A Study of $B^0 \rightarrow D^{*-} \tau^+ \nu_\tau$ with Three-Prong $\tau$ Decays at LHCb

Supervisor:
Dott. Concezio Bozzi

Additional Supervisor:
Dott. Guy Wormser

Examiner:
Prof. Paolo Lenisa

Candidate
Federico Betti

Academic year 2013-2014
# Contents

## Introduction

xi

## 1 Theory of flavour violation in semileptonic decays

1.1 The Glashow-Weinberg-Salam model ........................................... 1

1.2 The Cabibbo-Kobayashi-Maskawa matrix ....................................... 5

1.3 Quantum Chromodynamics .......................................................... 7

1.4 The theory of semileptonic decays ............................................. 8

1.4.1 The Heavy Quark Effective Theory ......................................... 8

1.4.2 Standard Model expectation for $B(B^0 \to D^{*-}(\tau^+\nu))/B(B^0 \to D^{*-}(\ell^+\nu))$ ...................................................... 9

1.5 Limitations of the Standard Model ............................................. 11

1.6 New Physics in $B^0 \to D^{*-}(\tau^+\nu)$ ..................................... 13

## 2 The LHCb experiment

2.1 The Magnet ................................................................. 18

2.2 The Tracking System .......................................................... 19

2.2.1 The VErtex LOcator ......................................................... 19

2.2.2 The Silicon Tracker ....................................................... 20

2.2.3 The Outer Tracker .......................................................... 23

2.3 Particle Identification .......................................................... 23

2.3.1 The Ring Imaging CHerenkov detectors .................................. 23

2.3.2 The Calorimeters ............................................................ 25

2.3.3 The Muon System ........................................................... 27

2.4 The Trigger System ............................................................. 28

2.5 Detector performance ............................................................ 30

2.5.1 Track reconstruction ......................................................... 30

2.5.2 Particle Identification ....................................................... 31
3 Events selection

3.1 Analysis strategy .................................................. 33
3.2 Data and MonteCarlo samples ................................. 37
3.3 First cut-based selection ........................................ 40

4 Study of the data sample composition and isolation of signal 47

4.1 Composition of surviving events .............................. 47
4.1.1 Combinatorial background .................................. 47
4.1.2 Misidentified kaon contribution ............................ 49
4.1.3 Branching ratios of $D^+_s \to \pi^+ \pi^- \pi^+ N$ .............. 49
4.1.4 Fit of exclusive $D^+_s$ and $D^+$ peaks and estimate of signal contribution ...................... 52
4.1.5 The “$D^+_s$-o-meter” ........................................... 53
4.2 Isolation of signal .................................................. 56
4.2.1 Partial reconstruction ........................ .................... 57
4.2.2 MultiVariate Analysis ......................................... 59

5 Conclusion and perspectives ................................. 71
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparison of the various measurements of $R_D$</td>
<td>12</td>
</tr>
<tr>
<td>1.2</td>
<td>Feynman diagram of $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$</td>
<td>14</td>
</tr>
<tr>
<td>2.1</td>
<td>The LHCb apparatus</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>Perspective view of the LHCb dipole magnet</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Magnetic field along the z axis</td>
<td>19</td>
</tr>
<tr>
<td>2.4</td>
<td>Cross section in the $(x,z)$ plane of the VELO silicon sensors</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Overview of the VELO vacuum vessel</td>
<td>21</td>
</tr>
<tr>
<td>2.6</td>
<td>Layout of the four detection layers of the TT station</td>
<td>22</td>
</tr>
<tr>
<td>2.7</td>
<td>View of the four IT detector boxes arranged around the LHC beampipe</td>
<td>22</td>
</tr>
<tr>
<td>2.8</td>
<td>Layout of a detection layer in the second IT station</td>
<td>22</td>
</tr>
<tr>
<td>2.9</td>
<td>Schematic view of the OT stations</td>
<td>23</td>
</tr>
<tr>
<td>2.10</td>
<td>Side view schematic layout of the RICH 1 detector</td>
<td>24</td>
</tr>
<tr>
<td>2.11</td>
<td>Top view schematic of the RICH 2 detector</td>
<td>24</td>
</tr>
<tr>
<td>2.12</td>
<td>Front view of one half of the SPD/PS</td>
<td>26</td>
</tr>
<tr>
<td>2.13</td>
<td>Individual scintillator pad with the WLS fibre</td>
<td>26</td>
</tr>
<tr>
<td>2.14</td>
<td>Downstream view of the ECAL installed</td>
<td>26</td>
</tr>
<tr>
<td>2.15</td>
<td>Outer, middle and inner type ECAL modules</td>
<td>26</td>
</tr>
<tr>
<td>2.16</td>
<td>View from upstream of the HCAL detector installed</td>
<td>27</td>
</tr>
<tr>
<td>2.17</td>
<td>A schematic of the internal cell structure of the HCAL</td>
<td>27</td>
</tr>
<tr>
<td>2.18</td>
<td>Schematic side view of the muon system</td>
<td>29</td>
</tr>
<tr>
<td>2.19</td>
<td>Front view of a quadrant of a muon station and division into logical pads</td>
<td>29</td>
</tr>
<tr>
<td>2.20</td>
<td>Flow-diagram of the different trigger sequences</td>
<td>30</td>
</tr>
<tr>
<td>2.21</td>
<td>Relative momentum resolution versus momentum</td>
<td>31</td>
</tr>
<tr>
<td>2.22</td>
<td>Kaon identification efficiency and pion misidentification rate</td>
<td>32</td>
</tr>
<tr>
<td>3.1</td>
<td>The inversion cut at $5\sigma$</td>
<td>41</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.2</td>
<td>The distributions of the variables on which cuts are applied (1)</td>
<td>44</td>
</tr>
<tr>
<td>3.3</td>
<td>The distributions of the variables on which cuts are applied (2)</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>Two dimensional distribution of $\Delta M$ as function of $M(D^0)$</td>
<td>46</td>
</tr>
<tr>
<td>3.5</td>
<td>The $3\pi$ mass distribution after the first cut-based selection</td>
<td>46</td>
</tr>
<tr>
<td>4.1</td>
<td>Fit of $\Delta M(D^{<em>-}) = M(D^{</em>-}) - M(\bar{D}^0)$ distribution</td>
<td>48</td>
</tr>
<tr>
<td>4.2</td>
<td>Fit of $D^+$ peak in $M(3\pi)$ distribution with a kaon hypothesis on $\pi^-$</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>Plot of the $3\pi$ mass with the kaon hypothesis</td>
<td>50</td>
</tr>
<tr>
<td>4.4</td>
<td>Fit of $D_s^+$ peak in $M(3\pi)$ distribution</td>
<td>53</td>
</tr>
<tr>
<td>4.5</td>
<td>Fit of $D^+$ peak in $M(3\pi)$ distribution</td>
<td>54</td>
</tr>
<tr>
<td>4.6</td>
<td>The $\min</td>
<td>M(\pi^+\pi^-)</td>
</tr>
<tr>
<td>4.7</td>
<td>The $\min</td>
<td>M(\pi^+\pi^-)</td>
</tr>
<tr>
<td>4.8</td>
<td>Schematic representation of $B^0 \rightarrow D^{*-}D_s^+(\rightarrow \pi^+\pi^-\pi^+N)$ decay</td>
<td>57</td>
</tr>
<tr>
<td>4.9</td>
<td>The profile of the correction $dz$ in function of the $3\pi$ mass</td>
<td>58</td>
</tr>
<tr>
<td>4.10</td>
<td>Illustration of a binary decision tree with feature variables $x_{i,j,k}$</td>
<td>61</td>
</tr>
<tr>
<td>4.11</td>
<td>Input variables used in BDT(1)</td>
<td>64</td>
</tr>
<tr>
<td>4.12</td>
<td>Input variables used in BDT(2)</td>
<td>65</td>
</tr>
<tr>
<td>4.13</td>
<td>Input variables used in BDT(3)</td>
<td>66</td>
</tr>
<tr>
<td>4.14</td>
<td>$m_{D_s,vn}$ distribution for data after first cut-based selection</td>
<td>67</td>
</tr>
<tr>
<td>4.15</td>
<td>$m_{D_s,vn}$ distribution for signal MC after first cut-based selection</td>
<td>67</td>
</tr>
<tr>
<td>4.16</td>
<td>Overtraining check</td>
<td>68</td>
</tr>
<tr>
<td>4.17</td>
<td>ROC curve</td>
<td>68</td>
</tr>
<tr>
<td>4.18</td>
<td>Signal and background efficiencies, signal purity and signal significance in function of the BDT cut</td>
<td>69</td>
</tr>
<tr>
<td>4.19</td>
<td>The $3\pi$ mass after the BDT cut</td>
<td>69</td>
</tr>
<tr>
<td>4.20</td>
<td>The $3\pi$ mass in the signal MC sample after the first cut-based selection</td>
<td>70</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Quantum numbers of leptons and quarks</td>
<td>3</td>
</tr>
<tr>
<td>3.1</td>
<td>Table of branching ratios for signal channel and physical background channels</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>MC samples used in the study</td>
<td>38</td>
</tr>
<tr>
<td>3.3</td>
<td>The stripping cuts applied to the tracks</td>
<td>39</td>
</tr>
<tr>
<td>3.4</td>
<td>The stripping cuts applied on the reconstructed particles</td>
<td>40</td>
</tr>
<tr>
<td>3.5</td>
<td>Efficiencies of the reconstruction and stripping pre-selection</td>
<td>41</td>
</tr>
<tr>
<td>3.6</td>
<td>The first cut-based selection</td>
<td>42</td>
</tr>
<tr>
<td>3.7</td>
<td>Efficiencies of the first cut-based selection</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>Branching ratios of $D_s^+$ into final states containing three charged pions and neutral particles (1)</td>
<td>51</td>
</tr>
<tr>
<td>4.2</td>
<td>Branching ratios of $D_s^+$ into final states containing three charged pions and neutral particles (2)</td>
<td>51</td>
</tr>
<tr>
<td>4.3</td>
<td>Branching ratios of $D^+$ into final states containing three charged pions and neutral particles</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>Ratios between $B$ and efficiencies used to extrapolate the contributions of the various decays</td>
<td>54</td>
</tr>
<tr>
<td>4.5</td>
<td>Estimated yields for signal and background contributions in the data sample</td>
<td>54</td>
</tr>
<tr>
<td>4.6</td>
<td>Input variables used in the BDT</td>
<td>62</td>
</tr>
</tbody>
</table>
Introduction

The $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ is sensitive to physics processes beyond the Standard Model (SM). It has recently gained attention after the measurement, performed at BaBar, of $R_{D^*} = \mathcal{B}(B^0 \rightarrow D^{*-}\tau^+) / \mathcal{B}(B^0 \rightarrow D^{*-}\ell^+\nu)$, which exceeds the SM prediction at the $3\sigma$ level. Therefore other measurements of this process are crucial to confirm or disprove this discrepancy.

The purpose of this thesis is to study the feasibility at LHCb of the precise measurement of $\mathcal{B}(B^0 \rightarrow D^{*-}\tau^+\nu_\tau)$, using the tau hadronic decay with three charged particles in the final state, i.e. $\tau^+ \rightarrow \pi^+\pi^-\pi^+(\pi^0)\nu_\tau$. The study of this tau decay mode is complementary in many ways to similar studies carried with the more abundant tau leptonic decays. In particular, three-prong tau decays:

- allow the reconstruction of the tau decay vertex, which gives a powerful criterion to discriminate between signal and the most abundant background source due to hadronic $B$ decays;
- avoid completely the difficult task of discriminating $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ signal decays, where the tau decays leptonically, from much more abundant semileptonic decays involving lighter leptons.

The theoretical framework for this study is described in Chapter 1 together with numerical predictions and experimental results. The study has been performed on the data taken by LHCb collaboration in 2011, corresponding to 1 fb$^{-1}$ of integrated luminosity. The LHCb detector is briefly illustrated in Chapter 2. The data sample has been subjected to a first cut-based selection, in order to suppress the most dominant background due to $B^0 \rightarrow D^{*-}\pi^+\pi^-\pi^+(\pi^0)$ decays. The efficiency of this selection has been evaluated on MonteCarlo (MC) samples. In Chapter 3, this selection and its effects are described together with the data and MC samples used. An estimation of various contributions in data sample after the first cut-based selection is given in the first part of Chapter 4. This estimation has been done with
the help of simulated and real data samples. One more stage necessary to isolate the signal component consists in applying a partial reconstruction technique, in order to get new dynamical variables which are used as input for a MultiVariate Analysis (MVA), described in Chapter 4. This MVA is useful to discriminate signal from the dominant background surviving after the first cut-based selection, which is given by $B^0 \rightarrow D^{*-}D^{+}_{(s)}X$ decays, where the $D^{+}_{(s)}$ decays in three charged pions and one or more undetected particles, and to get an estimate of the expected number of signal events. The results of this study are summarized and commented in Chapter 5, together with a list of improvements and next steps for the continuation of this analysis.
Chapter 1

Theory of flavour violation in semileptonic decays

The model which provides a unified and experimentally-established picture of electroweak interactions is the Glashow-Weinberg-Salam model (GWS) [1][2][3]. It is a non-abelian, Yang-Mills quantum field theory based on the \( \text{SU}(2) \times \text{U}(1) \) gauge symmetry group. The resulting theory obtained including the \( \text{SU}(3) \) symmetry of strong interactions and the Brout-Englert-Higgs spontaneous symmetry breaking mechanism, is called Standard Model of Particle Physics (SM), and it describes all the known interactions electroweak and strong and elementary particles (quarks, leptons, gauge and Higgs bosons) existing in nature, with the exception of gravity. The GWS model is described in the next section, while the \( \text{SU}(3) \) component of the SM will be described briefly in section 1.3.

1.1 The Glashow-Weinberg-Salam model

Before the electroweak symmetry breaking (which will be discussed later) the lagrangian of the GWS model can be divided into four main terms:

\[
\mathcal{L} = \mathcal{L}_B + \mathcal{L}_f + \mathcal{L}_H + \mathcal{L}_Y. \tag{1.1}
\]

\( \mathcal{L}_B \) is the kinetic term of the four gauge fields, that are \( B_\mu \) (the hypercharge field) and \( W^a_\mu (a = 1, 2, 3) \) (the weak isospin fields). Calling \( W^{a}_{\mu \nu} \) and \( B_{\mu \nu} \) their field strength tensors, this term is equal to:

\[
\mathcal{L}_B = -\frac{1}{4} W^{a \mu \nu} W^{a}_{\mu \nu} - \frac{1}{4} B^{\mu \nu} B_{\mu \nu}. \tag{1.2}
\]
The second term describes the kinetic of the fermions and their interactions with the gauge bosons:

$$\mathcal{L}_f = \overline{Q}^i i \not{D}_L Q^i + \overline{u}_R^i i \not{D}_R u_R^i + \overline{d}_R^i i \not{D}_R d_R^i + \overline{L}_L^i i \not{D}_L L_L^i + \overline{e}_R^i i \not{D}_R e_R^i,$$  \hspace{1cm} (1.3)

where $Q^j$ are the left-handed quark doublets, $u_R^j$ and $d_R^j$ are the up and down right-handed quark singlets, while $L^j$ are the left-handed lepton doublets and $e_R^j$ are the electron singlets.

$$Q^i = \left( \begin{array}{c} u_L^i \\ d_L^i \end{array} \right),$$  \hspace{1cm} (1.4)

$$L^j = \left( \begin{array}{c} \nu_L^j \\ e_L^j \end{array} \right).$$  \hspace{1cm} (1.5)

In Eq. (1.3), a sum over $j$ is assumed, where $j$ is the generation index, which runs from 1 to 3. $D_{L\mu}$ and $D_{R\mu}$ are the covariant derivative for the left-handed and right-handed fermion fields, needed to keep the lagrangian invariant under the SU(2) × U(1) local gauge transformation, and they are different depending on which field they are applied to. They are equal to:

$$D_{L\mu} = \partial_\mu + igW^a_\mu \sigma^a / 2 + ig' Y / 2 B_\mu,$$  \hspace{1cm} (1.6a)

$$D_{R\mu} = \partial_\mu + ig' Y / 2 B_\mu,$$  \hspace{1cm} (1.6b)

where $Y$ is the hypercharge of the field on which $D_\mu$ operates, $\sigma^a$ are the Pauli matrices, $g$ and $g'$ are the coupling constants. The fields that form the SU(2) doublets have weak isospin $T = 1/2$, with third component $T_3 = \pm 1/2$ for up and down-type fields, respectively. The right-handed fermion fields, which are SU(2) singlets, have $T = 0$.

