Chapter 3

The ${\Lambda_b}^0 \to {\Lambda_c}^{*+}\pi^-\pi^+\pi^-$ decay

As stated before the decay ${\Lambda_b}^0 \to {\Lambda_c}^{*+}\pi^-\pi^+\pi^-$ is a significant source of background in the search for violation of LFU in the R(${\Lambda_c}^{*+}$) measurement. In order to measure the BF of this background we study in this thesis an optimized selection of the ${\Lambda_b}^0 \to {\Lambda_c}^{*+}\pi^-\pi^+\pi^-$ signal and compute its efficiency.

3.1 Topology of signal

The decay that was used in this thesis is:

$${\Lambda_b}^0 \to {\Lambda_c}^{*+}\pi^-\pi^+\pi^-$$

where

- ${\Lambda_c}^{*+} = {\Lambda_c}^{+}(2625)$.
- ${\Lambda_c}^{*+} \to {\Lambda_c}^{+}\pi^-\pi^+$.
- ${\Lambda_c}^{+} \to pK^-\pi^+$.

In principle two ${\Lambda_c}^{*+}$ resonances can be used, as we can see form the distribution of mas difference between the ${\Lambda_c}^{*+}$ and the ${\Lambda_c}^{+}$ baryons in Figure 4.1b. In this study we consider only the ${\Lambda_c}^{+}(2625)$.
In order to separate the signal decay from combinatoric background we exploit the excellent capabilities of the LHCb detector concerning momentum, impact parameter resolution and particles identification.
This decay channel presents two important feature that cannot be used in semi-leptonic decays due to the presence of neutrinos: first it is kinematically closed i.e. all the particles in the final state can be reconstructed, second the reconstructed ${\Lambda_b}^0$ momentum points back to the Primary Vertex (PV).
Another advantage of this channel is that, thanks to resonance ${\Lambda_c}^{*+}$ which decays hadronically, the ${\Lambda_b}^0$ vertex can be found with a very good quality. Finally the reconstruction of the ${\Lambda_c}^{*+} \to {\Lambda_c}^{+}\pi^-\pi^+$ decay is almost free from down-feed contamination.
The ${\Lambda_b}^0$ baryons lifetime is large enough to allow them to fly on average about 1 cm before decaying.
Also ${\Lambda_c}^{+}$ decay displaced with respect to ${\Lambda_b}^0$ vertex, unlike ${\Lambda_c}^{+}(2625)$ resonance. These peculiarities help us to identify the searched channel: large impact parameters (IP) of the reconstructed tracks and displaced vertices with good qualities.
Finally, b hadrons are identified using LHCb PID system, section 2.2.3.

3.1.1 Relevant quantities

The data used in this work have been collected with the LHCb detector in pp collisions after various processing phases. The online event selection is performed by a trigger system, which consists of a hardware stage, based on information from the detectors, followed by a software stage, which applies a full event reconstruction. After the software stage and the offline event reconstruction data are finally obtained with a selection of the interesting quantities and are available as root files; this work used roughly 450.000 events with about 60 variables each identifying all the main characteristics of the events.
It is useful to define some relevant quantities and variables used in the analysis. All quantities are calculated in the laboratory frame, and in general, can be applied to each particle in the decay.

- Primary vertex (PV): the point of the space where the primary proton-proton interaction is reconstructed.
- Secondary decay vertex (SV): the point in the space where the decay of long-lived particles occurs. The displacement vector $\vec{d}$, defined as the spatial vector from primary to second vertex, is equal to $\vec{d} = \gamma\vec{\beta}ct = (\vec{p}/m)ct$, where c is the speed of light, m the mass, $\vec{p}$ is momentum and t the proper time decay of the decay particle.
- Impact Parameter (IP): given a vertex $\vec{v}$, the impact parameter is the vector formed by the nearest point of a track to $\vec{v}$.
- DIRA: the cosine of angle between the momentum of the particle $\vec{p}$ and the displacement vector, $DIRA = \vec{p} \cdot \vec{d}/|\vec{p}||\vec{d}|$ For a fully reconstructed particle its total momentum tends to be aligned to the displacement vector resulting into a DIRA close to unit (like in our case), differently for one partially reconstructed, like semileptonic decay.
- χ^2_{vertex} and χ^2_{track}: the tracks and the vertexes are recostructed by minimizing the χ^2 of the hints in the detectors. Small χ^2_{vertex} and χ^2_{track} ensures an agreement between the track and the vertex model and the reality.

