Master Thesis

A reconstruction and selection efficiency study at LHCb

Author
Raphael Märki - LPHE - EPFL

Supervisors
Prof. Olivier Schneider - LPHE - EPFL
Dr. Raluca Mureșan - LPHE - EPFL

January 14, 2010
1 Abstract

This report gives some insight into the methodology of the planned $\Lambda$ measurements at LHCb. It helps to understand the effect introduced by the LHCb detector structure, its geometrical acceptance, its reconstruction capabilities and the effect of chosen selection cuts on some of the distributions that are planned to be measured. The overall reconstruction efficiency based on Monte Carlo simulation for each bin in pseudorapidity and transverse momentum as well as the selection efficiency is also studied. Finally, in order to increase the quality of the sample, the selection cuts are tuned in a way that maximizes the signal significance.

2 Introduction

2.1 The Large Hadron Collider

The most powerful particle accelerator in the world is the Large Hadron Collider at CERN in Geneva. It is designed to collide two beams of protons at a nominal energy of 7 TeV per beam and generates altogether 600 million collisions every second. Four main experiments are situated along the 27 km of this circular accelerator, each at a point where protons collide.

The nominal luminosity and energy will only be progressively reached. At the launch of the experiment in the end of 2009, beams with an energy up to 1.2 TeV have been collided and the very first data at 450 GeV were recorded. The $V^0$ working group decided to analyze these first data although there are not so many events and not all detectors are perfectly calibrated.

Nevertheless, the first data with enough significant statistics for a $\Lambda$ study will be the 3.5 TeV per beam event planned to be taken in early 2010 before the next shutdown. This is the reason why mainly simulated data samples at precisely this energy are used in this report.

2.2 The LHCb experiment

One of the four main experiments at the Large Hadron Collider is the LHCb experiment [1] [2], a single-arm forward spectrometer. The main purpose of this experiment is to measure CP violation and rare decays in the b quark sector. The b hadrons will be produced by an almost head-on collision between the two proton beams of LHC with an energy of 7 TeV each. The approximately 20 m long LHCb detector is built of several sub-detectors which are capable to track and identify particles within an acceptance from 10 to 300 mrad horizontally and 250 mrad vertically. A complete picture of the LHCb detector is shown on Figure 2.1.
As said in the previous paragraph, the proton beams are not colliding exactly frontally at LHCb but there is a small crossing angle in the horizontal plane. This crossing angle is defined as being half of the acute angle between the two proton beams. It has two components, the internal and the external one. The internal crossing angle is there to compensate for the magnetic field of the dipole magnet [3]. The external crossing angle helps avoiding unwanted bunch collisions or any interaction between two bunches that are not meant to collide [4]. The value of the total crossing angle is chosen as a function of the intensity and polarity of the magnetic field, the beam energy and the number of bunches. This relation between the angle and the energy for the planned runs is shown in Table 2.1.

<table>
<thead>
<tr>
<th>Beam energy of the run</th>
<th>Crossing angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 TeV</td>
<td>285 mrad</td>
</tr>
<tr>
<td>5 TeV</td>
<td>340 mrad</td>
</tr>
<tr>
<td>3.5 TeV</td>
<td>270 mrad</td>
</tr>
</tbody>
</table>

Table 2.1: Total crossing angle for different energies and different running conditions [5].

The $\Lambda$ reconstruction will be based only on the tracking information, therefore only the tracking subdetectors are briefly described and their working principle explained.

- The VErtext LOcator (VELO) is the closest subdetector to the proton-proton collision region. It is composed of 42 semi-circular Si detector modules that can be moved as close as 8 mm from the beam. Each VELO station is able to make an azimuthal ($\phi$) and a radial ($r$) measurement.
of a particle’s trajectory. The VELO subdetector can also measure the position of the primary vertex with a resolution of 60 \( \mu \text{m} \) along the beam axis and 10 \( \mu \text{m} \) in the transverse plane.

- The Tracker Turicensis (TT) is a silicon microstrip detector located between RICH1 and the LHCb dipole magnet. It consists of four layers that cover a total surface of around 7 m^2.

- The Inner Tracker (IT) is also a silicon detector forming the inner part of three tracking stations (T1, T2 and T3) between the LHCb dipole magnet and the RICH2 subdetector. Its three stations of each four layers cover the region close to the beam pipe.

- The Outer Tracker (OT) surrounds the Inner Tracker and forms the outer part of the three tracking stations (T1, T2 and T3). This subdetector is composed of four layers of straw tube drift chambers at each tracking station. The external straw tube drift chambers are oriented vertically whereas the two inner layer are oriented at \( \pm 5^\circ \) respectively in the vertical plane. This gives a good three-dimensional track resolution and still keeps the ghost rate low.

Each subdetector can give hits whenever a particle is in its acceptance. In this report “in acceptance” means that a particle passes through the physical volume the detector occupies. Depending on the number of observed hits in different tracking station one defines the track types as shown in Figure 2.2.

![Figure 2.2: Track types.](image)

To be more precise, the track types are composed of track segments (for each tracking station) which have a clear definition [6]. The requirements for each track segment are listed hereafter:

- **VELO track segment** : at least three \( r \) and three \( \phi \) hits.
- **TT track segment** : at least one hit in the first four planes (TT1) and one hit in the last four planes (TT2).
- **T1, T2 and T3 track segment** : at least one hit in two planes of each tracking station.
Then, the exact definition of each track type is the following:

- **Velo track**: 3D tracks having solely a Vertex Locator (VELO) track segment.
- **T track**: tracks with solely a T1, T2 and T3 track segment.
- **Long track**: tracks traversing the detector, from the VELO to the T stations. They can have track segments of all tracking stations but only VELO and T track segments are required.
- **Upstream track**: tracks with VELO and Trigger Tracker (TT) track segments.
- **Downstream track**: tracks with TT and T1, T2 and T3 track segments.

In the following, a particle is defined as “reconstructed” whenever it has a long track, or for non final state particles (e.g. $V^0$) whenever all charged daughter particles have long tracks.