It is possible to define the electromagnetic charge $Q$ of a field in terms of its hypercharge $Y$ and third component of weak isospin $T_3$:

$$Q = \frac{Y}{2} + T_3.$$  \hspace{1cm} (1.7)

From this definition one can obtain the values of the hypercharges of all fermionic doublets and singlets according to their electromagnetic charge. The values of $T_3$, $Y$ and $Q$ of the fermion fields are reported in Tab. 1.1.
Table 1.1: Quantum numbers of leptons and quarks: third component of the weak isospin $T_3$, hypercharge $Y$ and electromagnetic charge $Q$.

<table>
<thead>
<tr>
<th>Fermion</th>
<th>$T_3$</th>
<th>$Y$</th>
<th>$Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_L$</td>
<td>1/2</td>
<td>1/3</td>
<td>2/3</td>
</tr>
<tr>
<td>$d_L$</td>
<td>-1/2</td>
<td>1/3</td>
<td>-1/3</td>
</tr>
<tr>
<td>$u_R$</td>
<td>0</td>
<td>4/3</td>
<td>2/3</td>
</tr>
<tr>
<td>$d_R$</td>
<td>0</td>
<td>-2/3</td>
<td>-1/3</td>
</tr>
<tr>
<td>$\nu_L$</td>
<td>1/2</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>$e_L$</td>
<td>-1/2</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$e_R$</td>
<td>0</td>
<td>-2</td>
<td>-1</td>
</tr>
</tbody>
</table>

The third term of Eq. (1.1) concerns the Higgs field and its coupling with the gauge bosons [4][5]:

\[
\mathcal{L}_H = (D^\mu \phi^\dagger)(D_\mu \phi) - V(\phi^\dagger \phi) = (D^\mu \phi^\dagger)(D_\mu \phi) - \left(-\mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2\right),
\]

where $\lambda$ and $\mu$ are positive parameters and $\phi$ is the SU(2) Higgs doublet with hypercharge 1:

\[
\phi = \left(\begin{array}{c} \phi^+ \\ \phi^0 \end{array}\right),
\]

with $\phi^+$ and $\phi^0$ electromagnetic charged and neutral complex scalar fields. $\phi^+$ is a SU(2) doublet with hypercharge 1, so the covariant derivative that operates on it is:

\[
D_\mu = \partial_\mu + ig W^a_\mu \sigma^a_2 + \frac{1}{2} ig' B_\mu,
\]

The last term of Eq. (1.1) gives the Yukawa interaction between the fermion fields and $\phi$, it is needed to generate the fermion masses, as described below:

\[
\mathcal{L}_Y = -\lambda_{d}^{ij} \bar{Q}^j \phi d^i_R - \lambda_{u}^{ij} \bar{Q}^j (i\sigma^2 \phi) u^i_R - g^i_e \bar{L}^j \phi e^i_R + h.c.,
\]

where a sum over indices $i$ and $j$ is assumed, $\lambda_{d,u}^{ij}$ are general complex-valued matrices and $g^i_e$ are coupling constants.

The Higgs potential $V(\phi^\dagger \phi)$ has a minimum when $\phi^\dagger \phi = \mu^2/\lambda^2 \equiv v^2/2$, where $v/\sqrt{2}$ is the vacuum expectation value (VEV) of $\phi$ ($v \simeq 246$ GeV). Using the fact that the four degrees of freedom of $\phi$ can be reduced to one through the SU(2) gauge invariance, the doublet can be written in the unitarity gauge and expanded...
around its own VEV:

\[ \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \]  
(1.12)

where \( H(x) \) is the Higgs field, that is scalar and real. Now the Lagrangian is no longer SU(2) invariant, because the symmetry has been spontaneously broken when a specific direction of vacuum has been chosen.

After defining the physical gauge fields

\[ W^\pm_\mu = W^1_\mu \pm iW^2_\mu / \sqrt{2}, \]  
(1.13a)
\[ Z_\mu = W^3_\mu \cos \theta_W - B_\mu \sin \theta_W, \]  
(1.13b)
\[ A_\mu = W^3_\mu \sin \theta_W + B_\mu \cos \theta_W, \]  
(1.13c)

where \( \theta_W \) is the Weinberg angle (\( \sin^2 \theta_W \simeq 0.23 \)), the following mass terms are obtained substituting (1.12) in (1.8):

\[ - \frac{1}{8} g^2 v^2 (W^+ \mu W^\mu + W^- \mu W^- \mu) - \frac{1}{8} v^2 (g^2 + g'^2) Z^\mu Z_\mu - \frac{1}{2} \lambda^2 v^2 H^2. \]  
(1.14)

From (1.14) is clear that the masses of the gauge bosons and \( H \) are:

\[ M_H = \lambda v, \]  
(1.15a)
\[ M_W = \frac{1}{2} g v, \]  
(1.15b)
\[ M_Z = \frac{1}{2} \sqrt{g^2 + g'^2} v, \]  
(1.15c)
\[ M_\gamma = 0. \]  
(1.15d)

So the three weak bosons acquire mass, while the photon remains massless.

Substituting (1.12) in (1.11) gives

\[ \mathcal{L}_Y = -\frac{v}{\sqrt{2}} \lambda_d L i \bar{d}_L d^i_R - \frac{v}{\sqrt{2}} \lambda_u i \bar{u}_L u^i_R - \frac{v}{\sqrt{2}} g_i e^i L e^i_R + \text{h.c.} \]  
(1.16)

Then the mass of \( e^i \) is equal to

\[ \frac{v}{\sqrt{2}} g^i e, \]  
(1.17)

proportional to the coupling between the electron (muon, tau) and the Higgs Boson, while the neutrino remains massless.
1.2 The Cabibbo-Kobayashi-Maskawa matrix

For what concerns the quarks, mass eigenstates are needed, since here $u^i$ and $d^i$ are weak eigenstates and not the physical fermions seen experimentally. This is why the matrices $\lambda_{u,d}$ have been introduced in (1.11). The mass terms for physical quarks are obtained by diagonalizing the $\lambda_{u,d}$ matrices. In order to do that, unitary matrices $U_{u,d}, W_{u,d}$ are defined such that [6]:

$$\lambda_{u,d}^\dagger \lambda_{u,d} = U_{u,d} D_{u,d}^2 U_{u,d}^\dagger, \quad (1.18a)$$

$$\lambda_{u,d}^\dagger \lambda_{u,d} = W_{u,d} D_{u,d}^2 W_{u,d}^\dagger, \quad (1.18b)$$

where $D_{u,d}$ are diagonal matrices. This leads to:

$$\lambda_{u,d} = U_{u,d} D_{u,d} W_{u,d}^\dagger, \quad (1.19)$$

Now the physical quark fields can be defined in this way:

$$u_i^L = U_{i,j}^u u_j^{i,phys}, \quad (1.20a)$$

$$u_i^R = W_{i,j}^u u_j^{i,phys}, \quad (1.20b)$$

$$d_i^L = U_{i,j}^d d_j^{i,phys}, \quad (1.20c)$$

$$d_i^R = W_{i,j}^d d_j^{i,phys}, \quad (1.20d)$$

which, together with (1.19), allows to get in (1.16) the terms:

$$-\frac{v}{\sqrt{2}} D_{ii}^{\mu} u_i^{\mu,phys} u_i^{\mu,phys} - \frac{v}{\sqrt{2}} D_{ii}^{\mu} d_i^{\mu,phys} d_i^{\mu,phys}. \quad (1.21)$$

So the quark masses are defined as:

$$m_{i,u,d}^i = \frac{v}{\sqrt{2}} D_{ii}^{\mu} \quad (1.22)$$

Besides the replacement of weak eigenstates quark fields with mass eigenstates, $W_{u,d}$ and $U_{u,d}$ have another important effect, which can be seen in the weak charged current. It is possible to demonstrate that in (1.3) there are the following terms:

$$-\frac{g}{\sqrt{2}} (J^{+\mu} W^+_{\mu} + J^{-\mu} W^-_{\mu}) - \frac{g}{\cos \theta_W} J^{N\mu} Z_{\mu}, \quad (1.23)$$

where $J^{\pm\mu}$ and $J^{N\mu}$ are the charged and neutral currents:

$$J^{+\mu} = \bar{\nu}_L^a \gamma^\mu e_L^a + \bar{u}_L^a \gamma^\mu d_L^a, \quad (1.24a)$$

$$J^{-\mu} = \text{h.c.}(J^{+\mu}), \quad (1.24b)$$

$$J^{N\mu} = \sum_{a,f} f_a \frac{\gamma^\mu}{2} \left[T_3 - (T_3 - 2\sin^2 \theta_W Q) \gamma^5\right] f^a, \quad (1.24c)$$
with $f^a$ standing for a generic fermion (neutrino, electron, quark) of the $a$-th generation, $T_3$ its weak isospin and $Q$ its electromagnetic charge. Using (1.20), the terms involving quarks in (1.24a) can be written in this form:

$$\bar{u}^{\mathrm{phys}}_L (U^\dagger u)_{ij} \gamma^\mu d^{\mathrm{phys}}_L.$$  \hfill (1.25)

The matrix $(U^\dagger u U_d \equiv V_{CKM})$ is the Cabibbo-Kobayashi-Maskawa matrix (CKM).

The charged-current interaction lagrangian for quarks can now be written in this way:

$$\mathcal{L}_{cc,\text{quarks}} = -\frac{g}{\sqrt{2}} \left( \bar{u} \gamma^\mu c \right) \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \gamma^\mu \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W^\pm + \text{h.c.},$$  \hfill (1.26)

where the spinors written here represent the physical quark fields. This makes clear that the $W^\pm$ bosons mediate interactions between up-type and down-type quarks also between different families, so the quark flavour can change in weak interactions. Every element of the CKM matrix describes the coupling strength between two different quarks, for example $|V_{tb}| \approx 1$ and $|V_{ub}| \approx 0.004$: this means that the coupling between $t$ and $b$ is very strong compared to the one between $u$ and $b$.

The CKM matrix can be parameterized by three mixing angles $\theta_{ij}$ and a $CP$-violating phase $\delta$ (after defining $s_{ij} \equiv \sin \theta_{ij}$ and $c_{ij} \equiv \cos \theta_{ij}$):

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} + \mathcal{O}(\lambda^4),$$  \hfill (1.27)

It is know experimentally that $s_{13} \ll s_{23} \ll s_{12} \ll 1$, so the CKM matrix can be written in the so-called Wolfenstein parameterization, which takes into account the hierarchy of the matrix elements:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4),$$  \hfill (1.28)

where:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}},$$  \hfill (1.29a)

$$s_{23} = A\lambda^2 = \lambda \frac{|V_{cb}|}{V_{us}},$$  \hfill (1.29b)

$$s_{13}e^{i\delta} = V_{ub}^* = A\lambda^3(\rho + i\eta).$$  \hfill (1.29c)
According to the experimental measurements [7], the magnitudes of the elements of the CKM matrix are:

\[
\begin{pmatrix}
0.97425 \pm 0.00022 & 0.2253 \pm 0.0008 & 0.00413 \pm 0.00049 \\
0.225 \pm 0.008 & 0.986 \pm 0.016 & 0.0411 \pm 0.0013 \\
0.0084 \pm 0.0006 & 0.0400 \pm 0.0027 & 1.021 \pm 0.032
\end{pmatrix}.
\]  

(1.30)

The transition between different quark families does not take place in the neutral currents. In fact, after the substitution of (1.20) in (1.24c), it is clear, thanks to the unitarity of \( U_{u,d} \) and \( W_{u,d} \), that in neutral-current interactions a fermion interacts with its antiparticle or with an identical fermion. So the SM does not foresee Flavour-Changing Neutral Currents (FCNC) at tree-level [8].

### 1.3 Quantum Chromodynamics

Quantum Chromodynamics (QCD), the theory that explains the strong interactions between quarks and gluons, is necessary to complete the description of the SM. QCD is a Yang-Mills non-abelian quantum field theory based on the exact and unbroken colour-SU(3) local gauge symmetry [7]. The QCD lagrangian is:

\[
\mathcal{L}_{QCD} = \sum_F \ol{\psi}_F \left( i\gamma^\mu \partial_\mu - g_s \gamma^\mu \frac{\lambda^C}{2} A_\mu^C - m_F \right) \psi_F - \frac{1}{4} A^C_{\mu\nu} A^C_{\mu\nu},
\]  

(1.31)

where \( \psi_F \) is a triplet in the SU(3) space of quark spinors of flavour \( F \) and mass \( m_F \), \( g_s \) is the strong coupling constant, \( A_\mu^C \) are the massless gluon fields \( (C = 1,\ldots,8) \), \( A^C_{\mu\nu} \) are the gluon field strength tensors and \( \lambda^C \) are the eight Gell-Mann matrices, generators of the SU(3) group.

The strong coupling constant has a dependence on the exchanged momentum \( q^2 \):

\[
g_s = \frac{8\pi^2}{(11 - \frac{2}{3} n_f) \log \frac{q}{\Lambda_{QCD}}},
\]  

(1.32)

where \( n_f \) is the number of flavours and \( \Lambda_{QCD} \) is the energy scale of strong interactions. Experimentally, \( \Lambda_{QCD} \approx 200 \text{ MeV} \).

Equation (1.32) shows that the magnitude of the coupling constant decreases with increasing \( q^2 \) or with decreasing distances, so quarks and gluons behave as quasi-free particles at high energies (when \( q \gg \Lambda_{QCD} \)). Conversely, at low energies \( (q \ll \Lambda_{QCD}) \) the coupling constant becomes very high. This means that at high energies quark-gluon interactions can be treated perturbatively, while at low energies a perturbation expansion in series of \( \alpha_s = g_s/4\pi \) has no meaning.
1.4 The theory of semileptonic decays

The amplitude of the semileptonic decay of a meson $M_{Qq}$ into a state containing a meson $M_{q'q}$ can be written as a term proportional to the product of a leptonic current $L^\mu$ and a hadronic current $H^\mu$ \cite{9}. If the exchanged four-momentum $q = p - p'$ between the two mesons of momenta $p$ and $p'$ is much smaller than the mass $M_W$ of the $W^\pm$ boson, the amplitude can be written in the following way:

$$M = -i\frac{G_F}{\sqrt{2}}V_{Qq}L^\mu H^\mu,$$

where $G_F$ is the Fermi constant.

The leptonic current is given by:

$$L^\mu = \bar{\ell}\gamma^\mu(1 - \gamma_5)\nu,$$

while the hadronic current is:

$$H^\mu = \langle M_{q'q}(p')|j^H_\mu|M_{Qq}(p)\rangle,$$

where $j^H_\mu$ can be expressed in term of the Lorentz-invariant quantities that are combinations of $(p_\mu + p'_\mu)$, $q_\mu$ and $q^2$. Since the two mesons interact also strongly, the hadronic current will contain some terms that parameterize the non-perturbative behaviour of the QCD. These terms are called form factors, they depend on the given initial and final state and they are measurable experimentally.

Several methods exist to calculate these form factors, one of them is the Heavy Quark Effective Theory (HQET), described in section 1.4.1. Through this effective theory, the calculation of the branching ratio of $B^0 \rightarrow D^{*-}\ell^+\nu$ is performed in section 1.4.2.

1.4.1 The Heavy Quark Effective Theory

The Heavy Quark Effective Theory (HQET) is an effective theory designed to describe the strong interaction between a single heavy quark ($b$ or $c$) and a light quark \cite{10}. The starting point of HQET is that the mass of the heavy quark, $m_Q$, can be considered infinite (Isgur-Wise limit), so the heavy quark acts like a stationary point source of colour charge. For heavy hadrons, the HQET lagrangian is then constructed by expanding the QCD lagrangian in a power series of $1/m_Q$ and neglecting higher-order terms. This expansion holds as far as $m_Q \gg \Lambda_{QCD}$. 
1.4 The theory of semileptonic decays

HQET is particularly useful to calculate the form factors of decays involving \(B \rightarrow D\) transitions, because in the associated heavy quark transitions, due to the large quark masses, the velocity transferred between them is small, leaving the colour source stationary to a good approximation.

