the signal mass window. a is defined as the significance of signal required, in this case a=5 (corresponding to a significance of 5σ. The results of this optimisation for each selection is shown in Figure 10, and the optmimum cuts are summarised in Table 6.



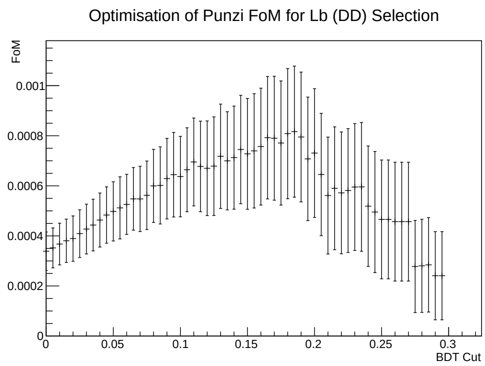
(a) B^0 LL Selection



(b) B^0 DD Selection



(c) Λ_b^0 LL Selection



(d) Λ_b^0 DD Selection

Figure 10: Optimisation of Punzi Figure of Merit as a function of BDT cut for different selections

Table 6: Optimum BDT cuts

Selection	**Optimum BDT Cut**
B^0 LL	0.14
B^0 DD	0.12
Λ_b^0 LL	0.2
Λ_b^0 DD	0.17

9 Efficiencies

The efficiencies have been measured using Monte Carlo samples which have been simulated with the same conditions as data in 2012. The efficiencies are summarised in Table 7 for the selection efficiencies and Table 8 for the trigger efficiencies. The trigger efficiencies are calculated for the Pre-June data taking period and for the Post-June period separately. They are then combined (weighted by the fraction of data taken in each period) for the total trigger efficiency. The errors shown are statistical only, a study of systematic uncertainties of the efficiencies is presented in Section 12. The stripping efficiency includes the efficiency of the reconstruction of the events and the selection of the stripping line. Each efficiency is calculated as the efficiency with respect to the the previous selection, such that the total efficiency is the product of all efficiencies.

With these efficiencies it is then possible to estimate the number of signal events we expect to see in the data used. This is shown in Table 9. For the Λ_b^0 decay, the branching ratio is assumed to be 20×10^{-6}. Clearly, with these efficiencies, and assuming this branching ratio it will probably not be possible to observe these decays in the current data sample.

Table 7: Summary of stripping, selection, MVA and total selection efficiencies for 2012 data

Selection	$\epsilon_{\mathrm{strip}}$ (%)	$\epsilon_{\mathrm{sel\|strip}}$ (%)	$\epsilon_{\mathrm{mva\|sel}}$ (%)	ϵ_{tot} (%)
$B^0 \to K_S^0(\mathrm{LL})(\eta' \to \rho^0\gamma)$	0.224 ± 0.005	54.0 ± 1.0	80.5 ± 1.1	0.0970 ± 0.003
$B^0 \to K_S^0(\mathrm{DD})(\eta' \to \rho^0\gamma)$	0.536 ± 0.007	$45.3 \pm 0.0.7$	25.9 ± 0.9	0.063 ± 0.002
$\Lambda_b^0 \to \Lambda^0(\mathrm{LL})(\eta' \to \rho^0\gamma)$	0.075 ± 0.004	45.6 ± 2.5	93.1 ± 1.92	0.032 ± 0.002
$\Lambda_b^0 \to \Lambda^0(\mathrm{DD})(\eta' \to \rho^0\gamma)$	0.204 ± 0.004	39.0 ± 1.1	24.8 ± 1.5	0.019 ± 0.001

10 Background Studies

Three different categories of backgrounds were considered for this analysis: the combinatoric background due to random tracks produced in the collisions, the peaking background which will be reconstructed with the mass of the signal particle, and the partially reconstructed backgrounds which are reconstructed in the left hand sideband of the mass plot. An exponential fit is used to model the combinatoric background.

The peaking bacgrounds investigated are $B^0 \to K_S^0\pi^+\pi^-\gamma$ and $\Lambda_b^0 \to \Lambda^0\pi^+\pi^-\gamma$, where the $\pi^+\pi^-\gamma$ are not produced from the η' resonance. 500k events are simulated with 2012 conditions. No events pass the selection and so there are not expected to be any peaking backgrounds in the results.

Table 8: Summary of trigger efficiencies for 2012 data

Selection	**Period (%)**	$\epsilon_{\mathrm{L0|MVA}}$ (%)	$\epsilon_{\mathrm{Hlt1|L0}}$ (%)	$\epsilon_{\mathrm{Hlt2|Hlt1}}$ (%)	$\epsilon_{\mathrm{Trig|MVA}}$ (%)
B^0(LL)	Pre-June	39.8 ± 2.2	66.5 ± 3.3	88.0 ± 2.8	23.2 ± 1.9
B^0(LL)	Post-June	36.4 ± 2.1	69.9 ± 3.4	64.8 ± 4.2	16.5 ± 1.6
B^0(LL)	Combined				17.6 ± 1.7
B^0(DD)	Pre-June	46.4 ± 2.8	60.9 ± 4.0	40.2 ± 5.1	11.4 ± 1.8
B^0(DD)	Post-June	44.5 ± 2.8	62.8 ± 4.0	76.9 ± 4.4	21.5 ± 2.3
B^0(DD)	Combined				19.79 ± 2.2
Λ_b^0(LL)	PreJune	46.9 ± 3.9	68.4 ± 5.3	82.7 ± 5.2	26.5 ± 3.5
Λ_b^0(LL)	PostJune	43.8 ± 3.8	78.9 ± 4.8	67.9 ± 6.2	23.5 ± 3.3
Λ_b^0(LL)	Combined				23.45 ± 3.4
Λ_b^0(DD)	PreJune	68.3 ± 4.6	43.5 ± 6.0	23.2 ± 7.7	6.9 ± 2.5
Λ_b^0(DD)	PostJune	59.4 ± 4.9	65.0 ± 6.2	64.1 ± 7.6	24.7 ± 4.2
Λ_b^0(DD)	Combined				21.8 ± 4.1