### 2.3 $V^0$ physics

By $V^0$ one understands a particle that is either a $\Lambda$, a $\bar{\Lambda}$ or a $K^0_S$, neutral strange hadrons decaying weakly into two charged daughter particles having a mean free path in their restframe ($c\tau$) of a few centimeters. These particles have a large production cross section in the LHC energy range, so their study doesn’t require a large integrated luminosity compared to $b$ physics analysis. Hence the publication about $V^0$ production is expected to be one of the first among the LHCb physics papers.

The samples used for this study are the so called minimum bias event samples. These events are recorded with a minimum trigger giving the least biased data one is capable to obtain. Since $V^0$s have a high production rate, no further trigger is needed to perform the analysis.

### 2.4 $V^0$ ratios and motivation

A very first measurement that can be done quite early, and even without knowing the luminosity, are the ratios between different $V^0$ types in bins of any kinematic or geometric variable. These ratios, for instance $\bar{\Lambda}/\Lambda$, are very useful for understanding the hadronization processes and tuning Monte Carlo generators. Indeed, the present tunings have been made with data produced at lower energy and then extrapolated to predict the distributions in the LHC range.

Especially the pseudorapidity\(^1\) dependence of the $\bar{\Lambda}/\Lambda$ fraction is expected to be very useful to determine the best simulation parameters and hence the best corresponding physical model. This comes from the combination of the extrapolation in energy with the extrapolation in pseudorapidity. Therefore the most significant differences between several Monte Carlo tunings are their predictions at high energies and high pseudorapidities.

Since LHCb is a one-arm forward spectrometer having an acceptance very close to the beam, it is able to reconstruct particles with a pseudorapidity up to 5. Hence, this experiment is the most appropriate among the LHC detectors for putting into evidence the differences between different Monte Carlo predictions of these $V^0$ ratios at high pseudorapidities. Indeed, some pseudorapidity distributions of the $\bar{\Lambda}/\Lambda$ ratio in the energy and pseudorapidity range of LHCb differ up to 5% for different Monte Carlo tunings. This difference is only 1% in the pseudorapidity range of the other central detectors at LHC.

Some pseudorapidity distributions for various Monte Carlo generator tunings are illustrated on Figure 2.3 at an energy of 5 TeV per beam. The pseudorapidity range where particles can be

---

\(^1\)The pseudorapidity is defined as $\eta = \frac{1}{2} \ln \left( \frac{p + p_L}{p - p_L} \right) = -\frac{1}{2} \ln \left( \tan \left( \frac{\theta}{2} \right) \right)$ where $p$ is the momentum, $p_L$ the longitudinal component of the momentum (along the $z$ axis, perpendicular to the detection planes) and $\theta$ the angle between the particle’s trajectory and the $z$ axis.
reconstructed at LHCb is between 2.5 to 5. One immediately sees how useful studies of data at such high pseudorapidities are for choosing between these tunings.

![Figure 2.3: \( \Lambda / \Lambda \) production ratio as a function of pseudorapidity for different LHCb Monte Carlo tunings [7] (generator level only).](image)

Beyond \( \Lambda \) and \( K^0_S \) this report also provides useful results for other particle measurements, for instance a \( \Lambda_b \) production study. In this analysis, \( V^0_S \)s decaying into two charged daughter particles are reconstructed but most of the presented results can also be used for the study of heavier hadrons that decay into more particles. As an example, the decay \( \Lambda_b^0 \rightarrow J/\Psi(1S)\Lambda(p\pi^-) \) needs even a \( \Lambda \) to be reconstructed. This illustrates the interest of this report for b physics, which really is the main purpose of the LHCb experiment.

### 2.5 How to reconstruct a \( V^0 \)

As said before, a \( V^0 \) has at least one decay mode containing only massive charged daughter particles. These particles leave tracks in the detector which allow to measure their momentum. The momentum measurement of the daughter particles is enough to reconstruct a \( V^0 \) using the following method.

First one duplicates all tracks for each expected daughter particle and associates to each of these tracks the mass of this daughter particle. Then one loops over all combination of two oppositely charged tracks and adds their previously calculated four-vectors. Whenever this combination gives a four-vector with a mass inside a defined window around the \( V^0 \) mass, the candidate is kept.

To be definitely kept, the \( V^0 \) candidate needs also to pass the loose selection cuts described in a later section. These cuts are based on other kinematic variables as well as on the \( \chi^2 \) of the reconstructed \( V^0 \) decay vertex.
3 Origin and decay of $Λ$s hyperons

The purpose of this study is to prepare $Λ$ production measurements, therefore the entire report starting from here refers now to $Λ$s and no other $V^0$s anymore.

3.1 Monte Carlo generation

For the analysis performed hereafter, only Monte Carlo generated events are used since no interesting data have been produced early enough to fit into the timeframe of this work. The Monte Carlo generation (MC09) was done with Gauss (v37r5 and higher), the LHCb Gaudi application based on PYTHIA and EvtGen tuned for LHCb. Proton-proton interactions at 3.5 TeV per beam colliding with a crossing angle of 270 mrad were simulated since such conditions are expected for the data in 2010.

3.2 Lambda mothers

A $Λ$ hyperon can be produced

1. at the primary vertex (PV) in a non-diffractive event;
2. at the primary vertex in a diffractive event (eg. $AB \rightarrow AX$);
3. in electromagnetic and strong decays of resonances;
4. in weak decays;
5. in interactions with detector matter.

Since the production of $Λ$s is studied, one ideally wants only prompt $Λ$s which are defined as those produced directly at the PV, either directly or in electromagnetic or strong decays of resonances. These prompt $Λ$s really describe the $Λ$ production processes in pp interactions, the final goal of this research.