1.4.2 Standard Model expectation for \(B(B^0 \rightarrow D^{*-\tau^+\nu})/B(B^0 \rightarrow D^{*-\ell^+\nu})\)

Following [9], an effective weak hamiltonian can be written for \(b \rightarrow c\ell\nu\) transitions:

\[
H_{\text{eff}} = \frac{4G_F V_{cb}}{\sqrt{2}} \sum_{\ell=e,\mu,\tau} (\bar{\ell} \gamma^\mu P_L \nu_\ell) + \text{h.c.},
\]

(1.36)

where \(J_{bc,\mu}\) is the effective SM \(b \rightarrow c\) charged current:

\[
J_{bc,\mu} = c \gamma^\mu P_L b.
\]

(1.37)

From this hamiltonian the decay rate of \(B \rightarrow D^{*-\ell^+\nu}\) in function of \(q^2 = (p_B - p_{D^*})^2\) is calculated:

\[
\frac{d\Gamma}{dq^2} = \frac{G_F^2 |V_{cb}|^2 |p| q^2}{96\pi^3 m_B^2} \left( 1 - \frac{m_\ell^2}{q^2} \right)^2 \times \left[ (|H_{++}|^2 + |H_{--}|^2 + |H_{00}|^2) \left( 1 + \frac{m_\ell^2}{2q^2} \right) + \frac{3m_\ell^2}{2q^2} |H_{0t}|^2 \right],
\]

(1.38)

where \(p\) is is the three-momentum of the \(D^*\) meson in the \(B\) rest frame:

\[
|p| = \frac{\sqrt{\lambda(m_B^2, m_{D^*}^2, q^2)}}{2m_B},
\]

(1.39)

\[
\lambda(a, b, c) = a^2 + b^2 + c^2 - 2(ab + bc + ca).
\]

(1.40)

\(H_{mn}\) are the hadronic relevant helicity amplitudes:

\[
H_{\pm\pm}(q^2) = (m_B + m_{D^*}) A_1(q^2) \mp \frac{2m_B}{m_B + m_{D^*}} |p| V(q^2),
\]

(1.41a)

\[
H_{00}(q^2) = \frac{1}{2m_{D^*} \sqrt{q^2}} \times \left[ (m_B^2 - m_{D^*}^2 - q^2)(m_B + m_{D^*}) A_1(q^2) - \frac{4m_B^2 |p|^2}{m_B + m_{D^*}} A_2(q^2) \right],
\]

(1.41b)

\[
H_{0t}(q^2) = \frac{2m_B |p|}{\sqrt{q^2}} A_0(q^2),
\]

(1.41c)
where $A_{0,1,2}(q^2), V(q^2)$ are the form factors of this particular transition.

To calculate these form factors it is useful to define the kinematical variable:

$$w = v_B \cdot v_{D^*} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2m_Bm_{D^*}},$$

where $v_B, v_{D^*}$ are the four-velocities of $B$ and $D^*$. It is possible to define $A_{0,1,2}(q^2), V(q^2)$ in terms of four variables $h_{A_1}(w), R_{0,1,2}(w)$:

$$A_1(q^2) = h_{A_1}(w)\frac{1}{2}(w + 1)R,$$  \hspace{1cm} (1.43a)
$$A_0(q^2) = \frac{R_0(w)}{R} h_{A_1}(w),$$  \hspace{1cm} (1.43b)
$$A_2(q^2) = \frac{R_2(w)}{R} h_{A_1}(w),$$  \hspace{1cm} (1.43c)
$$V(q^2) = \frac{R_1(w)}{R} h_{A_1}(w),$$  \hspace{1cm} (1.43d)

where $R = 2\sqrt{m_Bm_{D^*}}/(m_B + m_{D^*})$. According to the HQET computation of Ref. [11], the $w$ dependence of these quantities is given by:

$$h_{A_1}(w) = h_{A_1}(1)[1 - 8\rho^2z + (53\rho^2 - 15)z^2 - (231\rho^2 - 91)z^3],$$  \hspace{1cm} (1.44a)
$$R_0(w) = R_0(1) - 0.11(w - 1) + 0.01(w - 1)^2,$$  \hspace{1cm} (1.44b)
$$R_2(w) = R_2(1) + 0.11(w - 1) - 0.06(w - 1)^2,$$  \hspace{1cm} (1.44c)
$$R_1(w) = R_1(1) - 0.12(w - 1) + 0.05(w - 1)^2,$$  \hspace{1cm} (1.44d)

with $z = (\sqrt{w + 1} - \sqrt{2})/\sqrt{w + 1 + \sqrt{2}}$. The parameters $\rho^2, h_{A_1}(1), R_1(1), R_2(1)$ can be measured from the $B \rightarrow D^*(\tau^- \bar{\nu}_\tau)$ decay distributions [12], while $R_0(1)$ can be computed using those measured values and some HQET calculations.

Due to the different lepton masses, the last term of Eq. (1.38) is relevant only for decays involving taus. Moreover, the space factor term $(1 - m_\ell^2/q^2)^2$ suppresses decays into taus with respect to decays to lighter leptons. The branching fraction of $B^0 \rightarrow D^{*-} \tau^+ \nu$ is obtained integrating (1.38) over $q^2$. To reduce most theoretical uncertainties due to the evaluation of the hadronic form factors, it is useful to normalize the branching fraction of this decay to the modes with a light lepton $\ell$ in place of tau in the final state. The interesting observable is therefore the ratio:

$$R_{D^*} \equiv \frac{B(B^0 \rightarrow D^{*-} \tau^+ \nu)}{B(B^0 \rightarrow D^{*-} \ell^+ \nu)}.$$  \hspace{1cm} (1.45)
1.5 Limitations of the Standard Model

In terms of differential decay rates this rate is equal to:

\[ R_{D^*}(q^2) = \frac{d\Gamma_\tau/dq^2}{d\Gamma_\ell/dq^2} = \left(1 - \frac{m_{\tau}^2}{q^2}\right)^2 \left[1 + \frac{m_{\tau}^2}{2q^2} + \frac{3m_{\tau}^2}{2q^2} \left|H_{00}\right|^2 + \frac{|H_{00}|^2}{\left|H_{++}\right|^2 + \left|H_{--}\right|^2 + \left|H_{00}\right|^2}\right] \]  

(1.46)

where \(d\Gamma_\ell/dq^2\) has been calculated in an analogous way to \(d\Gamma_\tau/dq^2\). The SM expectation for this ratio is:

\[ R_{D^*}^{SM} = 0.252 \pm 0.003. \]  

(1.47)

BaBar reported a measurement of \(R_{D^*}\), which combines both \(B^0\) and \(B^-\) modes, exceeding by 2.7\(\sigma\) the SM expectation [13]:

\[ R_{D^*} = 0.332 \pm 0.024 \pm 0.018. \]  

(1.48)

A similar argument can be made for \(B^0 \rightarrow D^- \tau^+\nu\), for which the SM gives:

\[ R_{D}^{SM} \equiv \frac{\mathcal{B}(B^0 \rightarrow D^- \tau^+\nu)}{\mathcal{B}(B^0 \rightarrow D^- \ell^+\nu)} = 0.279 \pm 0.017. \]  

(1.49)

The BaBar measurement of \(R_D = 0.440 \pm 0.058 \pm 0.042\), combined with the \(R_{D^*}\) measurement, gives a disagreement with the SM at 3.4\(\sigma\). In Fig. 1.1 the various measurements of \(R_{D^*}\) and \(R_D\) are shown.

This could be a signal for New Physics (NP) beyond SM, so a measurement of this ratio with improved precision is of paramount importance.

1.5 Limitations of the Standard Model

There are several observed phenomena not explained by the SM, of which the most important are matter-antimatter asymmetry, dark matter, dark energy, neutrino masses and how gravity can be included in the model. Therefore the SM is an incomplete theory, and new physics phenomena must exist at a certain energy scale. In fact the SM could be considered as the low-energy limit of a more general theory, or as the renormalizable part of an effective field theory valid up to some still undetermined cut-off scale \(\Lambda\).

If there is NP, it should involve also the Yukawa sector, in fact the hierarchy of the CKM elements could suggest the existence of a still unknown structure generating the Yukawa coupling parameters, and, since Yukawa couplings make the
three generations non-degenerate, the NP that determines their values would cause also the origin of the three-generation structure. The Yukawa couplings play a role in the quadratically divergent Higgs mass too: in fact, the physical Higgs mass is shifted by the interactions of the Higgs with the gauge bosons and the fermions, and the correction caused by the fermions grows quadratically with increasing energy; so at high energies a large cancellation between these corrections and the Higgs self-interaction term is needed to obtain the Higgs mass experimentally observed. In order to avoid this large cancellation, NP should exist at TeV scale.

To give an idea of the reach in the NP scale from indirect searches in the flavour sector, one might follow Ref. [14], where processes with flavour violation $\Delta F = 2$ are examined. In general, for these processes one can expand the NP effective hamiltonian as a sum of operators which are multiplied by coefficients:

$$C_i(\Lambda) = \frac{F_i L_i}{\Lambda^2},$$

where $F_i$ are functions of the NP flavour couplings, $L_i$ are loop factors present in models without tree-level FCNC, and $\Lambda$ is the energy scale of NP. For a generic strongly-interacting theory with arbitrary flavour structure, one expects $F_i \sim L_i \sim 1$, so studying these processes allows to measure $C_i(\Lambda)$, putting constraints on $\Lambda$ [14], which turns out to be at the $10^{4-5}$ TeV level, much higher than what can be

![Figure 1.1: Comparison of the various measurements of $R_D$ with statistical and total uncertainties. The vertical bands represent the average of the measurements done before BaBar 2012 (light shading) and SM predictions (dark shading), for $R_D$ and $R_{D^*}$. The widths of the bands represents the uncertainties.](image-url)
reasonably obtained at present and future direct search facilities. These constraints decrease somewhat, down to the TeV range or below, if specific flavour structures or loop-suppression factors are taken into account.

1.6 New Physics in \( B^0 \to D^{*-} \tau^+ \nu \)

It has already been shown in section 1.4.2 that the experimental measurements of \( \mathcal{B} \) \((B^0 \to D^{*-} \tau^+ \nu)\) disagree with the SM expectation, and this means that NP with a different structure of the relevant weak charged interactions could be involved in this process. The \( B \) meson decays with tau lepton in the final state can show significant NP contributions not present in processes with light leptons, because large tau mass can reduce the helicity suppression of certain semileptonic decay amplitudes which are unobservable in decays with light leptons in the final state [9].

In general, to have NP leading to charged lepton helicity suppression in \( B^0 \to D^{*-} \tau^+ \nu \), one can add a total derivative of a scalar operator to the effective current \( J_{bc}^\mu \); so (1.37) becomes [9]:

\[
J_{bc,\mu} = \bar{\tau} \gamma^\mu P_L b + g_{SL} i \partial^\mu (\bar{\tau} P_L b) + g_{SR} i \partial^\mu (\bar{\tau} P_R b).
\]  

(1.51)

Only the \( H_{0t} \) helicity amplitude is affected by these NP quark charged currents, becoming:

\[
H_{0t} = H_{0t}^{SM} \left[ 1 + (g_{SR} - g_{SL}) \frac{q^2}{m_b + m_c} \right],
\]

(1.52)

where \( H_{0t}^{SM} \) is the SM amplitude defined in (1.41c). Finally, the ratio becomes [13]:

\[
R_{D^*} = R_{D^*}^{SM} + A_{D^*} \text{Re}(g_{SR} - g_{SL}) + B_{D^*} |g_{SR} - g_{SL}|^2.
\]

(1.53)

Similarly, for \( B^0 \to D^- \ell^+ \nu \) decays:

\[
R_D = R_D^{SM} + A_D \text{Re}(g_{SR} - g_{SL}) + B_D |g_{SR} - g_{SL}|^2.
\]

(1.54)

Therefore a measurement of \( R_{D^*} \) and \( R_D \) provide constraints on NP parameters.

There are several NP candidate models which can be tested in \( B^0 \to D^{*-} \tau^+ \nu \), such as the two Higgs doublet model of type II (2HDM-II), 2HDM of type III, the leptoquark model, models with new tensor mediators or coloured scalars. In a 2HDM of type II there is one Higgs doublet which couples to down quarks and charged leptons, while another one gives masses to the up quarks. In this model a
charged Higgs, in addition to a $W$ boson, would mediate the interaction between the hadronic and the leptonic currents (see Fig. 1.2). The difference between a 2HDM-II and a 2HDM-III is that in the latter both Higgs doublets couple to up quarks and down quarks as well. While in the 2HDM-III framework the $g_{SL}$ and $g_{SR}$ parameters are not constrained among each other, in 2HDM-II they are predicted to be

$$g_{SR} = -\frac{m_b m_\tau \tan^2 \beta}{m_{H^+}},$$

$$g_{SL} = 0,$$

(1.55a)

(1.55b)

where $\tan \beta$ and $m_{H^+}$ are respectively the ratio of the vacuum expectation values of the two Higgs doublets and the mass of the charged Higgs. The remaining models are based on new kinds of particles, like leptoquarks (scalar and vector bosons which decay into a quark and a lepton), tensor operators or coloured scalars (weak doublets with fractional electric charges). Another important NP model is the Minimal Supersymmetric Model (MSSM), according to which fermions have bosonic superpartners (called sfermions), and bosons have fermionic superpartners (called bosinos).

The BaBar result on $R_D$ and $R_{D^*}$ show that the two Higgs doublet model of type II (2HDM-II) is excluded at 99.8% confidence level [13], while in [15] it is demonstrated that is still possible to explain the observed branching fraction of $B \to D^* \tau \nu$ and $B \to D \tau \nu$ within the a 2HDM-III framework, although a significant portion of the parameters space is excluded by the analysis of $q^2$ spectra. The excess measured by BaBar can be explained also by models with leptoquarks [16], tensors mediators [17] and coloured scalars [18]. Indeed, the MSSM is not able to
accommodate this excess [19].
Theory of flavour violation in semileptonic decays
Chapter 2

The LHCb experiment

The LHCb (Large Hadron Collider beauty) detector is one of the four big experiments located at the Large Hadron Collider (LHC) of CERN. This experiment is mainly dedicated to the study of flavour physics, more specifically the physics of beauty and charmed hadrons. The main purpose of LHCb is the search for physics processes beyond SM through the study of $CP$ violation in $b$ or $c$ decays and the measurement of rare processes.

LHCb is a single-arm spectrometer with a forward angular coverage of $[10,250]$ mrad in the vertical plane and $[10,300]$ mrad in the horizontal plane [20]. This particular geometry has been chosen because, in $pp$ deep inelastic scattering, the most part of $b\bar{b}$ pairs come from virtual gluons that are highly boosted, so the $b$-hadrons coming from the hadronization of these pairs lie in the same forward (or backward) cone. The acceptance can be expressed also in terms of pseudorapidity $\eta$, that is defined in function of $\theta$, the polar angle with respect to the beam axis (taken as axis $z$):

$$\eta = - \log \tan \frac{\theta}{2}.$$  \hspace{1cm} (2.1)

The LHCb acceptance in terms of pseudorapidity is $1.6 \leq \eta \leq 4.9$.

The instantaneous luminosity of LHCb is kept constant during LHC running by constantly adjusting the optics at the interaction point. The experiment ran at lower luminosity with respect to the achievable one at LHC, to keep a low pile-up and to reduce radiation damage. The instantaneous luminosity was set to $2 \cdot 10^{32}$ cm$^{-2}$s$^{-1}$ in 2011 and $4 \cdot 10^{32}$ cm$^{-2}$s$^{-1}$ in 2012. The measured value of cross section of $b\bar{b}$ production at a center-of-mass-energy of 7 TeV is $\sigma_{b\bar{b}} = (284 \pm 20 \pm 49)$ µb [21].