The large amount of variables that define this decay process make it difficult to separate Background from Signal because both populate complicated regions in a large multidimensional space. The AI approach is mandatory to find where the main part of the Signal lies. To obtain these results an AI technique is trained with a mixture of real data and Monte-Carlo simulations and, once fully trained, the algorithm can be applied to additional data to enhance the Signal component.
To form the training sample in this work the Signal is taken from a Monte-Carlo simulation while the Background is taken from real data.

Chapter 4

The data sample analisys

This chapter describes the first step of the analysis of the data sample, in particular the identification of the variables that discriminate between signal and background.

4.1 Data sample

We used the data collected by the LHCb detector in pp collisions at 14TeV of energy in the center of mass. Generic hadronic events were selected at the trigger level by using impact parameter selection and the topology of the events.
After the trigger level selection the complete event reconstruction is performed and a second loose selection is applied to the events. This selection, called "stripping", requires the events to contain good tracks consistent with the decay of ${\Lambda_c}^+ \to pK^-\pi^+$ and 3 additional pions forming a good vertex. Finally requires that the reconstructed ${\Lambda_b}^0$ momentum points to the primary vertex associated to the ${\Lambda_b}^0$. After the stripping 7 million events are selected. In order to further reduce the number of events and for practical purposes additional quality cuts are applied to the sample Table 4.1 reducing the number of selected events to 450000.

$2260 < m({\Lambda_c}^+) < 2310 MeV$
Probability of the Λ_c^{*+} π^- identified as a π>0.4
Probability of the Λ_c^{*+} π^+ identified as a π>0.4
Probability of the other 3 π identified as a π>0.4
${\Lambda_b}^0$ DIRA PV>0.99999
${\Lambda_b}^0$ χ^2_{vertex}<15
${\Lambda_c}^+$ χ^2_{vertex}<10
${\Lambda_c}^{*+}$ χ^2_{vertex}<10

Table 4.1: Pre-selection cuts

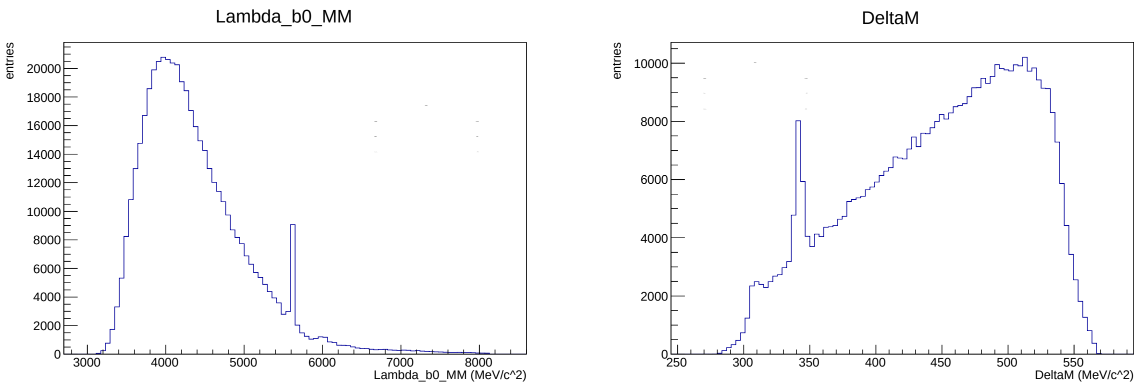
4.2 Selection of signal and background sample