As seen above, $Λ$ particles are not always directly produced in the primary collision of the two proton beams but also in secondary interaction with detector matter or in the decay of other particles coming themselves from the primary vertex or again from other particle decays. At LHCb, $Λ$s come mainly from the PV or the 15 mother particles listed in Table 3.1. The number of $Λ$s produced by diffractive interactions is very low as described later on and very few $Λ$s come from interaction with matter in the PV region since there is no matter present except residual gas.
Mother particle | \( c\tau \)
--- | ---
PV | 0
Short living hyperons | |
\( \Sigma^+ \) | < 10\(^{-14}\) m
\( \Sigma^0 \) | 2.22 \( \times \) 10\(^{-11}\) m
Long living hyperons | |
\( \Xi^0 \) | 8.71 cm
\( \Xi^- \) | 4.91 cm
\( \Omega^- \) | 2.461 cm

Table 3.1: \( \Lambda \) mother particles and their mean lifetime. The longest living strongly or electromagnetically decaying particle has a \( c\tau \) of 2.22 \( \times \) 10\(^{-11}\) m.

The longitudinal resolution on the primary vertex is of the order of 50-150 \( \mu m \), therefore one would like to know how many \( \Lambda \)s satisfy the definition of prompt within this region. Taking three times the resolution one shall have a good estimate of this contamination. Table 3.2 shows the percentage of mother particles for all of those seen in Table 3.1.

| Mother | \(< 150 \mu m\) | \(< 450 \mu m\) |
--- | --- | ---
PV | 46.84 % | 46.76 %
Short living baryons | |
\( \Sigma^+ \) | 31.48 % | 31.47 %
\( \Sigma^0 \) | 21.59 % | 21.61 %
Long living baryons | |
\( \Xi^0 \) | 0.0050 % | 0.0125 %
\( \Xi^- \) | 0.0105 % | 0.0293 %
\( \Omega^- \) | 0 % | 0.00067 %
Other particles | |
Mesons* | 0.0064 % | 0.0091 %

Table 3.2: Composition of \( \Lambda \) mothers for 3.5 TeV beam interactions. The 150 \( \mu m \) and 450 \( \mu m \) are the distance between PV and the \( \Lambda \) origin vertex.

Within 150 \( \mu m \), 99.9% of the \( \Lambda \)s are coming from the PV or fast decaying mother particles whereas for 450 \( \mu m \), 99.8% can be considered as prompt. It appeared also that no \( \Lambda \)s are produced by interaction with matter within 450 \( \mu m \) around the PV. This is quite intuitive since there is no detector matter in this region and beam-gas interactions are not included in our simulation.

Nevertheless, some short lived mothers might not come directly from PV but be the result of a non prompt production. To check if such a scenario is taking place for short lived mother particles and in what proportions, a full analysis over all decay trees containing a \( \Lambda \) was done and the result is presented on Figure 3.1.
Now one sees that 0.47 % of the short lived mothers of \( \Lambda \)s produced in the first 450 \( \mu \text{m} \) have an ancestor with a \( c\tau \) larger than \( 10^{-10} \text{ m} \). This fraction is 0.30 % for \( \Lambda \)s created less than 150 \( \mu \text{m} \) away from the PV. Therefore some short lived mother particles have themselves long lived mother particles. Thus the contamination of long lived mother particles is higher than initially expected.

Looking a bit closer to the ancestors of short lived mothers one sees that these \( \Sigma \)s are mainly coming from the PV or fast decays but they can also have longer living ancestors. One observes on Table 3.3 the case of \( \Lambda \)s produced less than 450 \( \mu \text{m} \) away from the PV.

<table>
<thead>
<tr>
<th>Ancestor particle</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Xi^+ )</td>
<td>0.0044 %</td>
</tr>
<tr>
<td>( \Xi^- )</td>
<td>0.027 %</td>
</tr>
<tr>
<td>( \Lambda^+ )</td>
<td>0.412 %</td>
</tr>
<tr>
<td>( \Lambda^0 )</td>
<td>0.0133 %</td>
</tr>
<tr>
<td>( \Xi^0 )</td>
<td>0.0051 %</td>
</tr>
<tr>
<td>( B^0 )</td>
<td>0.0006 %</td>
</tr>
<tr>
<td>( B^\pm )</td>
<td>0.0032 %</td>
</tr>
<tr>
<td>Total</td>
<td>0.471 %</td>
</tr>
</tbody>
</table>

Table 3.3: Composition of the longest living ancestor for \( \Sigma \)s decaying into a \( \Lambda \) within 450 \( \mu \text{m} \) around the PV.
Hence an additional 0.250 %\(^1\) of contamination has to be considered for 450 μm and 0.159 % for 150 μm lowering the percentage of “prompt” \(\Lambda\)s to 99.6% and 99.75% for 450 μm and 150 μm respectively. Those contaminations are even higher than the direct ones, where the \(\Lambda\) is coming directly from the long lived mother particle. Knowing this and assuming that our simulation is well tuned, one can correct the measurements for these effects as it was done in STAR [8] or CDF [9].

3.3 Lambda origin vertex distributions

To see where the \(\Lambda\)s are produced and to confirm if one can separate between prompt and non-prompt \(\Lambda\)s using a distance cut, the distribution of the distance between the primary vertex and the \(\Lambda\) origin vertex is plotted in Figure 3.2.

![Origin vertex](image)

Figure 3.2: Distance between PV and origin vertex of a \(\Lambda\) where the mother particle is a long living strange baryon (red) or a long living charmed or bottom baryon (black).

The direct contribution of long living strange baryons gives a constant low contamination within 1 mm around the PV. On the other hand, the mean lifetime of the long living charmed or bottom baryons corresponds to a \(c\tau\) close to the resolution, giving a exponentially decreasing contamination in that region.

Therefore being able to cut on the distance between PV and \(\Lambda\) origin vertex, would allow to remove non-prompt noise. Unfortunately it is rather difficult to determine this distance but it is easy to measure the impact parameter (IP), defined as the shortest distance between the PV and the reconstructed \(\Lambda\) track. Hence an IP cut will be used for the reconstructed \(\Lambda\)s and its efficiency and purity studied in a further section.

\(^{1}53.08\%\) of \(\Lambda\) mothers that are \(\Sigma\)s multiplied by 0.471 % of non prompt \(\Sigma\) ancestors.
3.4 Initial distributions for prompt lambdas

Before moving forward, one shall have a first quick look at how some distributions look like for produced prompt $\Lambda$s. The pseudorapidity and transverse momentum distributions are drawn in Figure 3.3 in a two-dimensional histogram. One sees that the shape is rather smooth which corresponds to a physical distribution of all prompt $\Lambda$s, i.e. not detector dependent nor modified by interaction with matter.