Table 9: Summary of expected yield in 1.96 fb^{-1} 2012 data

Selection	**Expected Yield**
$B^0 \to K_S^0(\mathrm{LL})(\eta' \to \rho^0\gamma)$	147
$B^0 \to K_S^0(\mathrm{DD})(\eta' \to \rho^0\gamma)$	106
$\Lambda_b^0 \to \Lambda^0(\mathrm{LL})(\eta' \to \rho^0\gamma)$	20
$\Lambda_b^0 \to \Lambda^0(\mathrm{DD})(\eta' \to \rho^0\gamma)$	11

Table 10 summarises the partially reconstructed background considered for the B^0 analysis. Since the probability of a π^0 beign misidentified as a γ is small, the only decay investigated is the $B^0 \to (K^* \to K^0\pi^0)\eta'$ where the π^0 is not reconstructed. Monte Carlo is simulated for this decay and the efficiency for passing the selection was calculated. Using this, the number of events in 1.96 fb^{-1} is expected to be less than 2. This would not show up in the fit to the sidebands, and so the partially reconstructed backgrounds are neglected.

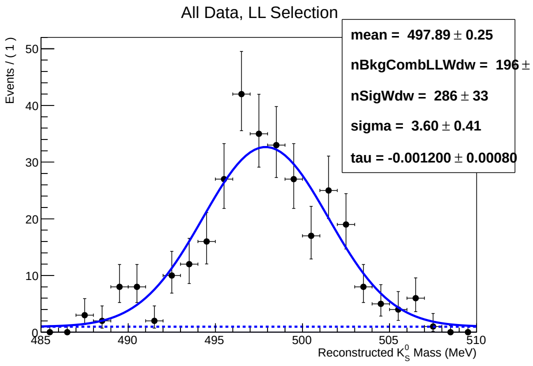
Table 10: Possible partially reconstructed backgrounds considered

Process	**Condition**	**B.r.**
$B^0 \to (K^* \to K^0\pi^0)\eta'$	π^0 not reconstructed	3.1e-6
$B^0 \to (D^- \to K_S^0\pi^+\pi^0)\pi^-$	π^0 mis-ID as γ	3.88e-5
$B^0 \to (D^0 \to K_S^0\pi^0)\pi^+\pi^-$	π^0 mis-ID as γ	1e-5
$B^0 \to (D^0 \to K_S^0\pi^0)\eta'$	π^0 not reconstructed	1e-6
$B^0 \to K_s^0(\omega \to \pi^+\pi^-\pi^0)$	π^0 mis-ID as γ	4.4e-6

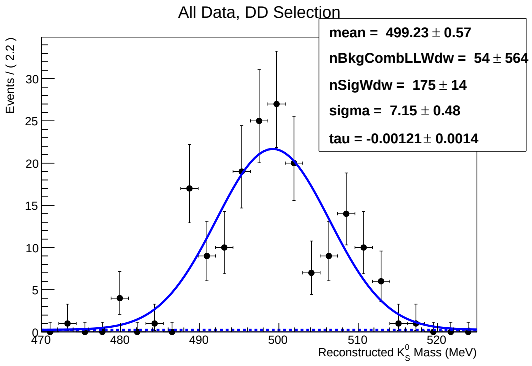
11 Results

The selection is applied to the data, and this section shows the fit to the reconstructed particles, and the yields obtained.

11.1 Mass Fits



(a) K^0_S reconstructed as Long-Long (LL) in data



(b) K^0_S reconstructed as Downstream-Downstream (DD) in data

Figure 11: Reconstructed K^0_S mass in data.

Figure 11 shows the reconstructed K^0_S after the selection is applied. Figure 12(a) shows the K^0_S reconstructed as LL, which is fit with a gaussian with a mass resolution of 3.6 MeV. Figure 12(b) shows the K^0_S reconstructed as DD, which is fit with a gaussian with a mass resolution of 7.15 MeV. Both fits show the K^0_S are reconstructed with a resolution which is consistent with the Monte Carlo samples.

Figure 12 shows the reconstructed η' mass fit with a Breit-Wigner curve to the data from the LL and DD selections combined. The width of the Breit-Wigner curve is 34.8±7.6 MeV. The width has a large uncertainty, and so is consistent with both models for the η' decay, however, the central value is closer to that of the phase space model.

Figure 13 shows the reconstructed mass of the B^0 after the selection has been applied, where the decay has been refitted using constraints on the mass of the daughter particles and the primary vertex. The yields determined for these decays are shown in Table 11. In both cases, the background level seems to be well modelled by an exponential fit, confirming the expectation that there is no partially reconstructed background to be concerned with. The same plots without the refitting of the decay are shown in Figure 14 to highlight the improvement available by refitting.