In order to obtain the distribution of the discriminating variables for the signal we use a sample about 6000 signal events simulated with GEANT4 [4] and reconstructed with the same algorithms and pre-selection used to reconstruct real data. For simulated events we also require that reconstructed candidates are correctly matched to signal decays.
In order to obtain the same distributions for background events we use our data sample and apply a selection to exclude signal events.
The first cut is selected on the mass of the ${\Lambda_b}^0$ which we know the exact value $m({\Lambda_b}^0) = 5619.51 \pm$

$0.23MeV$. In Figure 4.1a is represented the distrubution of this variable and it is easily to see that there is an important background signal. We decided to select a background cut on the $m({\Lambda_b}^0)$ equal to $m({\Lambda_b}^0) < 5550MeV$ or $m({\Lambda_b}^0) > 5680MeV$.
The second cut is selected on the distribution of the $\Delta m = m({\Lambda_c}^{*+}) - m(\Lambda_c^+)$. The exact values for this two variables are $m({\Lambda_c}^+) = 2286.46 \pm 0.14MeV$ and $m({\Lambda_c}^+(2625)) = 2628.11 \pm 0.19Mev$. Also in this case (Figure 4.1b) the plot of the Δm distribution has an important background signal. As we can see there is also a peak near 306 MeV which correspond to the Δm with the $m({\Lambda_c}^+(2595))$. So we set another background cut, this time on the Δm, equal to $\Delta m > 360MeV$. Doing this we select the background sample: all the data with this characteristics are cataloged as background. In the Table 4.2 the cuts that define the background sample are summarized.

$m({\Lambda_b}^0) < 5550MeV$ or $m({\Lambda_b}^0) > 5680MeV$
$\Delta m > 360MeV$

Table 4.2: Background selection cuts



(a) Plot of the mass of the ${\Lambda_b}^0$ for the data sample

(b) Plot of the Δm

Figure 4.1: Plot of the variable used for the background cut

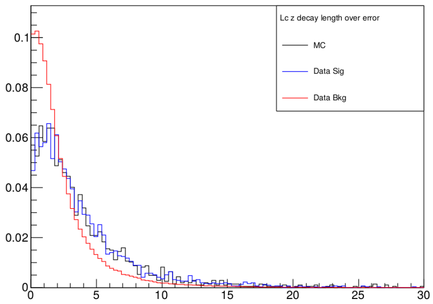
4.3 Comparison between variable for signal and background samples

The next step is to find some variables that discriminate the signal from the background. This plots on Figure 4.2 represent the normalized distributions of the most relevant variables for the signal and background of data sample and for montecarlo, in the region of interest. The signal of data sample is defined using the cuts in the Table 4.3. Since we want to get a signal enriched sample, we apply on our data sample the inverse cuts used for the selection which exclude signal events.

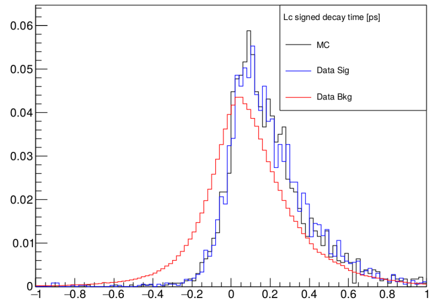
$5550 < m({\Lambda_b}^0) < 5680MeV$
$\Delta m < 360MeV$

Table 4.3: Signal selection cuts

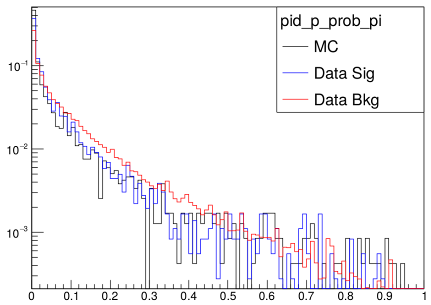
We select discriminating variables by requiring that their distribution on MC and the corresponding distribution on data is different. In addition we check that the distribution on MC is consistent with the signal enhanced data sample.