![Figure 3.3: Pseudorapidity and $p_t$ distribution for prompt $\Lambda$s.](image)

3.5 Lambda decays and absorption by matter

The $\Lambda$s can disappear in two different ways:

1. Decay
2. Interaction with matter

Nevertheless, most of the prompt $\Lambda$s (97 %) are actually decaying since they fly only a few cm and there is not much matter around the interaction point. Especially inside the VELO, $\Lambda$s have few chances to be absorbed by matter and the $\Lambda$s flying out of the VELO are not reconstructed anyway since both daughter particles are required to have long tracks which means that they have to produce hits in the VELO.

4 Reconstruction efficiency

Since there is much interest in measuring ratios like $\bar{\Lambda}/\Lambda$ or $K^0_S/(\Lambda+\bar{\Lambda})$ one needs a good understanding of the Monte Carlo based efficiency estimation. For $V^0$s, this efficiency depends on the interaction between the $V^0$ particles and matter, the matching of the daughter particles with the LHCb detector acceptance, the reconstruction capabilities and our selection.

To get this good understanding of the relation between measured $\Lambda$ distributions and the initial ones right after the pp interaction, some interesting variables (pseudorapidity, transverse momentum and the $Z$ coordinate of the $\Lambda$ end vertex) have been looked at. One wants to know the change
of these distributions at each step between the measured one and the initial one, i.e. the distributions of all prompt $\Lambda$s right after the collision.

The steps mentioned above are chosen to be the conditions that the $\Lambda$s decay within the simulated volume, that all daughter particles fly through the detector acceptance and finally that all daughter particles are reconstructed as long tracks. If one applies this last condition (that both daughter particles are reconstructed) to the initial distribution of prompt $\Lambda$s, one obtains the measured distribution of prompt $\Lambda$s without the effect of the selection. The efficiency of the selection cuts and noise are only studied in section 5.

4.1 Pseudorapidity distributions

Starting with the pseudorapidity in the center of mass distribution of all prompt $\Lambda$s and $\bar{\Lambda}$s on Figure 4.1 one sees a very particular structure. Indeed there is a big depletion in the zero region which is typical for pseudorapidity distributions even for uniformly distributed momentas. There are also two peaks at around $\pm 8.25$ for $\Lambda$s but not for $\bar{\Lambda}$s. These peaks come from the leading particle effect hence the pseudorapidities they are situated at correspond to the ones of the beam protons.

This Figure 4.1 contains also the distribution of the reconstructed $\Lambda$s (as defined in the first section) showing that only $\Lambda$s with a pseudorapidity between 2 and 6 can be reconstructed at LHCb using long tracks. Therefore this range of pseudorapidity is chosen for the following studies. On Figure 4.2 is shown a zoom on that particular pseudorapidity range and one sees a rather smooth structure of the distribution.

Figure 4.1: In black, pseudorapidity in the center of mass distribution for all prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right). In red, pseudorapidity in the center of mass distribution for prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right) decaying into $\pi$ and $p$ both reconstructed as long tracks.

Figure 4.2: Pseudorapidity in the center of mass distribution between 2 and 6 for all prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right).
4.1.1 Detector geometry effects on pseudorapidity distributions

The next step is to look at the pseudorapidity distribution of all prompt $\Lambda$s having a decay as end vertex. This is shown in Figure 4.3 where one observes a significant dip at a pseudorapidity of 4.35 and another one at 5.25. These pseudorapidities correspond to an angle between the $Z$ axis and the momentum of the $\Lambda$ of about 25.8 mrad and 10.5 mrad respectively.

![Figure 4.3: Pseudorapidity in the center of mass distribution for prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right) having a decay vertex as end vertex. One notices some structure in these pseudorapidity distributions.](image)

As one sees on Figure 4.4, the lack of $\Lambda$s at pseudorapidities of 4.35 and 5.25 comes from the interaction with matter. The minimal angular acceptance of the LHCb experiment is 10 mrad (around 5.3 of pseudorapidity) in the LHCb restframe. Nevertheless, the change in pseudorapidity between the restframe and the center of mass frame of the beams is small enough not to change these values significantly. Therefore the particles at a pseudorapidity of 5.25 in the center of mass are at the limit of the acceptance. They are actually passing into the so called beam pipe, the pipe in which the proton beams are circulating and which brings them to the interaction point. But what about the peak at 4.35?

![Figure 4.4: Pseudorapidity in the center of mass distribution for prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right) being absorbed by matter.](image)

For this peak at 4.35, the very first check was to look at the three-dimensional coordinates of those interaction vertices to see where exactly this matter interaction happens. It appears that these vertices are situated a few cm away from the beam trajectory in the RICH1 region, between the VELO and the TT. Figure 4.5 shows the concerned region and one sees that the beam pipe has there a conical shape with a quite large opening angle (25 mrad).
The lack of $\Lambda$s at a pseudorapidity of 4.35 comes from interaction with the matter of the beam pipe and can be explained by the geometry of the beam pipe in the RICH1 region.

The different scenarios are schematically illustrated on Figure 4.6. One sees that a certain fraction of particles are absorbed at a pseudorapidity of 4.35 by the beam pipe due to its shape.

Having explained the structure of the pseudorapidity distribution one can now look at the $\Lambda$s where both daughter particles are in the detector acceptance. This distribution is shown on Figure 4.7 where one sees a rather smooth shape except for the two regions discussed above.
Since no $\Lambda$s decaying after the VELO can have both daughters reconstructed as long tracks, no $\Lambda$s decaying in the RICH1 region will be reconstructed. Nevertheless, the daughter particles are likely to be absorbed in this region by the beam pipe. Therefore the dip at a pseudorapidity of 4.35 will be smeared but still appear on the pseudorapidity distributions of reconstructed $\Lambda$s. Indeed, one observes on Figure 4.8 that there is some structure at a pseudorapidity of 4.35. The dip at a pseudorapidity of 5.25 is not seen anymore since the detector is not able to measure particles that fly into the beam pipe.