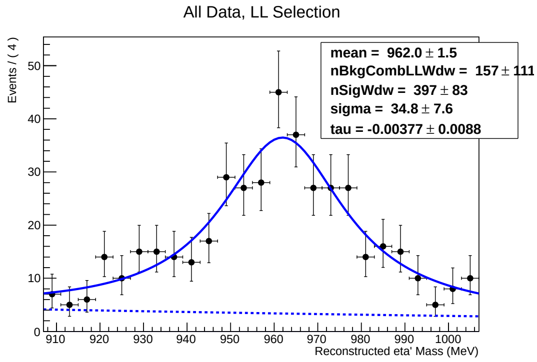


Figure 12: Reconstructed η' mass in data

Table 11: Yields determined from the fit to data

Decay	Selection	N Sig	N Bkg	Significance	mean (MeV)	width (MeV)
$B^0 \to K^0_S\eta'$	LL	177 ± 19	145.4 ± 5.5	9.84	5280.8 ± 3.6	30.1 ± 4.1
$B^0 \to K^0_S\eta'$	DD	100 ± 14	101.8 ± 4.3	7.05	5286.1 ± 4.6	28.7 ± 4.1
	Total	277	147.2	13.44		

Figure 14(a) shows the output of the LL selection. The background is described well by an exponential, which models the combinatoric background. There are 177 signal events described by a gaussian with a mass resolution of 30.1 ± 4.1 MeV, which is consistant with the expectation from the Monte Carlo samples in Figure **??**. The number of signal events is consistent with that expected from the efficiency calculations in Table 9.

Figure 15(b) shows the Output of the DD selection. The background is again described by an exponential, which models the combinatoric background. The background level is higher than for the LL selection, which is due to the poorer performance of the BDT for the DD selection. There are 100 signal events described by a gaussian with a mass resolution of 28.7 ± 4.1 MeV, which is again consistant with the expectation from the Monte Carlo samples in Figure 5. The signal observed has a statistical significance of 7.05σ. Once again the number of signal events is consistent with the expectation.

The statistical significance of the two selections combined is 13.44σ

12 Systematics

Various systematic uncertainties are considered when calculating the ratio of branching fractions given in equation 13. Each term will have a systemiatic uncertainty associated with it.

For the measured branching ratios, the systematic uncertainty is calculated from the uncertainties given in the PDG [3]. The systematic uncertainty of the ratio of branching fractions is therefore 0.83%.

The number of signal events determined for each decay, the dominant systematic uncertainty will be due to the model used to fit the data. Therefore different models will be used to model the data, and the spread of signal yields extracted will give the systematic uncertainty. The signal shape will be fit using a single gaussian, a double gaussian, a crystal ball and a breit wigner curve, each time using an exponential for the fit to the background. The background will then be fit using an exponential, a $2^{\rm nd}$ and $3^{\rm rd}$ order polynomial, and a $2^{\rm nd}$ and $3^{\rm rd}$ order ChebyChev polynomial, using a single gaussian to fit the signal shape. To add further checks, instead of allowing all parameters to float in the fit, the mass of the signal particle will the fix to the PDG value, the width will be fixed to the expectation from Monte Carlo and then the mass and width will be fixed to that measured by the other selection (i.e. the DD selection for LL). In each case, the fit will only be considered in the systematic uncertainty if the χ^2 of the fit is reasonable. Finally, the size of the mass window used will the varied from 2.5 times the mass resolution to 3.5 times the mass resolution to observe the variation in the number of signal events. The results of these checks are summarised in Table 12. The same checks will be applied to the Λ_b^0 after unblinding, and a systematic uncertainty will be assigned to the ratio.

The systematic uncertainties due to the efficiencies measured are also calculated. They are separated into the uncertainties on the selection efficiency and the trigger efficiency.

The selection efficiency will be calculated for different cuts, and different Monte Carlo samples. The kinematic cuts are varied by 1%, corresponding approximately to the resolution of the detector. The PID cuts on the photon and the BDT cuts are also varied by the same amount to test the variation in the selection efficiencies. The efficiencies are also calculated with different Monte Carlo samples, using the MayJune and JulySept samples separately, and also separating out the magnet-up and magnet-down polarities. The ratio of selection efficiencies for the B^0 and Λ_b^0 is calculated, and any variation in this ratio larger than the statistical uncertainties will be considered as a systematic uncertainty. The results of these tests are summarised in Table 13. Each systematic uncertainty is

Systematic errors on trigger efficiency

Systematic errors on f_l/f_d. How to deal with pT dependence?

Table 12: Variation of the number of signal events due to the fit model used.

Model/Test	$N_{B^0}(LL)$	**Fit** χ^2	$N_{B^0}(DD)$	**Fit** χ^2
Sig: Single Gaussian				
Sig: Double Gaussian				
Sig: Gaussian + Crystal Ball				
Sig: Breit Wigner				
Bkg: 2nd order Poly				
Bkg: 3rd order Poly				
Bkg: 2nd order ChebyChev				
Bkg: 3rd order ChebyChev				
Fix mass to PDG				
Fix width to MC				
Fix mass to other selection				
Fix width to other selection				
Mass window 2.5σ				
Mass window 3.5σ				