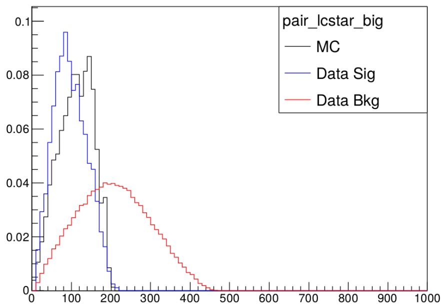
(a) Λ_c^+ z decay lenght over error



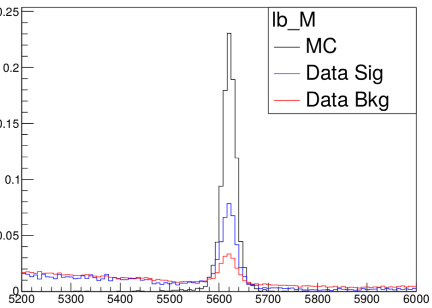
(b) Λ_c^+ decay time [ps]



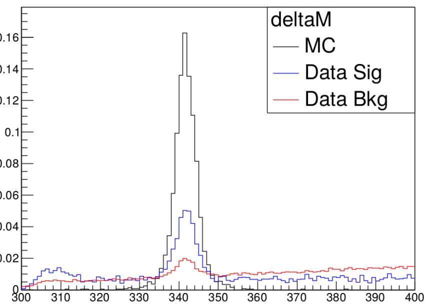
(c) Probability that a p is identified as a π



(d) Pair mass of the two π of the ${\Lambda_c}^{*+}$



(e) $m({\Lambda_b}^0)$ [MeV]



(f) Δm [MeV]

Figure 4.2: Comparison of some variables Plots

For example as we see from the plot 4.2c the variable plotted is not a good discriminant variable. Instead the variable plotted in 4.2d is a very good discriminant variable.

The final list of the discriminant variables is reported in the Table 4.4

Variable
$m(\Lambda_c^+)$
Probability of a Λ_c^+ p identified as a p
Λ_c^+ z decay lenght over error
Λ_c^{+*} z decay lenght over error
Λ_c^+ decay time
Λ_c^{+*} decay time
Λ_b^0 decay time
Probability of Λ_c^+ χ^2_{vertex}
Probability of Λ_c^{+*} χ^2_{vertex}
Λ_b^0 DIRA of PV
Pair mass of π_0 and π_1
Pair mass of π_1 and π_2

Table 4.4: Discriminant variables

The probability of a Λ_c^+ proton is identified as a proton is calculated by Neural Net from likelihood ratios from the data of RICH detectors. The MC and signal distributions have a peak near 1 greater than the background one, as we can see in the Figure 6.1(where I report all the variables used for training the BDT).
Λ_c^+ z decay length over error, and also Λ_c^{+*} with the appropriate substitutions, is equal to $\frac{d_z}{\sqrt{\sigma_{\Lambda_b^0 ENDVERTEX}{}^2 + \sigma_{\Lambda_c^+ ENDVERTEX}{}^2}}$ where $d_z = -z_{\Lambda_b^0 ENDVERTEX} + z_{\Lambda_c^+ ENDVERTEX}$. The background distribution of this variable stays closer to zero respect MC and signal distributions.
Λ_c^+ decay time is equal to $\frac{d \cdot m(\Lambda_c^+)}{c \cdot p(\Lambda_c^+)}$ where $d = \sqrt{d_x^2 + d_y^2 + d_z^2}$, c is the speed of light, m is the mass and p is the momentum. Also in this variable distribution the background stays closer to zero respect MC and signal, moreover the background distribution is symmetric to 0 while the MC and the signal ones are mostly displaced in the positive region.
The probabilities of χ^2 for the vertexes of the Λ_c^+ and Λ_c^{+*} are others important discriminant variables. The background distribution has an exponential increase near 0 unlike MC and signal distribution, which have a constant trend as we expect.
As it was said before, the Λ_b^0 DIRA of PV has to be near the unit for the signal, and also for MC.