To complete this section, the overall reconstruction efficiency is shown in Figure 4.9. This distribution looks very similar to the one for reconstructed $\Lambda$s since it is the ratio between precisely this distribution and the one of all prompt $\Lambda$s, $\bar{\Lambda}s$ which is rather flat in this region.
4.2 Transverse momentum distributions

For the transverse momentum ($p_t$) study one passes through the same steps as for the pseudorapidity distributions study. The $p_t$ distributions for all these steps are presented using a logarithmic scale on Figure 4.10.

![Figure 4.10: Transverse momentum distribution for:](image)

- □ All prompt $\Lambda$s.
- + All prompt $\Lambda$s decaying into $\pi$ and p.
- ○ All prompt $\Lambda$s decaying into $\pi$ and p both within the detector acceptance.
- △ All prompt $\Lambda$s decaying into $\pi$ and p both reconstructed as long tracks.

The shape of the distribution does not change significantly but at high $p_t$ there are very few events leading to bad statistics. This can be better seen on Figure 4.11 where the overall efficiency is shown as a function of the transverse momentum.

![Figure 4.11: Reconstruction efficiency for prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right) decaying into $\pi$ and p as a function of the transverse momentum.](image)

This efficiency goes down very steeply with low values of $p_t$. This can be explained by the fact that for a given momentum, $\Lambda$s with low $p_t$ have a higher pseudorapidity and might therefore be
out of the acceptance ($\eta > 5.25$). Since the proton of the charged $\Lambda$ decay is always flying in a direction close to the $\Lambda$ trajectory, it is also likely to have a large pseudorapidity and be out of the acceptance. Hence the proton has less chances to be reconstructed and without its proton no $\Lambda$ can be reconstructed.

4.3 Z coordinate of $\Lambda$ end vertex distributions

Another interesting variable one can look at is the Z coordinate of the $\Lambda$ end vertex (of both types, decay and matter interaction). This distribution is decreasing exponentially being modulated by the smeared distribution of the Z coordinate of the primary vertex and being slightly modified by the $\Lambda$ origin vertex distribution. Figure 4.12 shows this distribution for all prompt $\Lambda$s and $\bar{\Lambda}$s.

Figure 4.12: Z coordinate of $\Lambda$ decay vertex distribution for all prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right).

If one selects now only prompt $\Lambda$s decaying into $\pi$ and $p$, one sees on figure 4.13 that the shape of the distribution is not changed. Only the overall number of $\Lambda$s decreases by a factor of about 1.6.

Figure 4.13: Z coordinate of $\Lambda$ decay vertex distribution for prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right) decaying into $\pi$ and $p$.

When requesting now both daughter particles, the pion and the proton, to match the acceptance, one sees that the distribution remains still quite close to its initial shape (Figure 4.14).
4.3.1 Effects of reconstruction capabilities on Z coordinate of Λ end vertex distributions

An interesting effect appears as soon as one requires both daughter particles to be reconstructed as long tracks. When looking to Figure 4.15 one sees how the distribution completely changes in shape.

To understand where this structure comes from one needs to look at the efficiency since it depends less on the studied type of particles and the initial distribution. Figure 4.16 shows that the efficiency increases with the Z coordinate of the Λ end vertex until around 520 mm and then decreases again. Furthermore, one observes two dips at 250 mm and 450 mm that are explained in what follows.
First a reminder of what is exactly required: that both daughter particles are reconstructed as long tracks. This means that both daughter particles need to have three $r$ and three $\phi$ VELO hits and two hits in each of the three tracking stations (T1, T2, T3). Therefore both particles have to cross at least three VELO stations.

So if one looks a bit closer to the structure of the VELO on Figure 4.17 one sees that the density of VELO stations is significantly decreasing after 300 mm. This could explain the first dip in the efficiency at 250 mm since $\Lambda$s decaying after 250 mm have less chance to hit three VELO stations.

The second dip at 450 mm could be explained by the VELO station situated around this value of $Z$ coordinate. Indeed, if a particle decays right after 450 mm it will not hit this station and has quite few chances to hit all the last three VELO stations which is the only way it can produce a long track.

Figure 4.16: Overall reconstruction efficiency as a function of the $Z$ coordinate of the $\Lambda$ decay vertex for prompt $\Lambda$s (left) and $\bar{\Lambda}$s (right).

Figure 4.17: Internal structure of the VERTex LOcator. The shape of the reconstruction efficiency distribution is due to the lower density of VELO stations at the end part of the VELO.
Therefore one concludes that the distribution of the Z coordinate of the \( \Lambda \) end vertex is highly modified by reconstruction possibilities. Beyond 0.3 m the density of VELO stations decreases, leading to worse reconstruction capabilities.

### 4.4 Overall reconstruction efficiency

As a next step, the overall reconstruction efficiency was computed by dividing the distributions of reconstructed prompt \( \Lambda \)s by the distributions of all initial prompt \( \Lambda \)s. On Figure 4.19 one can observe that the best efficiency for \( \Lambda \) reconstruction can be achieved in a region around \( p_t = 1.5 \) GeV and \( \eta = 4 \).

**Figure 4.18**: Reconstruction efficiency for prompt \( \Lambda \)s (left) and prompt \( \bar{\Lambda} \)s (right) as a function of the transverse momentum and the pseudorapidity.

Reducing the bin size of this reconstruction efficiency plot shows its structure. This is represented on Figure 4.19 where the efficiency has a rather smooth structure for low values of \( p_t \) (below 2.5 GeV/c) but not above which is due to low statistics in the high \( p_t \) range, even when \( 10^7 \) events are looked at.

**Figure 4.19**: Reconstruction efficiency for prompt \( \Lambda \)s (left) and prompt \( \bar{\Lambda} \)s (right) as a function of the transverse momentum and the pseudorapidity.

This smoothness is even better shown on Figure 4.20 where the same distribution is plotted only for \( \Lambda \)s in a three-dimensional graph.
To get a deeper insight one shall look now at the projections of this three dimensional plot on Figure 4.21 and Figure 4.22.