Lb lifetime and polarisation

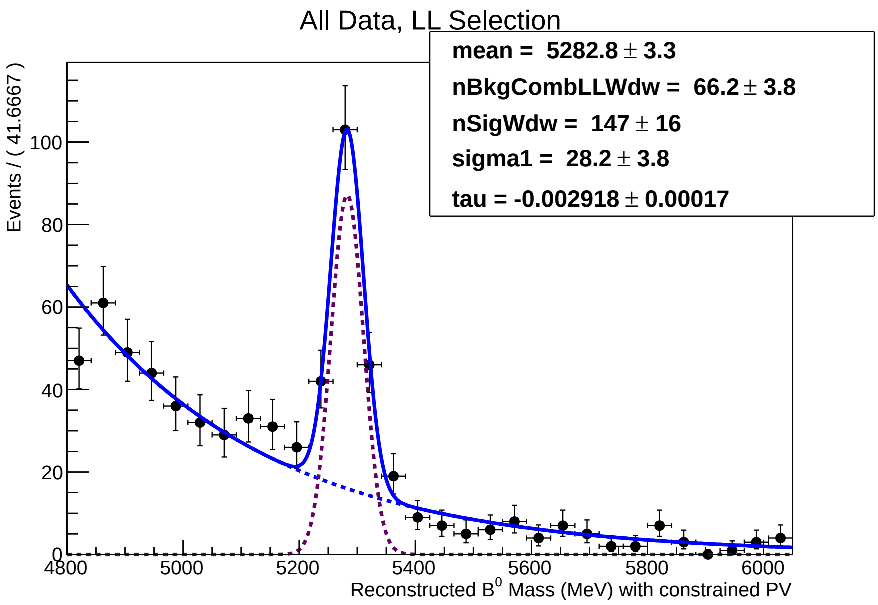
A summary of all systematic uncertainties is shown in Table 14. The errors are presented as percentages, and the total systematic error is a quadrature sum of all the systematic uncertainties.

Table 13: Variation of the selection efficiency calculated from Monte Carlo based on the cuts used, the data taking period, and the magnet polarity. Default cuts are the those used for the analysis

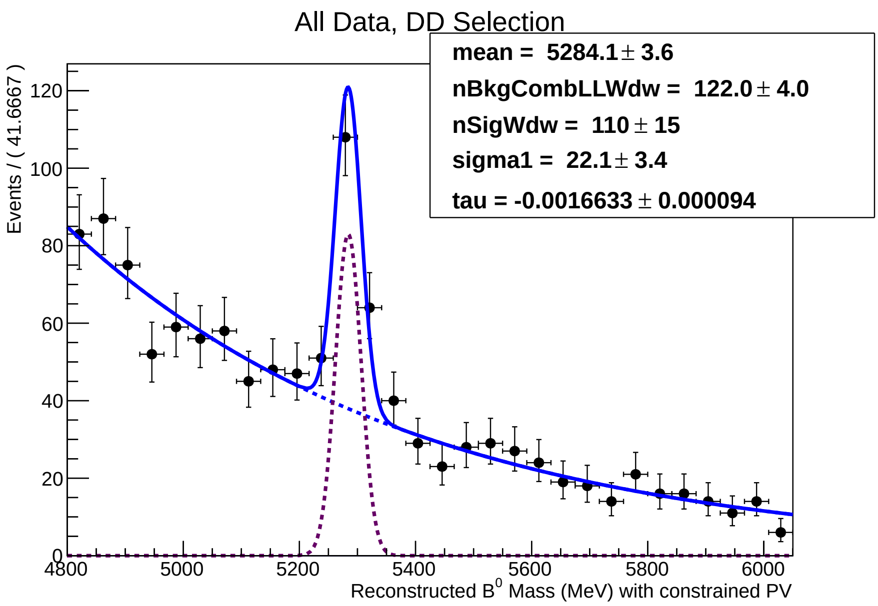
Sel	**Samp**	**Period**	**Mag**	**Cuts**	$\frac{\epsilon_{tot}(B^0)}{\epsilon_{sel}(\Lambda_b^0)}(LL)$	err	$\frac{\epsilon_{tot}(B^0)}{\epsilon_{sel}(\Lambda_b^0)}(DD)$	err
LL	PHSP	Both	Both	Default				
				Loose kinematic				
				Tight kinematic				
				Loose PID				
				Tight PID				
				Loose BDT				
				Tight BDT				
		May-June	Both	Default				
			up	Default				
			down	Default				
		July-Sept	Both	Default				
			up	Default				
			down	Default				
	SVP	Both	Both	Default				

Table 14: Summary of systematic uncertainties

Systematic effect	**Uncertainty (%)**
Measured b.r pf K_S^0 and Λ^0	0.83
Measured ratio of production fractions.	
Ratio of ϵ_{sel} due to...	
Ratio of ϵ_{trig} due to...	
Ratio of signal yields due to fit	
Total	