In Fig. 2.1 there is an overview of the apparatus and its sub-detectors, which will be described in next sections.
Figure 2.1: The LHCb apparatus. Starting from left, the following sub-components are visible: Vertex Locator (VeLo), Ring Imaging CHERENkov detector RICH-1, dipole magnet, tracking system (TT, T1, T2 and T3 stations), RICH-2, Electromagnetic CALorimeter (ECAL), Hadronic CALorimeter (HCAL), muon detector (M1, M2, M3, M4 and M5 stations).

2.1 The Magnet

To measure the momentum of the charged particles, a dipole magnet is used in LHCb. The magnet, located after the VELO and in between the two RICH detectors, is formed by two identical saddle-shaped coils, bent at 45° on the two transverse sides and mounted inside an iron yoke (see Fig. 2.2). The magnet gap is shaped to follow the detector acceptance.

The total weight of the yoke and the coils is 1500 tons and 54 tons respectively, and the maximum current in the coils is 6.6 kA.

The magnetic field has a vertical direction, so it deflects charged particles in the horizontal plane \((x, z)\). The strength of the field is variable along the \(z\)-axis (see Fig. 2.3), in order to be less than 2 mT in the region occupied by the RICH detectors and the highest possible between the tracking stations placed before and after the magnet. The integrated magnetic field is equal to 4 Tm for tracks of 10 m.
2.2 The Tracking System

The LHCb Tracking System (TS) is designed to reconstruct tracks of charged particles and measure their momentum. The TS consists of the VErrtex LOcator system (VELO) and four planar tracking stations: the Tracker Turicensis (TT) upstream of the dipole magnet and three stations T1-T3 downstream of the magnet. Each of the stations T1-T3 is divided in two parts, the Inner Tracker (IT), which together with the TT forms the Silicon Tracker (ST), and the Outer Tracker (OT).

In the following subsections all these detectors are briefly described.

2.2.1 The VErrtex LOcator

The VErrtex LOcator (VELO) is the LHCb subdetector placed closest to the interaction point \[22\]. Its purpose is to measure with a very high resolution the trajectories of particles, in order to reconstruct the primary and secondary vertices.

The most important information concerning the performance of the VELO are the following:

The direction of the \(B\) field is periodically reversed in order to study systematic effects depending on particle charge.
• Spatial resolution on primary vertex position of 42\,\mu m along the $z$-axis and 10\,\mu m in the transverse plane.

• Charged tracks impact parameter (with respect to PV) resolution of 20\,\mu m.

• Resolution on decay length of heavy flavoured hadrons ranging from 220\,\mu m to 370\,\mu m.

The VELO has also a pile-up veto counter to suppress events with more than 2 proton interactions per bunch crossing.

This sub-detector uses radiation-tolerant silicon strip sensors, 300\,\mu m thick, divided in 25 stations placed orthogonally to the beam. Each station, split in a lower and an upper part, consists of two types of modules, one measuring the radial coordinate, the other measuring the azimuthal angle of the charged tracks. The two parts can be moved farther from the beam during LHC injection, and closer to the beam during collisions.

The modules are placed in a vacuum vessel which is separated from the beam vacuum by a thin aluminum sheet.

### 2.2.2 The Silicon Tracker

The Silicon Tracker (ST) is composed by two different detectors that use silicon microstrip sensors: the Tracker Turicensis or Trigger Tracker (TT) and the Inner Tracker (IT).

The TT (shown in Fig. 2.6) is a planar tracking station with four detection layers made of 500\,\mu m thick silicon microstrips organized into either two or three readout sectors [23]. In the first and fourth layers the microstrips are parallel to the vertical axis ($y$-axis), while in the second and third layers the strips are rotated of $-5^\circ$ and $5^\circ$, to reconstruct the tracks in three dimensional space.

The IT is composed by the three inner parts of the stations T1, T2 and T3. Each IT station, arranged around the beampipe as shown in Fig. 2.7 and 2.8, consists of four individual detector boxes, each of which contains four detection layers. Each detection layer is composed by seven detector modules, and each module is formed by one or two silicon sensors and a readout hybrid.
2.2 The Tracking System

Figure 2.4: Cross section in the \((x, z)\) plane of the VELO silicon sensors, at \(y = 0\), with the detector in the fully closed position. The front face of the first modules is also illustrated in both the closed and open positions. The two pile-up veto stations are located upstream of the VELO sensors.

Figure 2.5: Overview of the VELO vacuum vessel.
Figure 2.6: Layout of the four detection layers of the TT station.

Figure 2.7: View of the four IT detector boxes arranged around the LHC beampipe.

Figure 2.8: Layout of a detection layer in the second IT station.
2.3 Particle Identification

2.2.3 The Outer Tracker

The Outer Tracker (OT) is a drift-time detector designed for charged particles tracking and momentum measurement over a large acceptance area. It is divided among the three outer parts of the stations T1, T2 and T3, and each station consists of four layers which are arranged with the same geometry of the TT (see section 2.2.2). The OT is designed as an array of individual, gas-tight straw-tube modules, each of which contains two layers of drift-tubes with inner diameters of 4.9 mm. The tubes are filled with a mixture of Argon (70%) and CO$_2$ (30%), which allows to get a fast drift time (below 50 ns), and a sufficient drift-coordinate resolution of 200 mm.

2.3 Particle Identification

2.3.1 The Ring Imaging CHERenkov detectors

A fundamental role in Particle Identification is played by the two Ring Imaging CHERenkov (RICH) detectors, one (RICH 1) located between the VELO and the TT, the other (RICH 2) between T3 and the first station of the muon detector. The working principle of a Cherenkov detector is Cherenkov radiation, that is emitted by a charged particle when it travels through a medium with a velocity greater than
The RICH 1 covers the low momentum charged particle range $[1, 60] \text{ GeV/c}$, while the RICH 2 covers the high momentum range from $15 \text{ GeV/c}$ to beyond $100 \text{ GeV/c}$. While the former covers the full LHCb acceptance, the latter has a limited angular acceptance of $[15, 120] \text{ mrad}$ in the horizontal plane and up to $100 \text{ mrad}$ in the vertical plane.

The RICH 1 uses aerogel and fluorobutane ($\text{C}_4\text{F}_{10}$) as gas radiators, while the RICH 2 uses the tetrafluoromethane ($\text{CF}_4$). In both detectors a combination of spherical and flat mirrors reflects the Cherenkov light, taking the image on the Hybrid Photon Detectors (HPD) that are out of the spectrometer acceptance.

The HPD are vacuum photon detectors in which a photoelectron, released from the conversion of an incident photon in a photocathode, is accelerated by an applied high voltage of typically 10 to 20 kV onto a reverse-biased silicon detector. The arrival of the photoelectron generates electron-hole pairs at an average yield of one for every $3.6 \text{ eV}$ of deposited energy. Finally, this current is read with very high efficiency by a carefully-designed readout electronics, that is able to detect single photoelectrons.

Schematic views of the RICH detectors are shown in Fig. 2.10 and 2.11.
2.3.2 The Calorimeters

The LHCb calorimeter system, placed between the stations M1 and M2 of the muon detector, selects and identifies hadron, electron and photon candidates for the Level 0 trigger (L0), measuring their transverse momentum, energy and position. A classical structure of an Electromagnetic CALorimeter (ECAL) followed by a Hadron CALorimeter (HCAL) has been adopted, with the addition of a Scintillator Pad Detector (SPD) to identify charged particles and a PreShower (PS) detector to make the electrons interact and generate a further cascade.

Each calorimeter is made of alternating layers of active (i.e. detecting) scintillating material and absorbing material. Because of the absorbing material, the particles lose energy via electromagnetic or hadronic cascades, which excite the molecules of the scintillating material, that emit ultraviolet radiation in an amount proportional to the energy of the particle. Then WaveLength Shifting (WLS) fibers carry the light to some photomultipliers, and, from the amount of light collected, the energy lost by the particle can be finally measured.

Since the hit density varies by two orders of magnitude over the calorimeter surface, a variable lateral segmentation in cells has been adopted. ECAL and SPD/PS have been divided into three different sections, HCAL in two; in each section the cell dimensions are different with respect to the other sections (the inner sections have smaller cells).

The SPD and the PS are the first calorimeters with which the particles interact. Their purpose is to improve spatial and energy resolutions of electromagnetic showers. The SPD/PS system consists of a 15 mm lead converter sandwiched between two identical planes of rectangular scintillator pads of high granularity (12032 readout channels). The sensitive area of this calorimeter is $7.6 \times 6.2 \text{ m}^2$.

The ECAL is dedicated to the measurement of electromagnetically-interacting particles, in particular electrons, photons and neutral pions. The structure of the ECAL is given by alternating 2 mm thick lead layers and 4 mm thick scintillating tiles. In the ECAL the inner section has cells of $4.04 \times 4.04 \text{ cm}^2$, the middle one $6.06 \times 6.06 \text{ cm}^2$ and the outer one $12.12 \times 12.12 \text{ cm}^2$. The energy resolution of the ECAL is given by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} + b,$$  

(2.2)

where $E$ is expressed in GeV, $8.5\% < a < 9.5\%$ and $b \simeq 0.8\%$. The total area of this detector is $7.76 \times 6.30 \text{ m}^2$.

The HCAL is optimized for the measurement of strongly-interacting particles.
Figure 2.12: Front view of one half of the SPD/PS.

Figure 2.13: Individual scintillator pad with the WLS fibre.

Figure 2.14: Downstream view of the ECAL installed but not completely closed.

Figure 2.15: Outer, middle and inner type ECAL modules.
2.3 Particle Identification

The structure of the HCAL is analogous to the one of the ECAL: iron spacers are alternated with scintillating material tiles, and in the longitudinal direction the length of tiles and iron spacers corresponds to the hadron interaction length in steel. The HCAL inner section has cells of $13.13 \times 13.13 \text{ cm}^2$, while the outer one $26.26 \times 26.26 \text{ cm}^2$.

The resolution of this calorimeter is:

$$\frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%,$$

with $E$ is expressed in GeV.

### 2.3.3 The Muon System

Muon triggering and identification are fundamental tasks in LHCb, because they are present in many sensitive $CP$-violating decays, such as the “golden modes” $B^0 \rightarrow J/\psi (\mu^+\mu^-)K^0_s$ and $B^0_s \rightarrow J/\psi (\mu^+\mu^-)\phi$, or in rare decays like $B^0_s \rightarrow \mu^+\mu^-$ which could be sensitive to New Physics. Muons are also used to tag the initial state flavour of the $B$ meson.

The muon system consists of five rectangular stations M1-M5; M1 is located upstream the calorimeters, while the other four stations are downstream them. M2-M5 are separated from each other by 80 cm thick iron absorbers to select the...
most penetrating muons. The minimum value of momentum needed to cross all the stations is 6 GeV/c. Every station is divided in four concentric regions (R1-R4), whose linear dimensions and segmentations scale as 1:2:4:8; in this way the particle flux is expected to be about the same over the four regions of a given station.

The first three stations have a high spatial resolution along the bending plane, to define the track direction and to calculate the $p_T$ of the candidate muon with a resolution of 20%. On the contrary M4 and M5 have a limited spatial resolution, because their main purpose is the identification of penetrating particles.

The muon system consists of Multi Wire Proportional Chambers (MWPC), except for region R1 of the station M1, where there are 12 chambers composed by two triple GEM detectors superimposed that can sustain a rate capability up to $500 \text{kHz/ cm}^2$ of charged particles, with a time resolution better than 3 ns. The GEM performance in terms of radiation hardness is such that ageing effects in 10 years of data taking are negligible. The MWPC are 1368, they have a gap of 5 mm filled with a combination of carbon dioxide, argon and tetrafluoromethane and a wire plane of 2 mm spacing placed in the middle of the gap. They have a time resolution of about 5 ns.

Every station is divided in logical pads that define the spatial resolution along $x$ and $y$, as shown in Fig. 2.19.

## 2.4 The Trigger System

The LHCb trigger is a system which allows to skim the interactions seen by the detector, before memorizing data for the analyses, reducing the rate from the bunch crossing one (40 MHz) to 2 kHz. This complex task is performed in two steps: the Level 0 (L0) and the High Level Trigger (HLT). The former utilizes custom-made electronics on the detectors, while the latter consists of software selections which run on a dedicated computing farm.

The task of the L0 trigger is to reduce the event rate from 40 MHz to 1 MHz. In order to do that, it selects the highest $E_T$ hadrons, electrons and photons in the calorimeters and the highest $p_T$ muons in the muon chambers: due to the large mass of heavy-flavoured hadrons, their decay products have high $p_T$ and $E_T$. The L0 trigger also uses information coming from the pile-up counter in the VELO, in order to reject events with many primary interactions.

The HLT further reduces the rate from 1 MHz to 2 kHz selecting the events of physical interest, by using the information coming from all sub-detectors. This
2.4 The Trigger System

Figure 2.18: Schematic side view of the muon system.

Figure 2.19: Left: front view of a quadrant of a muon station. Each rectangle represents one chamber. Each station contains 276 chambers. Right: division into logical pads of four chambers belonging to the four regions of station M1.
selection is performed by about 2000 CPUs located in the Event Filter Farm (EFF). The HLT is divided into two stages, HLT1 and HLT2:

- The HLT1 confirms or rejects L0 outputs. It consists of alleys, each of which is associated to a specific output type of L0 (e.g. muons, dimuons, electrons, hadrons and so on). Each alley uses different algorithms that combine information coming mostly from the VELO and the Tracking system.

- In the HLT2 stage two parallel selections are performed using cuts on invariant masses or pointing the B momentum towards the primary vertex: an inclusive selection, that selects partially reconstructed B decays, and an exclusive selection. The HLT2 output is the OR of these two selections.

A scheme which summarize the trigger flow is shown in Fig. 2.20.

2.5 Detector performance

2.5.1 Track reconstruction

The reconstruction of charged particle tracks starts with the search of seeds, that are hits in the VELO and the T stations and represent the initial track candidates [24]. Then the seeds are combined with hits in other detectors of the Tracking System, leading to the identification of tracks. Tracks are fitted with a pattern recognition algorithm (Kalman fitting) in order to account for multiple scattering and correct for energy losses due to ionization. After the fit, the reconstructed track is represented by state vectors \((x, y, dx/dz, dy/dz, \text{momentum } p)\), which are specified at given \(z\)-positions in the experiment.
2.5 Detector performance

Figure 2.21: Relative momentum resolution versus momentum for tracks traversing all detector components up to the muon stations.

The tracking efficiency is defined as the probability that the trajectory of a charged particle that has passed through the full tracking system is reconstructed. The efficiency has been measured using $J/\psi \rightarrow \mu^+ \mu^-$ decays: its average value is above 96\% in the momentum range $5 \text{ GeV}/c < p < 200 \text{ GeV}/c$ and in the pseudo-rapidity range $2 < \eta < 5$.

The momentum resolution has been measured studying $J/\psi \rightarrow \mu^+ \mu^-$ too. It is about 0.5\% for particles momenta below 20 GeV/c and about 0.8\% for particles momenta around 100 GeV/c. The momentum resolution $\delta p/p$ as function of the momentum $p$ is represented in Fig. 2.21.

2.5.2 Particle Identification

The combined information from the two RICH detectors, the calorimeters and the muon system allows to identify charged particles ($e$, $\mu$, $K$, $\pi$, $p$).

Muons are identified extrapolating their tracks from the hits in the muon stations. Electrons are identified matching the track momentum and the energy of the clusters in the ECAL and detecting separate clusters due to bremsstrahlung photons emitted by the electrons before the magnet. Neutral pions and photons identified by looking for clusters in the ECAL.
An overall particle likelihood is computed under a set of hypotheses. Since the pions are the most abundant particles, the first assumption is that all particles are pions. Then for each track in turn, the mass hypothesis is varied according to $e$, $\mu$, $K$, $\pi$, $p$ mass hypotheses, and the log-likelihood is computed again. For every particle a PID variable is defined: PID$^X$ of a particle is the difference between the log-likelihoods of the particle in the $X$ and in pion hypotheses respectively, where $X$ can be $e$, $\mu$, $K$ or $p$.

Figure 2.22 shows the kaon efficiency and the probability of misidentifying a pion as a kaon as function of the particle momentum. In this plot two different cuts on PID have been applied, in order to show how a PID cut can affect the misidentification rate and the efficiency.