Looking at the projections one has to remember that the two considered variables, $p_t$ and $\eta$, are not completely independent. For a given momentum, a particle with a low transverse momentum...
has a higher pseudorapidity. Thus the projections change in shape between the different regions of the second variable.

For $p_t$, the change is mainly expressed in the slope of the low $p_t$ part which increases with pseudorapidity. As previously said, low $p_t$ particles are more likely to have a high $\eta$ and therefore be out of acceptance leading to worse reconstruction capabilities. What is also observed again are the irregularities and huge statistical uncertainties for $\Lambda$s with a high transverse momentum due to low statistics in that region.

Except for these high $p_t$ values, the pseudorapidity distributions in bins of $p_t$ on Figure 4.22 do not look so different, they only slightly change in shape. Nevertheless, the mean value for the pseudorapidity distributions is moving towards lower values as one goes to higher transverse momentum regions.

![Figure 4.22: Reconstruction efficiency for prompt $\Lambda$s as a function of the pseudorapidity for different ranges of transverse momentum.](image)

The efficiency distributions have been studied along to the way in which interaction with matter, the limited acceptance of the detector and the reconstruction capabilities are affecting it.

To conclude this discussion, the fraction of decaying, accepted and reconstructed prompt $\Lambda$s are presented in Table 4.1.

<table>
<thead>
<tr>
<th>Prompt $\Lambda$s decaying into $\pi$ and $p$</th>
<th>$61.35 \pm 0.02%$(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompt $\Lambda$s decaying into $\pi$ and $p$ both in acceptance</td>
<td>$13.35 \pm 0.02%$</td>
</tr>
<tr>
<td>Prompt $\Lambda$s decaying into $\pi$ and $p$ both reconstructed as long tracks</td>
<td>$0.51 \pm 0.003%$</td>
</tr>
</tbody>
</table>

Table 4.1: Overall reconstruction efficiency.

The overall reconstruction efficiency, going from all prompt $\Lambda$s to all reconstructed ones is $0.51\%$. The biggest loss is due to the requirement that both daughter particles have long tracks which means that the $\Lambda$ has to decay before leaving the VELO and that both its daughters fly through the whole detector without being too strongly deflected by the magnetic field.

\(^1\)This fraction is not equal to the branching ratio since a few $\Lambda$s are absorbed by matter before they can decay.
5 Selection efficiency

To eliminate as much background as possible while still keeping many true $\Lambda$s among the $\Lambda$ candidates one applies selection cuts [10]. There are two different kind of cuts that are applied, the first ones, rather loose, remove only background and help reducing the size of the files and the amount of computing power. The second ones are tuned and help increasing as much as possible the signal significance which is defined as the ratio \( \frac{\text{Signal}}{\sqrt{\text{Signal} + \text{Background}}} \) in order to optimize the quality of the selection.

5.1 Loose selection cuts

The loose selection cuts are applied to the $\Lambda$ candidate as well as to its daughters. They are conditions on a set of kinematic and geometric variables.

For the $\Lambda$, the variables on which one cuts are the reconstructed mass of the $\Lambda$, the $\chi^2$ of the $\Lambda$ decay vertex, the transverse momentum ($p_t$) of the $\Lambda$, the impact parameter of the $\Lambda$ (IP)$^1$ and the $\Lambda$ flight distance (distance between PV and $\Lambda$ decay vertex). The conditions applied on the daughter particles are cuts on their transverse momentum ($p_t$), their momentum ($p$) and their impact parameter (IP)$^1$.

All the loose selection cuts that have been applied are listed in Table 5.1.

<table>
<thead>
<tr>
<th>Cuts on $\Lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda$ mass</td>
</tr>
<tr>
<td>Decay vertex $\chi^2$</td>
</tr>
<tr>
<td>$p_t$</td>
</tr>
<tr>
<td>$\ln(\text{IP [mm]})$</td>
</tr>
<tr>
<td>$\ln(\text{Flight distance [mm]})$</td>
</tr>
<tr>
<td>$\ln(\text{Flight distance [mm]})$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cuts on $\Lambda$ daughters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proton $p_t$</td>
</tr>
<tr>
<td>Proton $p$</td>
</tr>
<tr>
<td>Proton $\ln(\text{IP [mm]})$</td>
</tr>
<tr>
<td>Pion $p$</td>
</tr>
<tr>
<td>Pion $p_t$</td>
</tr>
<tr>
<td>Pion $\ln(\text{IP [mm]})$</td>
</tr>
</tbody>
</table>

Table 5.1: Loose selection cuts.

5.2 Tight selection cuts and tuning

To get a better purity without losing too much efficiency, tighter cuts are chosen and applied in such a way to optimize the quality of the sample. As previously explained, this is done by maximizing the significance. These tuned cuts are less numerous than the previous ones not to affect too much the distribution of the true $\Lambda$s. They are again based on the same variables as before with the exception for the IP cut ($\nu$) which is a combination of three previous variables. All these tight selection cuts are listed in Table 5.2.

---

$^1$Impact parameter is defined as the distance between the PV and the extrapolated track of the particle.
Cuts on Λ

<table>
<thead>
<tr>
<th>Λ mass</th>
<th>1115.7 ± 7 MeV/c²</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln( Flight distance [mm] )</td>
<td>&gt; 4</td>
</tr>
<tr>
<td>( \nu = \ln\left( \frac{\text{PionIP}[\text{mm}]}{\text{ProtonIP}[\text{mm}] \cdot \text{LambdaIP}[\text{mm}]} \right) )</td>
<td>&gt; 2.5</td>
</tr>
</tbody>
</table>

**TABLE 5.2:** Tight selection cuts.

To illustrate these tuned tight cuts which maximize the signal significance, **FIGURE 5.1**, **FIGURE 5.2** and **FIGURE 5.3** show on semi-logarithmic histograms where they are situated with respect to the signal and the noise.