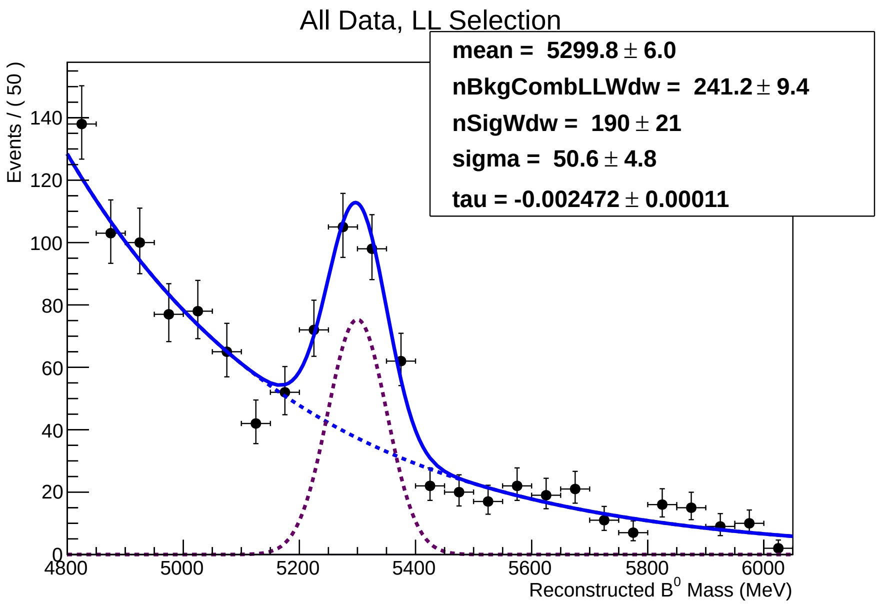


(a) Reconstructed B^0 with LL selection

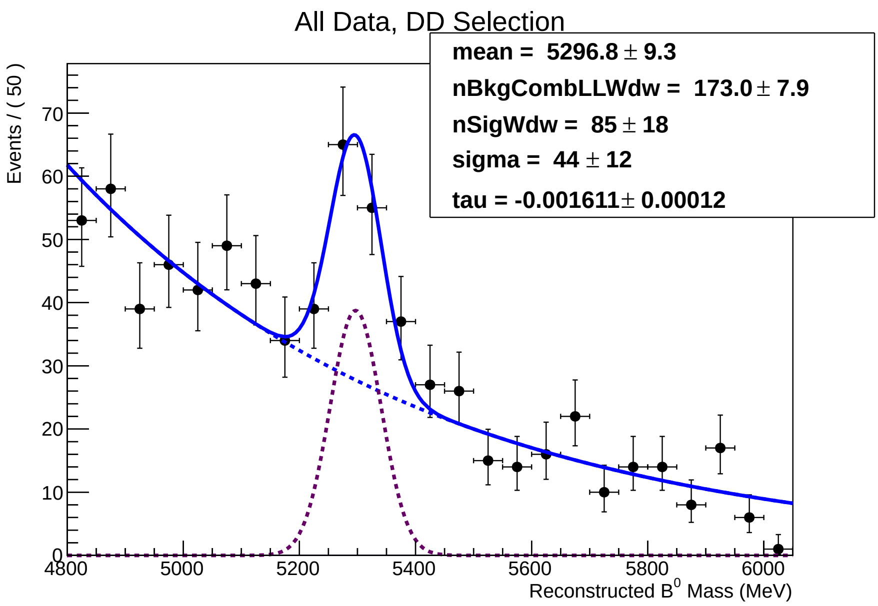


(b) Reconstructed B^0 with DD selection

Figure 13: Reconstructed B^0 mass in data with the decay refitted with constraints on the mass of the daughter particles and on the primary vertex.



(a) Reconstructed B^0 with LL selection



(b) Reconstructed B^0 with DD selection

Figure 14: Reconstructed B^0 mass in data.

References

[1] C. Di Donato, G. Ricciardi, and I. Bigi, *$\eta - \eta'$ Mixing - From electromagnetic transitions to weak decays of charm and beauty hadrons*, Phys. Rev. **D85** (2012) 013016, `arXiv:1105.3557`.

[2] N. Christ *et al.*, *The η and η' mesons from Lattice QCD*, Phys. Rev. Lett. **105** (2010) 241601, `arXiv:1002.2999`.

[3] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[4] D. Atwood and A. Soni, *$B \to \eta' + X$ and the QCD anomaly*, Phys. Lett. **B405** (1997) 150, `arXiv:hep-ph/9704357`.

[5] M. Ahmady, C. Kim, S. Oh, and C. Yu, *Heavy baryonic decays of $Lambda_b \to Lambda\eta'$*, Phys. Lett. **B598** (2004) 203, `arXiv:hep-ph/0305031`.

[6] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, `arXiv:1009.2731`.

[7] Heavy Flavor Averaging Group, D. Asner *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton Properties*, `arXiv:1010.1589`.

[8] CDF Collaboration, T. Aaltonen *et al.*, *Measurement of Ratios of Fragmentation Fractions for Bottom Hadrons in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96-TeV*, Phys. Rev. **D77** (2008) 072003, `arXiv:0801.4375`.

[9] BABAR Collaboration, B. Aubert *et al.*, *B meson decays to charmless meson pairs containing eta or eta' mesons*, Phys. Rev. **D80** (2009) 112002, `arXiv:0907.1743`.

[10] Belle Collaboration, J. Schumann *et al.*, *Evidence for $B \to$ eta-prime pi and improved measurements for $B \to$ eta-prime K*, Phys. Rev. Lett. **97** (2006) 061802, `arXiv:hep-ex/0603001`.

[11] CLEO Collaboration, S. Richichi *et al.*, *Two-body B meson decays to eta and eta-prime: Observation of $B \to$ eta K**, Phys. Rev. Lett. **85** (2000) 520, `arXiv:hep-ex/9912059`.

[12] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[13] LHCb collaboration, R. Aaij *et al.*, *Absolute luminosity measurements with the LHCb detector at the LHC*, JINST **7** (2012) P01010, `arXiv:1110.2866`.

[14] R. Aaij *et al.*, *The LHCb trigger and its performance*, `arXiv:1211.3055`.

[15] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[16] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, `arXiv:physics/0408124`.