Another variable useful to suppress misidentification is $ProbNNp\pi(K,...)$, that is the probability, computed through a MC-trained neural network, for a detected particle to be a pion, a kaon, and so on. $ProbNNp\pi$ will be used in the first cut-based selection, which will be described in Chapter 3.
Chapter 3

Events selection

This thesis is a feasibility study for the measurement of $B(B^0 \rightarrow D^{*-} \tau^+ \nu_{\tau})$ where the tau decays hadronically with three charged particles in the final state, so $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ (\pi^0) \nu_{\tau}$. The main advantage in studying this three-prong decay in comparison with a one-prong is that the tau decay vertex can be reconstructed with a good precision, which is very important to suppress backgrounds due to hadronic $B$ decays. Hadronic tau decays have also the advantage of a single neutrino in the final state, so the kinematics of the tau decay can be closed if the tau direction is known. Section 3.1 presents the definition of the signal and of the main background sources affecting the final state of this decay, and the strategy chosen to reject this background is explained. A description of the data and MonteCarlo samples used in this study is given in section 3.2, while in section 3.3 the first set of cuts used to suppress the most dominant background is illustrated.

3.1 Analysis strategy

The final aim of this analysis is to measure the ratio:

$$\frac{B(B^0 \rightarrow D^{*-} \tau^+ \nu_{\tau})}{B(B^0 \rightarrow D^{*-} \pi^+ \pi^- \pi^+)}$$

(3.1)

where the numerator is the branching fraction of the decay studied, while the denominator is the normalization channel, which is also the main background to be suppressed when isolating signal decays.

\footnote{Charge conjugated decay modes will be implied from now on}
The signal chosen for this study is defined according to the following decay chain:

\[ B^0 \rightarrow D^{*-} \tau^+ \nu_\tau, \]
\[ D^{*-} \rightarrow \pi^- D^0, \]
\[ D^0 \rightarrow K^+ \pi^-, \]
\[ \tau^+ \rightarrow \pi^+ \pi^- \pi^+ (\pi^0) \nu_\tau. \] \hfill (3.2)

In the normalization channel the same \( D^{*-} \) decay chain is required.

The visible final state for signal consists of the three pions coming from the tau and two pions plus a kaon from the \( D^{*-} \) decay chain, while the neutrinos are of course not detected. The tau decay vertex is reconstructed through a fit of the three pion tracks and the momentum assigned to it is the sum of the three pion momenta. Because of the presence of neutrino, the momentum and energy of the tau is not reconstructed. In an analogous way the \( D^0 \), the \( D^{*-} \) and the \( B^0 \) are reconstructed. Due to the small \( q \)-value of the \( D^{*-} \) decay into \( D^0 \pi^- \), the pion from the \( D^* \) has a low momentum and will be parallel to the \( D^0 \). Therefore, the reconstruction of \( D^{*-} \) vertex has not a good quality. Moreover, because of its negligible lifetime, its decay vertex will coincide with the \( B^0 \) vertex.

The number of reconstructed signal events after the selection will be equal to:

\[ N_{SIG} = L \cdot \sigma(pp \rightarrow b\bar{b}) \cdot \text{prob}(b \rightarrow B^0) \cdot \mathcal{B}(B^0 \rightarrow D^{*-} \tau^+ \nu_\tau) \cdot \mathcal{B}(D^{*-} \rightarrow \pi^- D^0) \cdot \mathcal{B}(D^0 \rightarrow K^+ \pi^-) \cdot \mathcal{B}(\tau^+ \rightarrow \pi^+ \pi^- \pi^+ (\pi^0) \nu_\tau) \cdot \varepsilon_{SIG}, \] \hfill (3.3)

where \( L \) is the integrated luminosity, \( \sigma(pp \rightarrow b\bar{b}) \) is the \( b\bar{b} \) production cross section, \( \text{prob}(b \rightarrow B) \) is the probability that \( b \) hadronizes into a \( B^0 \) meson. \( \varepsilon_{SIG} \) is the selection efficiency, that is evaluated on MonteCarlo (MC):

\[ \varepsilon_{SIG} = \frac{N_{MC}^{SIG}}{N_{GEN}^{SIG}}, \] \hfill (3.4)

with \( N_{GEN}^{SIG} \) equal to the number of MC signal generated events and \( N_{MC}^{SIG} \) equal to the number of MC signal events reconstructed by trigger and selection criteria. In an analogous way, the number of reconstructed events for the normalization channel is:

\[ N_{NORM} = L \cdot \sigma(pp \rightarrow b\bar{b}) \cdot \text{prob}(b \rightarrow B^0) \cdot \mathcal{B}(B^0 \rightarrow D^{*-} \pi^- \pi^- \pi^+) \cdot \mathcal{B}(D^{*-} \rightarrow \pi^- D^0) \cdot \mathcal{B}(D^0 \rightarrow K^+ \pi^-) \cdot \varepsilon_{NORM}. \] \hfill (3.5)
3.1 Analysis strategy

Table 3.1: Table of branching ratios for signal channel and background channels with the same final state. For the signal $B$ and its error the SM prediction was used, for the other channels PDG values were used \[7\]. The last three background channels have two more tracks in the final state, and they represent a background when these extra tracks are not detected. Through vertex isolation criteria, the background with extra tracks can be suppressed. The $B$ of the $D^+_s$ decays reported in this table are related to final states with at most three charged pions coming from $D^+_s$, and their calculation is described more in details in section 4.1.3.

<table>
<thead>
<tr>
<th>Signal channel</th>
<th>$\mathcal{B}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow D^{*-}\pi^+(-\rightarrow \pi^+\pi^-\pi^+(\pi^0)\nu_\tau)$</td>
<td>0.173 ± 0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Background channels</th>
<th>$\mathcal{B}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow D^{*-}\pi^+\pi^+\pi^-(\pi^0)$</td>
<td>2.5 ± 0.3</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{<em>-}D^</em>_s^+/D^{*+}+(2317)/D^+_s(2457)(\rightarrow \pi^+\pi^-\pi^+)$</td>
<td>0.040 ± 0.004</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{<em>-}D^</em>_s^+/D^{*+}+(2317)/D^+_s(2457)(\rightarrow \pi^+\pi^-\pi^+\pi^-)$</td>
<td>0.56 ± 0.06</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}K^0D^+(\rightarrow D^+(\rightarrow \pi^+\pi^-\pi^+(\pi^0))\gamma/\pi^0)$</td>
<td>0.0050 ± 0.0005</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}K^0D^+(\rightarrow \pi^+\pi^-\pi^+(N))$</td>
<td>0.0123 ± 0.0011</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}D^+(\rightarrow \pi^+\pi^-\pi^+(N))$</td>
<td>0.0012 ± 0.0001</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}K^0D^+(\rightarrow D^+(\rightarrow K^-\pi^+\pi^-\pi^+(\pi^0))\pi^+)$</td>
<td>0.067 ± 0.006</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}K^0D^0(\rightarrow D^+(\rightarrow K^-\pi^+\pi^-\pi^+(\pi^0))\gamma/\pi^0)$</td>
<td>0.130 ± 0.012</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}K^+D^0(\rightarrow K^-\pi^+\pi^-\pi^+(\pi^0))$</td>
<td>0.030 ± 0.003</td>
</tr>
</tbody>
</table>

Using (3.3) and (3.5), the ratio (3.1) can be written as:

$$\frac{\mathcal{B}(B^0 \rightarrow D^{*-}\pi^+\nu_\tau)}{\mathcal{B}(B^0 \rightarrow D^{*-}\pi^+\pi^-\pi^+)} = \frac{N_{SIG}}{N_{NORM}} \frac{\varepsilon_{NOM}}{\varepsilon_{SIG}} \frac{1}{\mathcal{B}(\pi^+ \rightarrow \pi^+\pi^-\pi^+\nu_\tau)}.$$  (3.6)

It is clear from (3.6) that measuring a ratio of branching ratios is useful to cancel theoretical and experimental uncertainties on luminosity, cross sections and hadronization probability.

There are several decays with the same visible final state of the signal, they are listed in Tab. 3.1 together with their measured or expected branching fractions.

The most dominant background source is $B^0 \rightarrow D^{*-}\pi^+\pi^-\pi^+(\pi^0)$, in fact it has a branching ratio about 14 times larger than the signal. The method, recently proposed by Guy Wormser \[26\], used to remove this kind of background is the “vertex inversion”, i.e. the requirement that the tau decay vertex must lay downstream with respect to the $B^0$ vertex by a significance of at least 5σ along the $z$ axis, to be sure that the three pions do not come directly from the $B^0$ vertex. This method was not achievable in $B$-factories, since there the $B$ mesons are not boosted enough to generate daughter particles with large flight lengths, which are needed to resolve well
the vertices along a certain direction. Since the tau and $D^0$ have a similar lifetime ($(290.3\pm0.5)\text{ fs}$ for tau and $(410.1\pm1.5)\text{ fs}$ for $D^0$), this procedure keeps a reasonably high signal efficiency, while suppressing the $B^0 \to D^{*-}\pi^-\pi^+\pi^+$ background almost completely. Moreover, the vertex inversion requirement can be reversed thus giving a sample of $B^0 \to D^{*-}\pi^-\pi^+\pi^+$ decay, which can be used to obtain the yield of the normalization sample.

Together with the inversion cut, there is a set of cuts which are applied in order to increase signal purity. These cuts are discussed in section 3.3.

Once the events where the three pions come directly from the $B^0$ vertex are removed, the only remaining physics background is due to decays where the three pions originate from a particle with a detectable lifetime. The most important contribution to these decays is given by $B^0 \to D^{*-}D^{s(*)}_s$, with $D^{s(*)}_s \to (\pi^+\pi^-\pi^+\pi^-\pi^+\pi^+\pi^-\pi^+\pi^+\pi^-\pi^+\pi^+(N))$ (from now on $D^+_s$ means also $D^{(*)}_s$, $D^{(*)}_s(2317)$ or $D^{(*)}_s(2457)$). This kind of background is treated using a partial reconstruction technique that should give meaningful results for this background but not for the signal. Variables related to this partial reconstruction are calculated and used as input variables in a MultiVariate Analysis (MVA) to separate signal from $B^0 \to D^{*-}D^+_s$ decay. The events with $B^0 \to D^{*-}D^+$, where $D^+ \to \pi^+\pi^-\pi^+\pi^-\pi^+\pi^+\pi^-\pi^+\pi^+\pi^-\pi^+\pi^+(N)$, are treated in the same way, since the topology of this decay is similar to the one involving $D^+_s$ mesons. The partial reconstruction technique and the MVA application are explained in section 4.2.

For what concerns the other $B$ decays resulting in a $D^{*-}$, a $D^+ \to \pi^+\pi^-\pi^+\pi^-\pi^+\pi^+\pi^-\pi^+\pi^+\pi^-\pi^+\pi^+(N)$, and an additional neutral kaon in the final state, the partial reconstruction may fail, so their contribution is estimated by looking at the MC. The MC is used to check also the contribution of all other background decays.

The last three decays reported in Tab. 3.1 result in other charged track from $D^0$ decays, in addition to three pions. This kind of background should be suppressed by a factor of about 4. This factor 4 comes from a crude estimation based on the fact that the track reconstruction efficiency for a track in the acceptance is around 75%, so this extra track is not seen $\sim 1/4$ of the times. This factor has to be measured more precisely by looking at MC where an extra track is present either at the $B$ vertex or at the $3\pi$ vertex.

If extra tracks are seen, vertex isolation criteria can be applied, by which additional tracks are added to the $B$ vertex to see if the quality of the vertex (measured by a $\chi^2$) improves. If the vertex quality improves, it means that those added tracks are probably coming from that vertex, and they were missed during reconstruction.

There are other kinds of background, which are non-physical:
3.2 Data and MonteCarlo samples

- combinatorial: events where one or more particles (especially pions) coming from other processes are wrongly associated with the reconstructed $D^{*-}$ decay. The effect of this background is to introduce a broad and smooth component in the distribution of $M(D^{*-}) - M(D^0)$, so the high-mass sideband of this distribution is used to evaluate and subtract the contribution of this background.

- misidentification: events where one or more particles (especially kaons) are seen as pions. Because of this misidentification the decay $B^0 \rightarrow D^{*-}D^+(\rightarrow \pi^+K^-\pi^+(N))$ can affect the signal when the $K^-$ is seen as a $\pi^-$. A way to estimate the contribution of this background is to assign the kaon mass to the pion with opposite charge with respect to the $D^+$ and look if there is a peak in the mass distribution of the three charged hadrons.

The goal of this thesis is to select signal in the data sample with a purity of 60% and a statistical uncertainty of about 3%, which is much less than the statistical uncertainty of 7.2% obtained by BaBar. This is the first time that a very pure tau sample is isolated for a NP study, and this method will be useful for further searches of NP effects, beyond the scope of this analysis. Since the decay $\tau^+ \rightarrow a_1^{+}\nu_\tau$, followed by $a_1^{+} \rightarrow \pi^{+}\pi^{+}\pi^{+}$ [27], the invariant mass distribution of the three pions for signal is expected to follow the $a_1^+$ lineshape. This is not true for background events where the three pions originate from a $D_s^+$ or a $D^+$ decay chain.

3.2 Data and MonteCarlo samples

The data sample used in this study has been collected by LHCb in 2011 at a center of mass energy of 7 TeV, and it corresponds to an integrated luminosity of 1 fb$^{-1}$.

The MonteCarlo (MC) samples are generated using PYTHIA software [28] to simulate $pp$ collision and hadronization, and EvtGen software [29] for particle decays. The MC samples used in this study are reported in Tab. 3.2. These samples are used to evaluate selection efficiencies, to train the MVA and to evaluate background contributions in data. In the last MC sample the interaction of the particles with the detector has been not simulated because of technical problems, therefore only quantities at the generation level were available. This $MC$-generator contains a mixture of $B^0 \rightarrow D^{*-}D_s^+\pi^+(N)/D^+(\rightarrow \pi^+\pi^-\pi^+(N))$ events generated according to the values reported in Tab. 4.1, 4.2 and 4.3. It contains also events with the same
Table 3.2: MC samples used in the study. For each sample the number of generated events is reported. In the second sample a cut at generator level on the $B^0$ transverse momentum has been applied, reducing the number of events by a factor 2.

<table>
<thead>
<tr>
<th>Decay</th>
<th>Generated events</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow D^{*-}(\rightarrow \pi^-D^0(\rightarrow K^+\pi^-))\tau^+(\rightarrow \pi^+\pi^-\pi^+\nu_\tau)\nu_\tau$</td>
<td>4500000</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}(\rightarrow \pi^-\bar{D}^0(\rightarrow K^+\pi^-))\pi^+\pi^-\pi^+$</td>
<td>1500000</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{*-}(\rightarrow \pi^-\bar{D}^0(\rightarrow K^+\pi^-))D_s^+(\rightarrow \pi^+\eta(\rightarrow \pi^+\pi^-\pi^0))$</td>
<td>3060000</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^{<em>-}D_s^{+(\pi^{(</em>)})}/D^+(\rightarrow \pi^+\pi^-\pi^+(N))$ at generator level</td>
<td>5000</td>
</tr>
</tbody>
</table>

final state where one of the three pions is actually a kaon, an electron or a proton, in order to emulate the misidentification. These misidentified events are reduced by a factor 10 to emulate the PID misidentification probability. From 5000 events of this MC-generator, 3771 and 253 originate from $D_s^+$ and $D^+$ decays, respectively.

Through the so-called stripping process at LHCb, some preselection cuts are applied to data and MC for the decay $B^0 \rightarrow D^*+\tau^+\nu_\tau$. Stripping is a filtering procedure whereby real and simulated events are selected and classified in several output streams according to criteria specific to each analysis being performed, with the aim of reducing the needed computing resources. The stripping cuts on the charged tracks, reported in Tab. 3.3, are applied to the following variables:

- the momentum $p$ and the transverse momentum $p_T$ of the particle;
- the probability of the track (measured in terms of $\chi^2$) to originate from the Primary Vertex (PV), that is the $pp$ interaction point;
- the ghost probability, i.e. the probability that the track is reconstructed from random combinations of hits in the detector;
- the $\chi^2$ of the fit used to determine the track from the information given by the tracking system;
- the PID$K$ of the particle, that is defined as $\log \mathcal{L}(K) - \log \mathcal{L}(\pi)$ [24], the difference between the log-likelihoods of the particle in the kaon and in pion hypotheses respectively; the likelihoods are computed using the information coming from the RICH system.