**FIGURE 5.1:** In green, the Λ mass distribution for reconstructed Λ (left) and \( \bar{\Lambda} \) (right) candidates having an associated Monte Carlo truth. Hatched in light red, the Λ mass distribution for all reconstructed Λ (left) and \( \bar{\Lambda} \) (right) candidates. The black arrows show the mass cut applied on the Λ candidates.

**FIGURE 5.2:** In green, the natural logarithm of the Λ flight distance distribution for reconstructed Λ (left) and \( \bar{\Lambda} \) (right) candidates having an associated Monte Carlo truth. Hatched in light red, the natural logarithm of the Λ flight distance distribution for all reconstructed Λ (left) and \( \bar{\Lambda} \) (right) candidates. The black arrow shows the flight distance cut applied on the Λ candidates.
\[
\nu = \ln(c IP \text{ [mm]} \ast p IP \text{ [mm]}) / \Lambda IP \text{ [mm]} = \ln(\nu)\]

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
 & Loose sel. cuts & Loose + tight sel. cuts \\
\hline
\text{All candidates} & 9543460 & 10318900 & 29004 & 29438 \\
\text{True } \Lambda & 31884 & 31308 & 22121 & 21940 \\
\text{Diffractive true } \Lambda & 1248 & 1278 & 823 & 833 \\
\text{True } \Lambda \text{ from matter interaction} & 35 & 26 & 15 & 7 \\
\text{True prompt } \Lambda & 30614 & 30126 & 21412 & 21266 \\
\text{True non-prompt } \Lambda & 1270 & 1182 & 709 & 674 \\
\text{Daughters from } K_S & 36262 & 37188 & 4076 & 4428 \\
\hline
\end{tabular}
\end{table}

**TABLE 5.3: Tight selection cut efficiency and purity**

- Total number of events: \(9.4686 \times 10^6\)
- Tight selection efficiency: \(69.4 \pm 0.3\% / 70.1 \pm 0.3\%\)
- Purity without tight selection cuts: \(0.334 \pm 0.002\% / 0.303 \pm 0.002\%\)
- Purity with tight selection cuts: \(76.3 \pm 0.3\% / 74.5 \pm 0.3\%\)
- \(K_S\) contamination after tight selection: \(14.1 \pm 0.2\% / 15.0 \pm 0.2\%\)

In order to illustrate quantitatively the improved purity, **FIGURE 5.4** shows the pseudorapidity distribution of true \(\Lambda\)s and the one of false candidates having passed the tight selection cuts.
5.4 Reducing $K^0_S$ contamination

The biggest contamination of our $\Lambda$ candidates are the $K^0_S$ misidentified as $\Lambda$ (around 15%). Therefore, getting rid of these particles was the first thing to start with.

A very promising method is to do again the calculation for each candidate’s four-vector as explained before in section 2, but associating this time the mass of a pion to both measured daughter tracks. If both daughter particles (initially assumed to be $\pi$ and $p$) are actually pions coming from $K^0_S \rightarrow \pi^+ \pi^-$, one should see a peak at the $K^0_S$ mass. The reconstructed mass distribution where the daughter tracks are given the mass of a pion is shown in Figure 5.5 where one clearly sees the $K^0_S$ contamination mass peak.

![Figure 5.4](image1.png)  ![Figure 5.4](image2.png)

**Figure 5.4:** In green, pseudorapidity distribution for selected $\Lambda$ (left) and $\bar{\Lambda}$ (right) candidates having an associated Monte Carlo truth. In red, pseudorapidity distribution for selected $\Lambda$ (left) and $\bar{\Lambda}$ (right) candidates having no associated Monte Carlo truth.

![Figure 5.5](image3.png)

**Figure 5.5:** Reconstructing the $K^0_S$ mass with 2 hypothetical $\pi$s gives a mass peak which one can cut out to get rid of the $K^0_S$ noise (in red).
In order to determine the best cut one looks at the significance as a function of the mass window of the cut on the hypothetical $K^0_S$ mass. As shown on Figure 5.6 the optimal window is 5.5 MeV/c$^2$ on each side of the $K^0_S$ mass (total mass window of 11 MeV/c$^2$).

![Figure 5.6: The significance as a function of the mass window of a cut on the hypothetical $K^0_S$ mass.](image)

Using this cut, the $K^0_S$ contamination is lowered from 7771 to 1570 (-80 %) whereas the number $\Lambda$s and $\bar{\Lambda}$s is only decreasing from 40380 to 38386 (-5 %). The significance of true $\Lambda$s increased by 3 %.

So far this cut looks nice, but it is important that it doesn’t affect too much the shape of the measured distributions. Therefore one looks at the way it changes the true $\Lambda$ distributions, for instance pseudorapidity or $p_t$. To check that the removed true $\Lambda$s are uniformly distributed in $p_t$ and pseudorapidity one takes the ratio between the true $\Lambda$s that are removed by the cut and the true $\Lambda$s that pass it, as presented in Figure 5.7.
Figure 5.7: Ratio between the true Λs that are removed by the $K^0_S$ mass cut and the true Λs that pass it as a function of pseudorapidity. Where again true Λ refers to Λs having a Monte Carlo correspondent.

One sees that there is a slight linear dependence in pseudorapidity. Nevertheless, the dependence looks weak enough compared to the efficiency of the cut and hence the $K^0_S$ mass cut seems sane.

5.5 Reducing diffractive contamination

Another contamination one would like to get rid of are diffractive events (2.8 % after the tight selection). There are actually two type of such events, single and double diffractive ones. The single diffractive events are those where only one of both colliding protons produces particles whereas double diffractive events give particles from both protons. In both cases there is no exchange of quantum number between the beam protons. It is very difficult to distinguish between double diffractive events and non-diffractive events since the only difference is that no quantum number is exchanged between the colliding protons. Therefore one tries mainly to get rid of single diffractive events.

The most promising method is to look at the number of backward VELO tracks since single diffractive events are mainly going into one direction. Hence those particles are highly boosted into the direction of the initial proton. Therefore if a Λ is detected in the forward direction which comes from such an event, there should be very few particles hitting the VELO in the backward direction.