A Stripping Cuts

Table 15: Summary of stripping cuts in Stripping 20

Particle	**Variable**	**Cut**
B^0	m(B^0)	±750 MeV
	Vtxχ^2	<12
	p_T	>800 MeV
	DOCA χ^2	<20
	DIRA	>0.999
	IP χ^2	<20
	FD χ^2	>20
Λ_b^0	m(Λ_b^0)	±750 MeV
	Vtxχ^2	<12
	p	>6 GeV
	DOCA χ^2	<20
	DIRA	>0.999
	IP χ^2	<20
η'	m(η')	±50 MeV
	Vtxχ^2	<10
	DOCA χ^2	<12
	p_T	>800 MeV
η	m(η)	±50 MeV
	p_T	>600 MeV
$K_S^0(LL)$	m(K^0)	±18 MeV
	Vtxχ^2	<12
	FDχ^2	>15
	p_T	>1 GeV
	DOCA χ^2	<25
$K_S^0(DD)$	m(K^0)	±28 MeV
	Vtxχ^2	<12
	FDχ^2	>500
	p_T	>1 GeV
	DOCA χ^2	<25
$\Lambda^0(LL)$	m(Λ^0)	±15 MeV
	Vtxχ^2	<12
	p_T	>1.5 GeV
	DOCA χ^2	<30
$\Lambda^0(DD)$	m(Λ^0)	±20 MeV
	Vtxχ^2	<12
	p_T	>1.5 GeV
	DOCA χ^2	<25
γ	p_T	>200 MeV
Tracks $\pi^{\pm}$,p	Track χ^2	<3
	p_T	>400 MeV
	IPχ^2	>20

B Variables and Output from BDT training

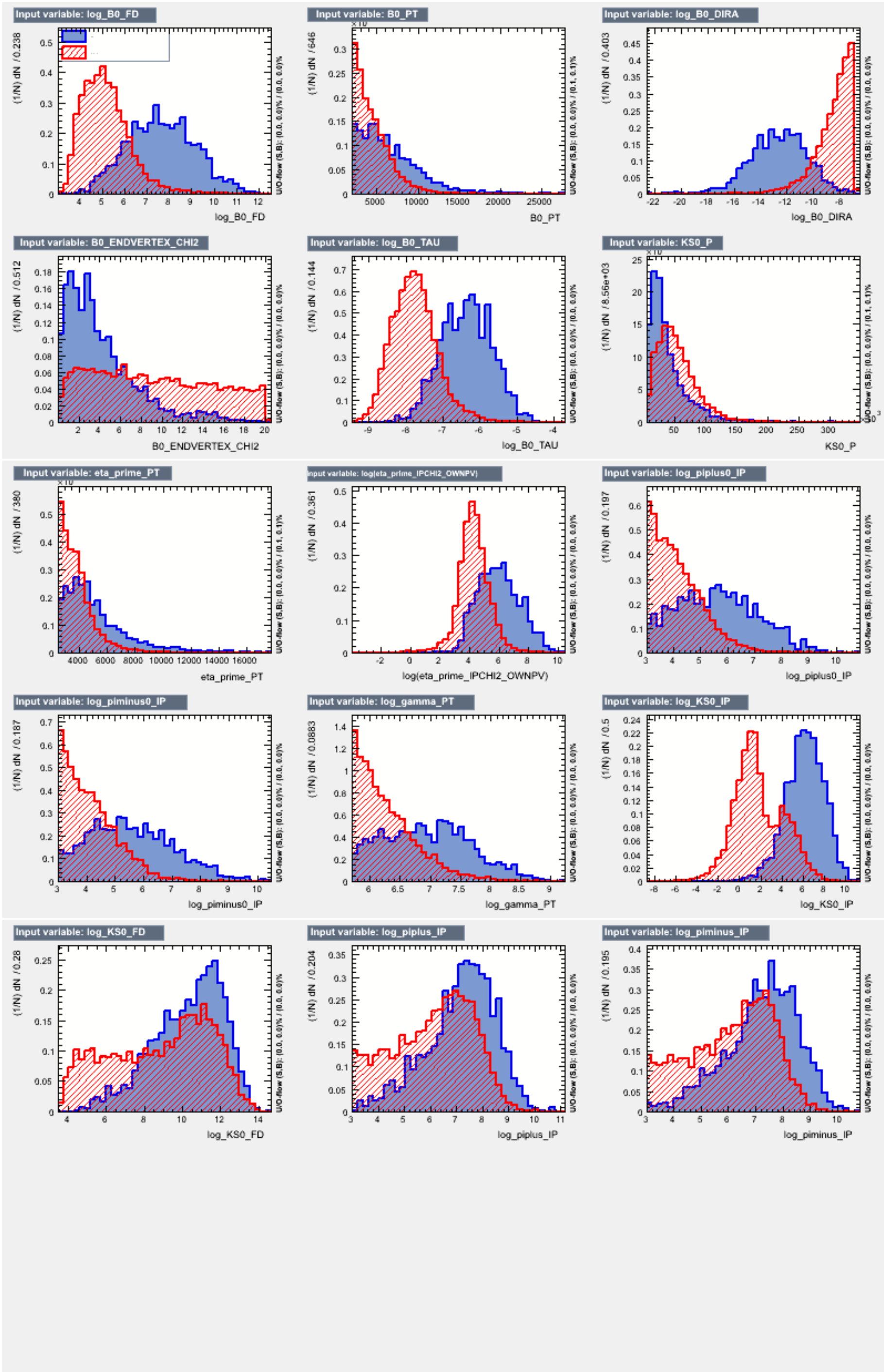


Figure 15: Variables trained for the BDT (LL Selection) showing the distribution for signal and background events for the B^0 selection.