In Tab. 3.4 the stripping cuts on the composite particles are listed. The variables involved in these selection are:
### 3.2 Data and Monte Carlo samples

Table 3.3: The stripping cuts applied to the tracks of the final state.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Variable</th>
<th>Selection cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^+$ from $D^0$</td>
<td>$p_T$ &gt; 250 MeV</td>
<td>$\min(\chi^2)$ distance w.r.t. PV &gt; 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p$ &gt; 2 GeV</td>
</tr>
<tr>
<td></td>
<td>track fit $\chi^2/D.O.F.$ &lt; 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>track ghost probability &lt; 0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PID $K$ &gt; -5</td>
<td></td>
</tr>
<tr>
<td>$\pi^-$ from $D^0$</td>
<td>$p_T$ &gt; 250 MeV</td>
<td>$\min(\chi^2)$ distance w.r.t. PV &gt; 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p$ &gt; 2 GeV</td>
</tr>
<tr>
<td></td>
<td>track fit $\chi^2/D.O.F.$ &lt; 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>track ghost probability &lt; 0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PID $K$ &lt; 8</td>
<td></td>
</tr>
<tr>
<td>$\pi^-$ from $D^{*-}$</td>
<td>$p_T$ &gt; 110 MeV</td>
<td>track fit $\chi^2/D.O.F.$ &lt; 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>track ghost probability &lt; 0.6</td>
</tr>
<tr>
<td>$\pi^+, \pi^-, \pi^+$ from $\tau^+$</td>
<td>$p_T$ &gt; 250 MeV</td>
<td>$\min(\chi^2)$ distance w.r.t. PV &gt; 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>track fit $\chi^2/D.O.F.$ &lt; 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>track ghost probability &lt; 0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PID $K$ &lt; 8</td>
</tr>
</tbody>
</table>

- the $p_T$ of the composite particle and its decay products;
- the difference between the reconstructed mass and the nominal mass value given by the PDG;
- the masses of the particle daughters;
- the $\chi^2$ of the fit used to determine the decay vertex position;
- the probability of the particle (measured in terms of $\chi^2$) to originate from the PV;
- the DOCA, i.e. the Distance Of Closest Approach between the tracks of the particle daughters;
- $\cos(\theta_X)$, where $\theta_X$ is the angle between the momentum of the particle $X$ and the direction of flight from the primary interaction vertex and the decay.
### Events selection

**Table 3.4:** The stripping cuts applied on the reconstructed particles.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Variable Selection cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^0$</td>
<td>$p_T &gt; 1600$ MeV</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>vertex $\chi^2/D.O.F. &lt; 10$</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$ distance w.r.t. PV $&gt; 50$</td>
</tr>
<tr>
<td></td>
<td>$p_T(\pi^-) + p_T(K^+) &gt; 1200$ MeV</td>
</tr>
<tr>
<td></td>
<td>$DOCA(\pi^-, K^+) &lt; 0.5$ mm</td>
</tr>
<tr>
<td></td>
<td>$DOCA(\pi^-, K^+) \chi^2 &lt; 15$</td>
</tr>
<tr>
<td></td>
<td>$\cos(\theta_X) &gt; 0.999$</td>
</tr>
<tr>
<td>$D^{*-}$</td>
<td>$p_T &gt; 1250$ MeV</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>vertex $\chi^2/D.O.F. &lt; 25$</td>
</tr>
<tr>
<td></td>
<td>$M - M(D^0) &lt; 165.5$ MeV</td>
</tr>
<tr>
<td>$\tau^+$</td>
<td>$M \in [400$ MeV $\div 3500$ MeV $]$</td>
</tr>
<tr>
<td></td>
<td>$\cos(\theta_X) &gt; 0.99$</td>
</tr>
<tr>
<td></td>
<td>vertex $\chi^2 &lt; 25$</td>
</tr>
<tr>
<td></td>
<td>$\min(M(\pi^+\pi^-)) &lt; 1670$ MeV</td>
</tr>
<tr>
<td></td>
<td>$n^0$ of daughters with $p_T &lt; 300$ MeV $\leq 1$</td>
</tr>
<tr>
<td></td>
<td>$\max(DOCA(\pi^+, \pi^-, \pi^+)) &lt; 0.15$ mm</td>
</tr>
<tr>
<td>$B^0$</td>
<td>$M(D^{*-}) - M(\tau^+) \begin{cases} -2579$ MeV $\div 300$ MeV \ or \ -720$ MeV $\div 1721$ MeV \end{cases}$</td>
</tr>
<tr>
<td></td>
<td>$DOCA(D^{*-}, \tau^+) &lt; 0.13$ mm</td>
</tr>
<tr>
<td></td>
<td>$\cos(\theta_X) &gt; 0.999$</td>
</tr>
</tbody>
</table>

In Tab. 3.5 the efficiencies of these cuts are reported. After the reconstruction and the stripping process, 12147799 events are selected in the data sample. No specific requirements on the trigger have been imposed.

### 3.3 First cut-based selection

As explained in section 3.1, a first cut-based selection is necessary in order to suppress the most dominant background $B^0 \rightarrow D^{*-}\pi^+\pi^-\pi^+$ and to get a cleaner signal. The set of these cut is reported in Tab. 3.6. The most important cut is the first one, which will be referred to as inversion cut at 5 σ: it requires that the tau vertex...
3.3 First cut-based selection

Table 3.5: Efficiencies of the reconstruction and stripping pre-selection.

<table>
<thead>
<tr>
<th></th>
<th>Signal MC</th>
<th>$D^*^- D_s^+ $ MC</th>
<th>$D^*^- 3\pi$ MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before reconstruction</td>
<td>4500000</td>
<td>3060000</td>
<td>$2 \cdot 1500000$</td>
</tr>
<tr>
<td>After reconstruction and stripping</td>
<td>31887</td>
<td>37598</td>
<td>33278</td>
</tr>
<tr>
<td>$\varepsilon_{REC}(%)$</td>
<td>0.709 ± 0.004</td>
<td>1.229 ± 0.006</td>
<td>1.109 ± 0.006</td>
</tr>
</tbody>
</table>

Figure 3.1: The inversion cut at $5\sigma$.

is downstream with respect to the $D^0$ vertex with a significance of at least $5\sigma$, as shown in Fig. 3.1.

Some of the other cuts are topological, for example there are two requirements on the $\tau^+$ flight distance (with respect to the PV) along the z-axis and in the transverse plane. Another important requirement is that the $\bar{D}^0$ and the pions from $\tau^+$ must not originate directly from the PV, so a cut on the Impact Parameter (IP) with respect to the PV is applied. The $\bar{D}^0$ and the $\tau^+$ must originate from the same PV in order to reconstruct particles that belong to the same decay chain. Two cuts on the $\chi^2$ and on the error of the $3\pi$ vertex are applied in order to select a well reconstructed $\tau^+$ vertex.

In order to suppress misidentified kaons, there are cuts on pions $\text{ProbNNpi}$, that is the probability, computed through a MC-trained neural network, for a detected particle to be a pion (see section 2.5.2). The $3\pi$ mass region selected looking at the MC distribution is between 850 MeV and 2000 MeV. The last cut is on $\Delta M(D^{*-}) = M(D^{*-}) - M(D^0)$, applied to remove most of the combinatorial background; however some combinatorial is still present in the selected region and it will be estimated in section 4.1.1. The distributions of the variables listed in Tab. 3.6 are shown in Fig. 3.2 and 3.3 for MC samples of signal, $B^0 \rightarrow D^{*-} 3\pi$ and $B^0 \rightarrow D^{*-} D_s^+$.
Table 3.6: The first cut-based selection.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[vtx_z(D^0) - vtx_z(\tau^+)]/\text{error}$</td>
<td>$&lt;-5$</td>
</tr>
<tr>
<td>$[vtx_z(\tau^+) - vtx_z(PV)]/\text{error}$</td>
<td>$&gt;10$</td>
</tr>
<tr>
<td>$\sqrt{[vtx_x(\tau^+)]^2 + [vtx_y(\tau^+)]^2}$</td>
<td>$0.4 \text{ mm} \div 6 \text{ mm}$</td>
</tr>
<tr>
<td>$\chi^2[IP_{PV}(D^0)]$</td>
<td>$&gt;10$</td>
</tr>
<tr>
<td>$\chi^2[IP_{PV}(\pi)]$, $\pi$ from $\tau^+$</td>
<td>$&gt;15$</td>
</tr>
<tr>
<td>$\text{PV}(D^0)$ = $\text{PV}(\tau^+)$</td>
<td></td>
</tr>
<tr>
<td>number of $B^0$ candidates</td>
<td>$= 1$</td>
</tr>
<tr>
<td>$\chi^2[vtx(\tau^+)]$</td>
<td>$&lt;10$</td>
</tr>
<tr>
<td>$\text{error}[vtx(\tau^+)]$</td>
<td>$&lt;1 \text{ mm}$</td>
</tr>
<tr>
<td>$\text{ProbNNpi}$ of each $\pi$ from $\tau^+$</td>
<td>$&gt;0.6$</td>
</tr>
<tr>
<td>$\text{ProbNNpi}$ of $\pi$ from $D^{*-}$</td>
<td>$&gt;0.1$</td>
</tr>
<tr>
<td>$M(\tau^+)$</td>
<td>$850 \text{ MeV}/c^2 \div 2000 \text{ MeV}/c^2$</td>
</tr>
<tr>
<td>$\Delta M = M(D^{*-}) - M(D^0)$</td>
<td>$143 \text{ MeV}/c^2 \div 147 \text{ MeV}/c^2$</td>
</tr>
</tbody>
</table>

After this cut-based selection the $D^{*-} 3\pi$ events are expected to be completely removed, according to what happens applying these cuts on the $D^{*-} 3\pi$ MC (see Tab. 3.7). From the same table it comes out, as expected, that this selection affects the $D_s^+$ component in more or less the same way of the signal, since the three pions originate from a particle with a detectable lifetime in both decays.

11261 events survive the cut-based selection in data sample. In Fig. 3.4 the two dimensional plot of $\Delta M$ versus $M(D^0)$ after the cut-based selection (without the cut on $\Delta M$) is shown for data.

In Fig. 3.5 the $3\pi$ mass distribution after the first cut-based selection is shown. The peaks due to the exclusive decays $D^+ \rightarrow \pi^+\pi^-\pi^+$ and $D_{s}^{+} \rightarrow \pi^+\pi^-\pi^+$ are clearly visible at their respective masses. These peaks will be used in section 4.1.4 to estimate the signal and the $D_{(s)}^{+} \rightarrow \pi^+\pi^-\pi^+N$ contribution in data sample.
### 3.3 First cut-based selection

Table 3.7: Efficiencies of the first cut-based selection. The last line shows the total efficiency, i.e. the product of stripping (see Tab. 3.5) and selection efficiencies.

<table>
<thead>
<tr>
<th>Cuts applied in cascade</th>
<th>Signal MC</th>
<th>$D^- D^+_s$ MC</th>
<th>$D^-$ $3\pi$ MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before selection</td>
<td>31887</td>
<td>37598</td>
<td>33278</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cuts applied in cascade</th>
<th>$\varepsilon$ (%)</th>
<th>$\varepsilon$ (%)</th>
<th>$\varepsilon$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$vtx(D^0) - vtx(\tau^+)$/error $&lt;-5$</td>
<td>12.23 ± 0.18</td>
<td>17.6 ± 0.2</td>
<td>0.51 ± 0.04</td>
</tr>
<tr>
<td>$vtx(\tau^+) - vtx(PV)/error &gt; 10$</td>
<td>12.15 ± 0.18</td>
<td>17.4 ± 0.2</td>
<td>0.46 ± 0.04</td>
</tr>
<tr>
<td>0.4 mm $&lt; \sqrt{[vtx(\tau^+)]^2 + [vtx(\tau^-)]^2} &lt; 6$ mm</td>
<td>11.7 ± 0.2</td>
<td>16.2 ± 0.2</td>
<td>0.39 ± 0.03</td>
</tr>
<tr>
<td>$\chi^2[IP_{PV}(D^0)] &gt; 10$</td>
<td>10.1 ± 0.2</td>
<td>15.1 ± 0.2</td>
<td>0.26 ± 0.03</td>
</tr>
<tr>
<td>$\chi^2[IP_{PV}(\pi)], \pi$ from $\tau^+, &gt; 15$</td>
<td>9.1 ± 0.2</td>
<td>13.0 ± 0.2</td>
<td>0.16 ± 0.02</td>
</tr>
<tr>
<td>$PV(D^0) = PV(\tau^+)$</td>
<td>8.6 ± 0.2</td>
<td>12.2 ± 0.2</td>
<td>0.16 ± 0.02</td>
</tr>
<tr>
<td>number of $B^0$ candidates $= 1$</td>
<td>7.46 ± 0.15</td>
<td>8.78 ± 0.15</td>
<td>0.054 ± 0.013</td>
</tr>
<tr>
<td>$\chi^2[vtx(\tau^+)] &lt; 10$</td>
<td>6.54 ± 0.14</td>
<td>7.64 ± 0.14</td>
<td>0.045 ± 0.012</td>
</tr>
<tr>
<td>error[$vtx(\tau^+)] &lt; 1$ mm</td>
<td>5.79 ± 0.13</td>
<td>6.05 ± 0.12</td>
<td>0.027 ± 0.009</td>
</tr>
<tr>
<td>ProbNN of each $\pi$ from $\tau^+, &gt; 0.6$</td>
<td>4.85 ± 0.12</td>
<td>4.73 ± 0.11</td>
<td>0.018 ± 0.007</td>
</tr>
<tr>
<td>ProbNN of $\pi$ from $D^{*-} &gt; 0.1$</td>
<td>4.80 ± 0.12</td>
<td>4.67 ± 0.11</td>
<td>0.018 ± 0.007</td>
</tr>
<tr>
<td>580 MeV/c^2 $&lt; M(\tau^+) &lt; 2000$ MeV/c^2</td>
<td>4.77 ± 0.12</td>
<td>4.10 ± 0.10</td>
<td>0.006 ± 0.004</td>
</tr>
<tr>
<td>143 MeV/c^2 $&lt; \Delta M &lt; 147$ MeV/c^2</td>
<td>4.49 ± 0.12</td>
<td>3.79 ± 0.10</td>
<td>0.003 ± 0.003</td>
</tr>
</tbody>
</table>

After selection: 1431 1424 1

$\varepsilon_{SEL} (%)$ | 4.49 ± 0.12 | 3.79 ± 0.10 | 0.003 ± 0.003 |

$\varepsilon_{REC} \cdot \varepsilon_{SEL} (%)$ | 0.0318 ± 0.0008 | 0.0465 ± 0.0012 | $(3 \pm 3) \cdot 10^{-8}$ |
Figure 3.2: The distributions of the variables on which cuts are applied for MC samples of signal, $B^0 \rightarrow D^{*-3\pi}$ and $B^0 \rightarrow D^{*-}D^+_s$ (1).
3.3 First cut-based selection

![Figure 3.3: The distributions of the variables on which cuts are applied for MC samples of signal, $B^0 \rightarrow D^{*-}3\pi$ and $B^0 \rightarrow D^*-D^+_s$ (2).]
Figure 3.4: Two dimensional distribution of $\Delta M$ as function of $M(D^0)$.

Figure 3.5: The $3\pi$ mass distribution after the first cut-based selection.
Chapter 4

Study of the data sample composition and isolation of signal

4.1 Composition of surviving events

After the first cut-based selection, the data sample is assumed to consist of signal, \( B^0 \rightarrow D^{*-}D_+(s) \) decays where the \( D_+(s) \) mesons result in three charged pions and zero or more neutral particles in the final state, decays with misidentified kaons, combinatorial background due to false \( D^{*-} \) and some surviving \( B^0 \rightarrow D^{*-}\pi^+\pi^-\pi^+N \) decays. The yields of each of these contributions is estimated in different ways, explained in the next sections.