If one looks now at the distribution of the Λs as a function of the number of backward VELO tracks on Figure 5.8 one sees that the diffractive contamination is really located at low values.
To determine the best cut one could apply, the significance of non-diffractive events as a function of the cut is considered. Unfortunately one sees on Figure 5.9 that the best cut optimizing this significance would be no cut at all. But for some studies where high purity is required one may apply such a cut though.

The problem is coming from the fact that the diffractive contamination is too low and spread over too many bins of backward VELO tracks. The same approach has also been tested with less...
promising variables such as the total number of tracks, the total number of VELO tracks or the ratio between backward and forward VELO tracks but those lead to even worse results. So it seems impossible to increase the significance of $\Lambda$s coming from non-diffractive events using a cut on the number of backward VELO tracks or other cuts based on track counting.

5.6 Final selection efficiency

It is interesting to look at the final values of efficiency and purity including the cut on the hypothetical $K_s^0$ mass. The relative purities are shown in TABLE 5.4.

<table>
<thead>
<tr>
<th>Fraction</th>
<th>Purity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of true $\Lambda$s</td>
<td>85.8 ± 0.2 %</td>
</tr>
<tr>
<td>$K_s^0$ background</td>
<td>3.5 ± 0.1 %</td>
</tr>
<tr>
<td>Diffractive $\Lambda$ contamination</td>
<td>3.8 ± 0.1 %</td>
</tr>
<tr>
<td>Non-prompt $\Lambda$ contamination</td>
<td>3.1 ± 0.1 %</td>
</tr>
</tbody>
</table>

TABLE 5.4: Overall selection purity.

Concerning the efficiency, 66.4 ± 0.2 % of all prompt and non-diffractive $\Lambda$s survive the cuts and one finally gets a purity of 80.0 ± 0.2 % of prompt and non-diffractive $\Lambda$s.

6 Efficiency corrected distributions

The previous sections gave a detailed insight into the overall reconstruction and selection efficiencies. Now these results will be applied to correct a distribution from a sample obtained without Monte Carlo truth knowledge. The purpose is to check if the obtained distributions are correct and to estimate their statistical errors. To do so one takes the overall efficiency (reconstruction and selection) in bins of pseudorapidity and transverse momentum and divides the reconstructed distributions by it.

6.1 Pseudorapidity distributions

First, the shape of the overall (reconstruction + selection) efficiency as a function of pseudorapidity is shown on FIGURE 6.1.

For this study, $10^7$ events have been simulated. To get an efficiency correction that is independent of the sample on which it is applied, half of the events are used to calculate the efficiency correction distribution. Then this correction is applied to the other half and finally compared to the
correspondent Monte Carlo truth. Therefore the effective number of events is $5 \times 10^6$. Nevertheless, the first real data sample that will be used for such a study is expected to contain $10^8$ events.

In the following efficiency distributions, the number of bins have been chosen small enough to guarantee good statistics and small statistical errors. The purpose here is to get less than 5% errors on the corrected distributions for $10^8$ events. As an example, the result of an efficiency correction is compared to the associated Monte Carlo truth on FIGURE 6.2.

![Figure 6.2](image)

**FIGURE 6.2:** Selected and corrected $\Lambda$ pseudorapidity distribution in black and associated Monte Carlo truth in red.

One sees that both distributions coincide and that the statistical error on the reconstructed distribution is varying down to 10% in the optimal region, between 3.5 and 4.5 in pseudorapidity.

### 6.2 $\overline{\Lambda}/\Lambda$ Ratio

As previously explained, the $\overline{\Lambda}/\Lambda$ ratio might be of great help to determine the most adequate Monte Carlo tuning and best physics model. Therefore we’ll look at the dependence of this ratio with respect to the pseudorapidity at Monte Carlo level and for selected $\Lambda$s on FIGURE 6.3.

![Figure 6.3](image)

**FIGURE 6.3:** $\overline{\Lambda}/\Lambda$ ratio at Monte Carlo Truth level and for selected $\Lambda$s.

One sees that the distribution of the $\overline{\Lambda}/\Lambda$ ratio is not the same for selected $\Lambda$s and for the true ones. The previous pseudorapidity distribution corrections are now applied to this ratio to check its accuracy and to obtain again an estimate of the statistical errors (FIGURE 6.4).
Figure 6.4: Corrected $\bar{\Lambda}/\Lambda$ ratio for selected $\Lambda$s in black and associated Monte Carlo truth in red.

The number of bins has again been decreased to gain on the statistical errors. The smallest error bars in the optimal pseudorapidity range are now 15%. This study has been performed on a sample with $5 \times 10^6$ events, for a sample with $10^8$ events these statistical errors are approximately divided by a factor 4.5. Hence statistical errors of less than 5% can be obtained using $10^8$ and taking the number of bins presented above.

7 Conclusion

This study provides a basis for the $\Lambda$ measurements that are expected to be done at LHCb in early 2010.

First the origin and the decay or absorption of $\Lambda$s have been well documented. It was seen that within the resolution of the primary vertex, mainly prompt $\Lambda$s are produced (99.75%).

Several distributions and their change in shape due to the detector and the reconstruction capabilities have been studied. It was understood that the beam pipe in the RICH1 region absorbs particles at a pseudorapidity of 4.35 and that the VELO detector geometry leads to irregularities in the $Z$ coordinate of the $\Lambda$ end vertex distribution of reconstructed $\Lambda$s.

The selection cuts have also been described and their efficiency documented. Furthermore, a new cut that helps to reduce $K_0^0$ contamination has been introduced and discussed. The final efficiency is now 66.4% and the purity 80.0% when only prompt and non-diffractive $\Lambda$s are considered as signal.

To complete the efficiency study, some test distributions have been reconstructed and corrected. This helped to calculate the statistical errors one will have when correcting real data distributions. It appeared that the statistical errors are small enough to distinguish between different Monte Carlo tunings and their associated physics models.

The reconstruction and selection efficiency also depends on the polarization of the $\Lambda$s but in the simulated data used for this study the $\Lambda$s were not polarized. Thus, further simulations and studies will be done including that effect what may change the efficiency distributions.

Finally, real data will soon be used to continue this study and will allow extending it later to $b$ baryons.
References