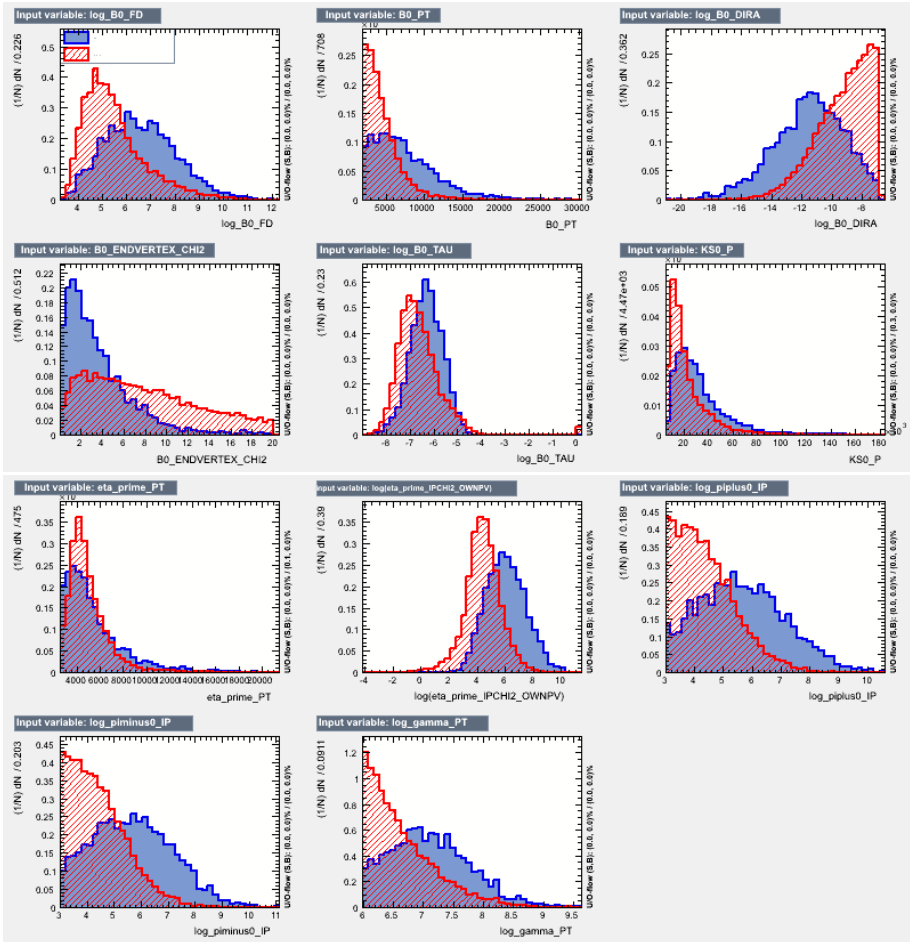


Figure 16: Variables trained for the BDT (DD Selection) showing the distribution for signal and background events for the B^0 selection.