4.1.1 Combinatorial background

The first component which has been studied is the combinatorial background, that is given by events in which one or more charged tracks are wrongly associated with the reconstructed \( D^{*-} \) decay. In order to determine its contribution, the \( \Delta M(D^{*-}) = M(D^{*-}) - M(D^0) \) distribution is studied after the first cut-based selection (after releasing the cut on this variable). This distribution, showing a peak in the region around 145 MeV and a smooth shape that is due to combinatorial background, is fitted with a function equal to the sum of two components, which model the signal and the background respectively. The integral of the background function in the region 143 MeV ÷ 147 MeV defined by the cut-based selection gives the number of combinatorial events in the selected sample.

For this fitting procedure (and for the following ones) the software RooFit is used [30]. The function used for signal is equal to the sum of three gaussian distributions
Study of the data sample composition and isolation of signal

Figure 4.1: Fit of $\Delta M(D^{*-}) = M(D^{*-}) - M(D^0)$ distribution.

with same mean and different standard deviations. For background modelization a phase-space function with a threshold effect is adopted. This function, provided by RooFit, depends on four parameters in the following way:

$$f(\Delta M) = \left[1 - e^{-\frac{(\Delta M - d)}{c}}\right] \left(\frac{\Delta M}{d}\right)^a + b \left(\frac{\Delta M}{d} - 1\right).$$

In this fit the parameters $a$ and $b$ are set to 0 in order to keep the fit stable.

The fit results are reported in Fig. 4.1. The fitted values of the gaussian mean is $(145.454 \pm 0.009)$ MeV/$c^2$, while the three standard deviations are $\sigma_1 = (3.2 \pm 0.5)$ MeV/$c^2$, $\sigma_2 = (0.93 \pm 0.06)$ MeV/$c^2$, $\sigma_3 = (0.46 \pm 0.04)$ MeV/$c^2$. The relative fractions of the three gaussians with respect to the total signal function are 0.05 ± 0.04, 0.60 ± 0.07 and 0.35 ± 0.08 respectively. According to the fit results there are $9700 \pm 300$ events in the signal peak and $12300 \pm 300$ background events. The number of combinatorial events in the selected region, obtained normalizing the background PDF to the number of background events and integrating it in range 143 MeV ÷ 147 MeV, is $2120 \pm 50$. 
4.1 Composition of surviving events

4.1.2 Misidentified kaon contribution

Events due to $B^0 \rightarrow D^- D^+ (\rightarrow \pi^+ K^- \pi^+ (N))$ decays where the kaon is misidentified as pion, can be studied by looking at the $3\pi$ mass distribution with the hypothesis that the $\pi^-$ has a kaon mass, to see if there is a peak in this distribution around the $D^+$ mass. Fitting this peak gives the number of $B^0 \rightarrow D^- D^+ (\rightarrow \pi^+ K^- \pi^+ (N))$ events. The number of events with partially reconstructed $D^+$ decays can be estimated by multiplying this number by $K_{D^+,missID}$

\[ K_{D^+,missID} \equiv \frac{B(D^+ \rightarrow \pi^+ K^- \pi^+ (N))}{B(D^+ \rightarrow \pi^+ K^- \pi^+ )} \]  

By assuming the same efficiency for both channels, since the only decay of the type $D^+ \rightarrow \pi^+ K^- \pi^+ N$ is $D^+ \rightarrow \pi^+ K^- \pi^+ \pi^0$, which has a branching ratio of $(5.99 \pm 0.18)\%$, and since $B(D^+ \rightarrow \pi^+ K^- \pi^+) = (9.13 \pm 0.19)\%$, it comes out that $K_{D^+,missID}$ is equal to $1.66 \pm 0.06$.

Figure 4.2 shows the fit to the $\pi^+ K^- \pi^+$ mass near the $D^+$ peak, according to which there are $320 \pm 30$ $D^+ \rightarrow \pi^+ K^- \pi^+$ events where the kaon has been misidentified as a pion. A gaussian distribution has been chosen for signal and a first-order polynomial for background. Multiplying this number by $K_{D^+,missID}$, the resulting number of misidentified $B^0 \rightarrow D^- D^+ (\rightarrow \pi^+ K^- \pi^+ (N))$ is $530 \pm 50$.

This kind of background can be reduced by applying a cut on $\text{ProbNNp}_i$ of the $\pi^-$, for instance requiring it to be greater than 0.9. In Fig. 4.3 the plot of the $3\pi$ mass with the kaon hypothesis is illustrated before (solid line) and after (dashed line) this cut. This cut impacts significantly on the signal efficiency, therefore it will not be used in the following.

4.1.3 Branching ratios of $D^+_{(s)} \rightarrow \pi^+ \pi^- \pi^+ N$

The yield of $B^0 \rightarrow D^s - D^+_s (\rightarrow \pi^+ \pi^- \pi^+)$ background can be determined by fitting the $3\pi$ invariant mass around the $D^+_s$ peak. The number of $B^0 \rightarrow D^s - D^+_s (\rightarrow \pi^+ \pi^- \pi^+ (N))$ events can be extrapolated from the number of exclusive decay $B^0 \rightarrow D^s - D^+_s (\rightarrow \pi^+ \pi^- \pi^+)$ events, by using the ratio:

\[ K_{D^+_s} \equiv \frac{B(D^+_s \rightarrow \pi^+ \pi^- \pi^+ (N))}{B(D^+_s \rightarrow \pi^+ \pi^- \pi^+ )} \]  

between the inclusive and exclusive branching ratios of $D^+_s$ into $3\pi$, assuming the same reconstruction and selection efficiency for the two processes.
Fit of exclusive $D^+$ peak (kaon hypothesis)

Figure 4.2: Fit of $D^+$ peak in $M(3\pi)$ distribution with a kaon hypothesis on $\pi^-$.  

Figure 4.3: Plot of the $3\pi$ mass with the kaon hypothesis before and after $Prob_{NN\pi \pi}(\pi^-) > 0.9$.  

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fit_dplus}
\caption{Fit of $D^+$ peak in $M(3\pi)$ distribution with a kaon hypothesis on $\pi^-$.}
\end{figure}
Tables 4.1 and 4.2 show the branching ratios of $D_s^+$ into final states containing three charged pions and neutral particles, passing through all intermediate resonances in which $D_s^+$ can decay in order to obtain such a final state. Decays with more than three charged pions are not considered because the isolation criteria (see section 3.1) suppresses them very efficiently when applied to the 3π vertex. It comes out from the table that $\mathcal{B}(D_s^+\rightarrow \pi^+\pi^-\pi^+) = (15.4 \pm 1.2)\%$. Since $\mathcal{B}(D_s^+\rightarrow \pi^+\pi^-\pi^+) = (1.09 \pm 0.05)\%$, $K_{D_s^+}$ can be calculated:

$$K_{D_s^+} = \frac{(15.4 \pm 1.2)\% + (1.09 \pm 0.05)\%}{(1.09 \pm 0.05)\%} = 15.1 \pm 1.3.$$  \hspace{1cm} (4.4)

**Table 4.1:** Branching ratios of $D_s^+$ into final states containing three charged pions and neutral particles, passing through $\eta/\eta'$ intermediate resonances. According to PDG, $\mathcal{B}(D_s^+\rightarrow \eta'\rho^0)$ could be too high, so it has been reduced by a factor 2 to match the inclusive $\mathcal{B}(D_s^+\rightarrow \eta'X)$.

<table>
<thead>
<tr>
<th>Intermediate decays</th>
<th>$\mathcal{B}$ (%)</th>
<th>$D_s^+$ decays $\mathcal{B}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta \rightarrow \pi^+\pi^-\pi^0/\gamma$</td>
<td>$27.1 \pm 0.3$</td>
<td>$\eta\pi^+$ $\eta\rho^+$ $\eta'\pi^+$ $\eta'\rho^+$</td>
</tr>
<tr>
<td>$\eta' \rightarrow \rho^0\gamma$</td>
<td>$29.1 \pm 0.5$</td>
<td>$0.46 \pm 0.03$ $2.4 \pm 0.2$</td>
</tr>
<tr>
<td>$\eta' \rightarrow \eta(\rightarrow N)\pi^+\pi^-$</td>
<td>$30.9 \pm 0.5$</td>
<td>$1.15 \pm 0.08$ $1.8 \pm 0.6$</td>
</tr>
<tr>
<td>$\eta' \rightarrow \eta(\rightarrow \pi^+\pi^-\pi^0/\gamma)\pi^0\pi^0$</td>
<td>$6.0 \pm 0.2$</td>
<td>$1.22 \pm 0.08$ $1.9 \pm 0.7$</td>
</tr>
<tr>
<td>$\eta' \rightarrow \omega(\rightarrow \pi^+\pi^-\pi^0)\gamma$</td>
<td>$2.5 \pm 0.2$</td>
<td>$0.237 \pm 0.018$ $0.38 \pm 0.13$</td>
</tr>
</tbody>
</table>

**Table 4.2:** Branching ratios of $D_s^+$ into final states containing three charged pions and neutral particles, passing through $\omega$, $\phi$ and $\tau$ decays.

<table>
<thead>
<tr>
<th>Intermediate decays</th>
<th>$\mathcal{B}$ (%)</th>
<th>$D_s^+$ decays $\mathcal{B}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega \rightarrow \pi^+\pi^-\pi^0$</td>
<td>$89.2 \pm 0.7$</td>
<td>$\omega\pi^+$ $\omega\rho^+$ $\omega\phi$ $\omega\tau$</td>
</tr>
<tr>
<td>$\omega \rightarrow \pi^+\pi^-$</td>
<td>$1.53 \pm 0.13$</td>
<td>$0.24 \pm 0.06$ $2.8 \pm 0.7$ $4.5 \pm 0.4$ $8 \pm 2$</td>
</tr>
<tr>
<td>$\phi \rightarrow \pi^+\pi^-\pi^0$</td>
<td>$15.3 \pm 0.3$</td>
<td>$0.043 \pm 0.011$ $0.69 \pm 0.06$ $1.3 \pm 0.4$</td>
</tr>
<tr>
<td>$\tau \rightarrow \pi^+\pi^-\pi^+(\pi^0)\nu_\tau$</td>
<td>$13.93 \pm 0.08$</td>
<td>$0.77 \pm 0.03$</td>
</tr>
</tbody>
</table>

The same calculation can be performed for the ratio:

$$K_{D^+} = \frac{\mathcal{B}(D_s^+\rightarrow \pi^+\pi^-\pi^+(N))}{\mathcal{B}(D_s^+\rightarrow \pi^+\pi^-\pi^+)}.$$  \hspace{1cm} (4.5)
Study of the data sample composition and isolation of signal

Table 4.3: Branching ratios of \( D^+ \) into final states containing three charged pions and neutral particles, passing through \( \eta \), \( \eta' \), \( \omega \) resonances. The \( B \) of the resonant decays that give \( \pi^+\pi^-\pi^0 \) final state must be subtracted from \( B(D^+ \rightarrow \pi^+\pi^-\pi^+) \).

<table>
<thead>
<tr>
<th>Intermediate decays</th>
<th>( B(%) )</th>
<th>( \pi^+\pi^-\pi^0 )</th>
<th>( \eta\pi^- )</th>
<th>( \eta'\pi^- )</th>
<th>( \eta\pi^+ )</th>
<th>( \eta'\pi^+ )</th>
<th>( \omega\pi^+ )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta \rightarrow \pi^+\pi^-\pi^0 )</td>
<td>22.9 ± 0.3</td>
<td>0.081 ± 0.005</td>
<td>0.032 ± 0.008</td>
<td>0.13 ± 0.08</td>
<td>0.35 ± 0.02</td>
<td>0.14 ± 0.04</td>
<td>0.47 ± 0.03</td>
</tr>
<tr>
<td>( \eta \rightarrow \pi^+\pi^-\pi^0 )</td>
<td>4.22 ± 0.08</td>
<td>0.0149 ± 0.0009</td>
<td>0.0058 ± 0.0015</td>
<td>0.136 ± 0.009</td>
<td>0.144 ± 0.009</td>
<td>0.050 ± 0.015</td>
<td>0.028 ± 0.002</td>
</tr>
<tr>
<td>( \eta \rightarrow \eta\pi^- )</td>
<td>29.1 ± 0.5</td>
<td>0.144 ± 0.009</td>
<td>0.050 ± 0.015</td>
<td>0.028 ± 0.002</td>
<td>0.010 ± 0.001</td>
<td>0.0315 ± 0.0012</td>
<td>0.0029 ± 0.0013</td>
</tr>
<tr>
<td>( \eta \rightarrow \eta'\pi^- )</td>
<td>6.0 ± 0.2</td>
<td>0.028 ± 0.002</td>
<td>0.010 ± 0.001</td>
<td>0.0315 ± 0.0012</td>
<td>0.0029 ± 0.0013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \eta \rightarrow \omega\pi^- )</td>
<td>2.5 ± 0.2</td>
<td>0.0135 ± 0.0002</td>
<td>0.0018 ± 0.0002</td>
<td>0.0018 ± 0.0002</td>
<td>0.0006 ± 0.0002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega \rightarrow \pi^+\pi^-\pi^0 )</td>
<td>0.38 ± 0.04</td>
<td>0.0018 ± 0.0002</td>
<td>0.0006 ± 0.0002</td>
<td>&lt; 0.0003 ± 0.0002</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The branching ratios are computed in the same way as the \( D_s^+ \) ones, and they are reported in Tab. 4.3. The value of \( B(D_s^+ \rightarrow \pi^+\pi^-\pi^+) \) is calculated to be \((1.61 ± 0.08)\)\%.

The exclusive branching ratio \( B((D \rightarrow \pi^+\pi^-\pi^+)) = (0.318 ± 0.018)\)\%, so

\[
K_{D^+} = \frac{(1.61 ± 0.08)\% + (0.318 ± 0.018)\%}{(0.318 ± 0.018)\%} = 6.1 ± 0.4.
\] (4.6)

4.1.4 Fit of exclusive \( D_s^+ \) and \( D^+ \) peaks and estimate of signal contribution

The estimation of signal and inclusive \( D_s^+ \) contributions is done through some extrapolations which are explained in this section.

First of all, the \( B^0 \rightarrow D^{*-}D_s^+(\rightarrow \pi^+\pi^-\pi^+) \) events numbers are directly measured fitting the \( D_s^+ \) and the \( D^+ \) peaks in the 3\( \pi \) mass distribution. Then these numbers are multiplied by \( K_{D_s^+} \) and \( K_{D^+} \) to get the number of inclusive \( B^0 \rightarrow D^{*-}D_s^+(\rightarrow \pi^+\pi^-\pi^+\pi^0) \) events.

Since for these decays only the MC at generator level is available, it is assumed that the efficiency for the exclusive and the inclusive decays is the same, except for what concerns the cut on the 3\( \pi \) mass, which is, of all the variables involved in the cut-based selection, the only one that can be looked at generator level. In the MC-generator there are 3771 \( D_s^+ \) events and 253 \( D^+ \) events; the 3\( \pi \) mass cut selects 2874 \( D_s^+ \) and 238 \( D^+ \).

In order to evaluate the number of signal events, the number of exclusive \( B^0 \rightarrow D^{*-}D_s^+(\rightarrow \pi^+\pi^-\pi^+) \) events is scaled according to the ratio between the branching ratios and the ratio between the efficiencies. The efficiencies are evaluated on the MC samples of signal and \( D^{*-}D_s^+ \).
4.1 Composition of surviving events

The results of the fit of $D_s^+$ and $D^+$ peaks are shown in Fig. 4.4 and 4.5. The signal is described by a gaussian and the background by a first-order polynomial. The estimated yields of exclusive $D_s^+$ and $D^+$ decays are $350 \pm 30$ and $92 \pm 16$, respectively.

\[\begin{align*}
\text{Fit of exclusive } D_s \text{ peak} & \\
\hline
\text{Events } / (1.2 \text{ MeV/c}) & \\
\hline
10 & 15 \quad 20 \quad 25 \quad 30 \quad 35 \\
\hline
\hline
\text{Data} & \text{Signal + Background} & \text{Gaussian Signal} & \text{Polynomial background} \\
\hline
\text{Ng1} & 349 \pm 28 \\
\text{Npol} & 92 \pm 23 \\
\text{mu} & 1970.02 \pm 0.63 \text{ MeV/c}^2 \\
\text{sigma1} & 8.80 \pm 0.62 \text{ MeV/c}^2 \\
\text{slope} & -0.000428 \pm 0.00018 (\text{MeV/c}^2)^{-1} \\
\end{align*}\]

**Figure 4.4:** Fit of $D_s^+$ peak in $M(3\pi)$ distribution.

Using the values of $K_{D_s^+}$ and $K_{D^+}$ of Eq. (4.4) and (4.6), the $B$ and efficiency ratios reported in Tab. 4.4 and the fit results, the data sample turns out to be split according to the contributions listed in Tab. 4.5. This estimation seems to explain the nature of 8300 events on the total of 11261, so about 3000 remaining events could be surviving $B^0 \rightarrow D_s^{*-}3\pi$ or some source background that has not been considered, as $B^0 \rightarrow D^{*-}DK$. This gap may be due to the fact that the differences in reconstruction efficiency of partial reconstructed background have been neglected.