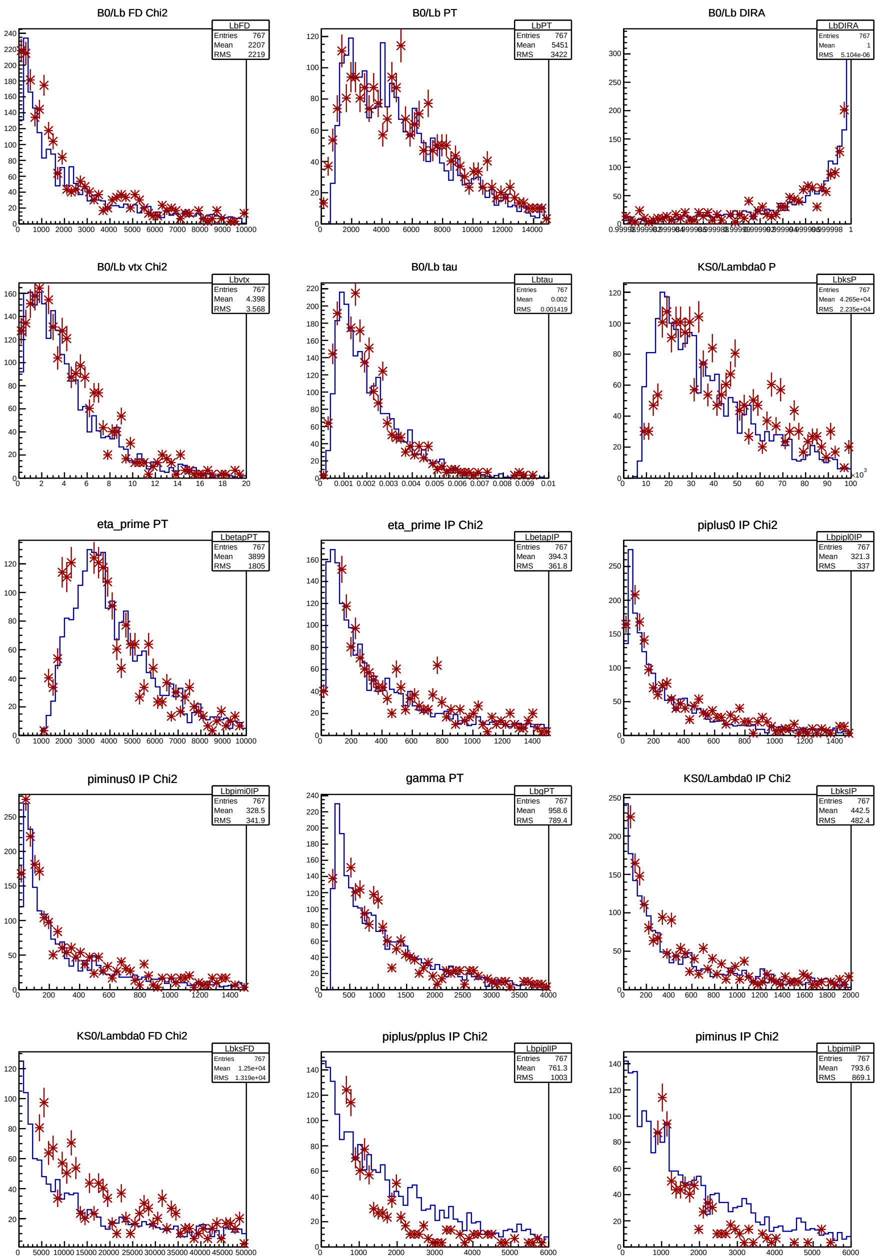


Figure 17: Variables trained for the BDT (LL Selection) showing the distribution for B^0 signal (histogram) and Λ_b^0 (points) signal events overlaid.

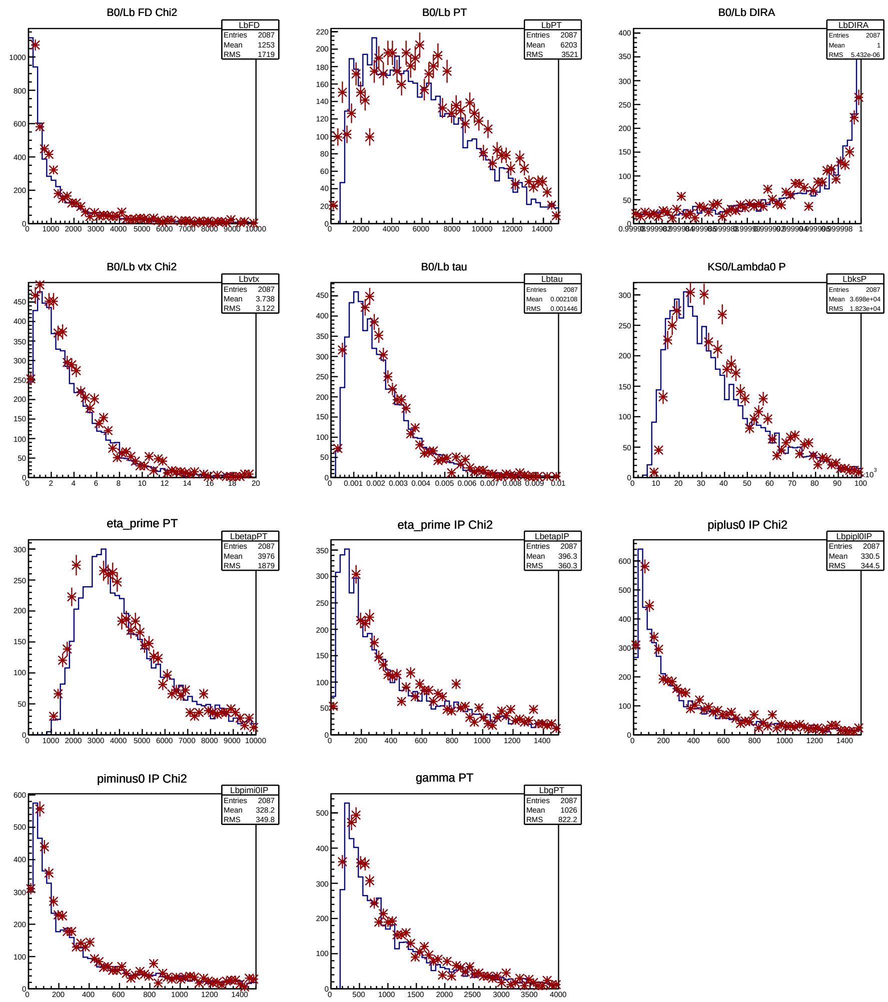


Figure 18: Variables trained for the BDT (DD Selection) showing the distribution for B^0 signal (histogram) and Λ_b^0 (points) signal events overlaid.

Figure 19: Correlations of variables trained for the BDT

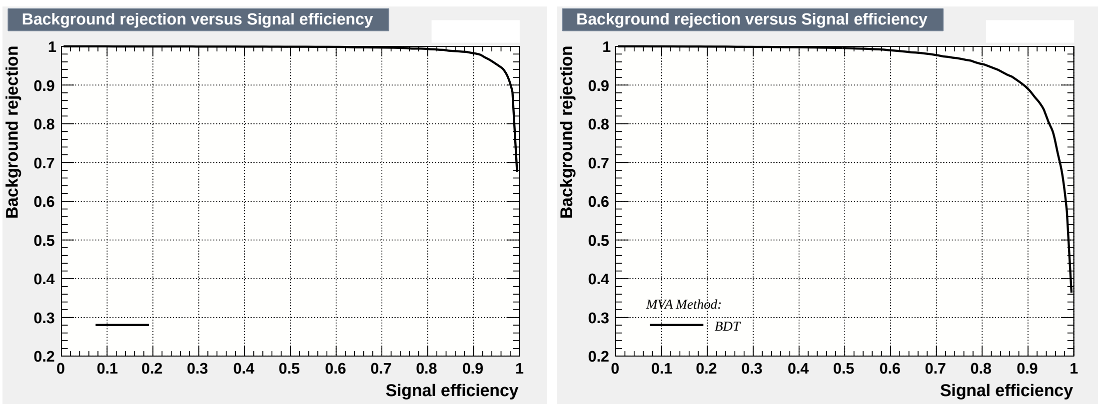


Figure 20: ROC curves for the output from LL(left) and DD(right) selections