4.1.5 The “$D_s^{+}$-o-meter”

The sample selected so far should be dominated by $B^0 \rightarrow D_s^{*-}D_s^+(\rightarrow 3\pi N)$ decays and, according to Tab. 4.1, in a large part of these decays one $\pi^+\pi^-$ pair comes from $\eta \rightarrow \pi^+\pi^-\pi^0$ or $\eta' \rightarrow \eta\pi^+\pi^-$ decays. If this is the case, the $\pi^+\pi^-$ mass should
Study of the data sample composition and isolation of signal

Figure 4.5: Fit of $D^+$ peak in $M(3\pi)$ distribution.

Table 4.4: Ratios between $B$ and efficiencies used to extrapolate the contributions of the various decays. The reconstruction and cut efficiencies are taken from Tab. 3.5 and 3.7.

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{B}{B}$ (signal) / $\frac{B}{B}$ (exclusive $D_s^+$ background)</td>
<td>4.4 ± 0.4</td>
</tr>
<tr>
<td>$\frac{\varepsilon_{REC}(signal)}{\varepsilon_{REC}(exclusive D_s^+ background)}$</td>
<td>0.577 ± 0.004</td>
</tr>
<tr>
<td>$\frac{\varepsilon_{CUT}(signal)}{\varepsilon_{CUT}(exclusive D_s^+ background)}$</td>
<td>1.18 ± 0.04</td>
</tr>
<tr>
<td>$K_{D_s^+}$</td>
<td>15.1 ± 1.3</td>
</tr>
<tr>
<td>$\varepsilon_{CUT}(D_s^+ from gen. MC) = 2874/3771$</td>
<td>(76.2 ± 0.7)%</td>
</tr>
<tr>
<td>$K_{D^+}$</td>
<td>6.1 ± 0.4</td>
</tr>
<tr>
<td>$\varepsilon_{CUT}(D^+ from gen. MC) = 238/253$</td>
<td>(94.1 ± 1.5)%</td>
</tr>
</tbody>
</table>

Table 4.5: Estimated yields for signal and background contributions in the data sample.

<table>
<thead>
<tr>
<th>Decays</th>
<th>Number of events</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_0 \rightarrow D_s^- D_s^+ (\rightarrow \pi^+ \pi^- \pi^+ (N))$</td>
<td>4000 ± 500</td>
</tr>
<tr>
<td>$B_0 \rightarrow D_s^- D_s^+ (\rightarrow \pi^+ \pi^- \pi^+ (N))$</td>
<td>500 ± 100</td>
</tr>
<tr>
<td>$B_0 \rightarrow D_s^- D_s^+ (\rightarrow \pi^+ \pi^- \pi^+ \pi^0 (N))$</td>
<td>1040 ± 140</td>
</tr>
<tr>
<td>combinatorial</td>
<td>2120 ± 50</td>
</tr>
<tr>
<td>misidentified $B_0 \rightarrow D^+ D^- (\rightarrow \pi^+ K^- \pi^+ (N))$</td>
<td>530 ± 50</td>
</tr>
</tbody>
</table>
be peaked at low mass. This can be checked by looking at the minimum of the di-pion \((\pi^+\pi^-)\) mass distribution in the \(D^*_-D_s^+\) MC-generator (see Fig. 4.6). This feature can be used as a "\(D_s^+\)-o-meter", i.e. one can look for the presence of this characteristic shape for the minimum di-pion mass in data, to check qualitatively whether background is actually dominated by \(D_s^+\) decays.

In order to apply this technique, the minimum of the di-pion \((\pi^+\pi^-)\) mass \(\text{min}[M(\pi^+\pi^-)]\) is plotted for data events and for signal MC, \(D^*_-D_s^+\) generator MC, and combinatorial in Fig. 4.7. Signal, \(D_s^+\) and combinatorial contributions are normalized according to the numbers reported in Tab. 4.5. \(D^*_-D_s^+\) decays with a \(\eta\) or a \(\eta'\) in the three pions decay chain are drawn separately from decays where there are no \(\eta\) or \(\eta'\). The plot shows also the contribution of events with one \(e/\mu/K\) misidentified as a pion; these events are included in the \(D^*_-D_s^+\) MC, as explained in section 3.2.

The \(\text{min}[M(\pi^+\pi^-)]\) shape for combinatorial events is taken from the sideband of \(\Delta M(\pi^+\pi^-)\), defined in the range \(148.097\) MeV \(\div\) \(160\) MeV; the sideband is \(3\sigma\) away from the \(\Delta M(\pi^+\pi^-)\) peak, where \(\sigma\) is the weighted average of the three standard deviations obtained from the fit described in section 4.1.1.

**Figure 4.6:** The \(\text{min}[M(\pi^+\pi^-)]\) distribution for the \(D^*_-D_s^+\) MC at generator level when \(\eta\) or \(\eta'\) are in the decay chain of the three pions. The clear peak on the left is due to the decay \(\eta \rightarrow \pi^+\pi^-\pi^0\).
Study of the data sample composition and isolation of signal

4.2 Isolation of signal

In this section a technique used to reconstruct the $B^0$ and $D^+_s$ momenta in $D^+_s$ decays with three pions and one or more neutral particles is illustrated, together with its utilization in a MultiVariate Analysis (MVA) performed to suppress the dominant background and isolate the signal in the data sample.

Figure 4.7: The min[$M(\pi^+\pi^-)$] distribution for data. The superimposed green histogram is the sum of the the same distribution for signal MC, combinatorial, $D^+_s$ with $\eta/\eta'$, $D^+_s$ without $\eta/\eta'$ and events with misidentified $e/\mu/K$.

The application of this “$D^+_s$-o-meter” confirms that the majority of events surviving the first cut-based selection is indeed due to $B^0 \rightarrow D^+_s D^+_s (\rightarrow \pi^+\pi^-\pi^+)$ decays. The peak at low mass is in fact well described by the sum of components (even if the $D^* D^+_s$ MC is at generator level only), and the tail at high mass is well modeled too. The missing 3000 events are located in the middle region.
4.2 Isolation of signal

4.2.1 Partial reconstruction

In the $B^0 \rightarrow D^{*-}D_s^+$ decay the following equation can be written (momentum conservation, see Fig. 4.8):

$$|\vec{p}_B|\hat{u}_B = |\vec{p}_{D_s^+}|\hat{u}_{D_s^+} + \vec{p}_{D^*},$$

(4.7)

where $\vec{p}_B$, $\vec{p}_{D_s^+}$ and $\vec{p}_{D^*}$ are the momenta of the mesons involved in the decay, and $\hat{u}_B, \hat{u}_{D_s^+}$ their unit vectors.

![Figure 4.8: Schematic representation of $B^0 \rightarrow D^{*-}D_s^+ (\rightarrow \pi^+\pi^-\pi^+ N)$ decay.](image)

Starting from this relation and applying some vectorial algebra, it is possible to reconstruct the values of $|\vec{p}_B|$ and $|\vec{p}_{D_s^+}|$ in two different ways, that will be called “vectorial” and “scalar” approaches. The magnitudes of the momenta obtained with these two approaches are equal to:

$$P_{B,v} = \frac{|\vec{p}_{D^*} \times \hat{u}_{D_s^+}|}{|\hat{u}_B \times \hat{u}_{D_s^+}|},$$

(4.8a)

$$P_{B,s} = \frac{\vec{p}_{D^*} \cdot \hat{u}_B - (\vec{p}_{D^*} \cdot \hat{u}_{D_s^+})(\hat{u}_B \cdot \hat{u}_{D_s^+})}{1 - (\hat{u}_B \cdot \hat{u}_{D_s^+})^2},$$

(4.8b)

for the $B^0$ momentum, and

$$P_{D_{s,v}} = \frac{|\vec{p}_{D^*} \times \hat{u}_B|}{|\hat{u}_{D_s^+} \times \hat{u}_B|},$$

(4.9a)

$$P_{D_{s,s}} = \frac{(\vec{p}_{D^*} \cdot \hat{u}_B)(\hat{u}_B \cdot \hat{u}_{D_s^+}) - \vec{p}_{D^*} \cdot \hat{u}_{D_s^+}}{1 - (\hat{u}_B \cdot \hat{u}_{D_s^+})^2},$$

(4.9b)

for the $D_s^+$ momentum.

Therefore the $B^0$ and $D_s^+$ momenta are reconstructed starting from three vectors:
Study of the data sample composition and isolation of signal

Figure 4.9: The profile of the correction $dz$ in function of the $3\pi$ mass. The chosen fitting function is a parabola.

- $\vec{p}_{D^*}$, which is fully reconstructed;
- $\hat{u}_B$, that is the direction of the line joining the $B^0$ vertex and the PV;
- $\hat{u}_{D^+_s}$, which at first approximation is given by the flight direction of the three-pion system, i.e. without considering possible neutral particles in $D^+_s$ decay.

This technique can be extended to take into account the presence of neutral particles in the $D^+_s$ decay chain, by applying a correction to the $B^0$ vertex position. In order to do that, a parametrization of this correction as function of $M(3\pi)$ has been obtained using the $D^{*-}D^{+}_s$ MC. Figure 4.9 shows the profile of $-dz = vt x_z (B^0) - vt x_z^{TRUE} (B^0)$, the correction on the $B^0$ vertex position along $z$, as function of the $3\pi$ mass. A parabolic function has been chosen in order to fit this profile, and, after fitting, the dependence can be expressed in the following way:

$$-dz = 3.0346 - 0.00439605 \cdot M(3\pi) + 1.6 \cdot 10^{-6} [M(3\pi)]^2,$$

(4.10)

with $dz$ in mm and $M(\pi^+\pi^-\pi^\pm)$ in MeV/$c^2$. 

4.2 Isolation of signal

This correction allows to recompute the $B^0$ vertex position, in order to get new momentum values at a next-level of approximation. They are called $P_{B,vn}$, $P_{B,sn}$, $P_{D_s,vn}$ and $P_{D_s,sn}$, and they are calculated by using (4.8a), (4.9a), after plugging in the new directions $\hat{u}_{B,n}$ and $\hat{u}_{D^+_s,n}$. Another set of variables which can be reconstructed by using $P_{B,v(s,m,sn)}$ and $\hat{u}_{B(B,n)}$ are the squared $D^+_s$ masses built in the various hypotheses, $m^2_{D^+_s,v(s,m,sn)}$, which are calculated using the nominal $B^0$ and $D^*$ masses. This variables will be useful in defining a $D^+_s$ mass region for the training of the MVA.

Relation (4.7) does not hold for signal decays, since in these decays the additional neutrino from the $B^0$ decay should be added to the second term of the equation. For this reason, variables coming from this partial reconstruction can be used to discriminate signal from $D^+_s$ background. This is the topic of next section.

4.2.2 MultiVariate Analysis

The aim of a MultiVariate Analysis (MVA) is to discriminate signal from background events. Since an event is described by multiple feature variables, the MVA consists in finding the best approximation $f(\vec{x}, \vec{w})$ of the unknown discriminating function $f(\vec{x})$, where $\vec{x}$ is the vector of the feature variables and $\vec{w}$ the set of parameters typical of the method adopted. In fact the feature variables can be thought as describing a multidimensional space, whose dimension has to be reduced by the MVA in order to classify the events enclosed in this reduced region. This approach takes into account the correlations between the feature variables. There is no a priori knowledge of the approximated function, which is inferred directly from the data, in fact the MVA uses a machine learning approach: the function is trained on a sample (usually a simulated one) in order to extract the parameters $\vec{w}$ of the model, and then tested on a statistically independent sample.

There are many different MVA methods: the most important are rectangular cut optimization, multidimensional likelihood estimator, k-Nearest Neighbour classifier, Fisher discriminants, artificial neural network, support vector machine, boosted decision trees. The classifier used in this study is the Boosted Decision Tree (BDT), and its training, evaluation and application have been obtained by means of the Toolkit for Multivariate Analysis (TMVA) [31], a ROOT integrated software package.
4.2.2.1 The Boosted Decision Trees

A Decision Tree (DT) is a binary tree structured classifier in which sequential rectangular cuts are applied. At each step, the best cut for any variable is found, dividing the feature space in two sub-partitions, for which additional best cuts are found, and so on. The starting point is called root node, where the entire sample is analyzed. In the successive stages (branch nodes), increasingly smaller portions of the feature space are analyzed.

Usually the criterion adopted to separate the sample at a given branch node is to reduce the impurity of the sample, parametrized with the so-called Gini Index:

$$G = P(1 - P), \quad (4.11)$$

where $P = S/(S + B)$ is the signal purity, with $S$ and $B$ equal to the amounts of signal and background respectively. The Gini Index assumes maximum value when the sample is fully mixed ($P = 1/2$) and null value when the sample is completely pure ($P = 1$). The feature space partition ends when no further impurity reduction can be achieved. This happens at the last tree nodes in the DT, which are called leaf nodes. A schematic structure of a DT is pictured in Fig. 4.10.

The main advantages of a DT are the tolerance to missing variables in both training and test samples and insensitivity to irrelevant variables, while the most important disadvantage is the instability with respect to statistical fluctuations: in fact, if two variables have similar discriminating power at a split node, a statistical fluctuation can lead to the choice of one variable instead of the other one, giving rise to an altered classifier response. For this reason a DT is usually boosted, i.e. a forest of decision trees is generated and each event is classified according to the major number of times it has been recognized as signal-like or background-like in every tree. This process allows to stabilize the classifier response and drastically improve the separation power.

The boosting algorithm used in this study is the adaptive boost (AdaBoost). The working principle of AdaBoost is that events that were misclassified during the training of a DT are given a higher event weight during the training of the following DT. It works in the following way: given the $i$-th DT, the events that have been misclassified by the previous DT are weighted using the boost weight $\alpha$:

$$\alpha_i = \frac{1 - \epsilon_i^{-1}}{\epsilon_i^{-1}}, \quad (4.12)$$

where $\epsilon_{i-1}$ is the misclassification rate of the previous tree.
4.2 Isolation of signal

The output of the $i$-th DT is $h_i(\vec{x}) = 1$ for signal and $h_i(\vec{x}) = -1$ for background. The BDT output of a forest of $N$ DTs is defined in this way:

$$BDT(\vec{x}) = \frac{1}{N} \sum_{i=1}^{N} \log(\alpha_i)h_i(\vec{x}),$$

that is a number between $-1$ and $1$.

In this way a BDT output distribution for a sample can be built. If the classifier has been efficient, there are two peaks in this distribution, one towards $BDT = 1$ for signal-like events and one towards $BDT = -1$ for background-like events. The BDT has a good discriminating performance if these two peaks are not overlapped.

4.2.2.2 Input variables and training samples for BDT

The input variables chosen for the BDT are listed in Tab. 4.6. Besides variables cal-
Table 4.6: Input variables used in the BDT.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M(B^0)$</td>
<td>the visible $B^0$ mass, obtained by adding the $D^{*-}$ and $3\pi$ 4-momenta</td>
</tr>
<tr>
<td>$\max[M(\pi^+\pi^-)]$</td>
<td>the maximum of di-pion masses in the $D_s^+$ decay</td>
</tr>
<tr>
<td>$\log</td>
<td>(P_{B,s} - P_{B,v})/P_{B,v}</td>
</tr>
<tr>
<td>$\log</td>
<td>(P_{B,sn} - P_{B,vn})/P_{B,vn}</td>
</tr>
<tr>
<td>$\log[</td>
<td>P_{D_s,sn}/P_{D_s,v}</td>
</tr>
<tr>
<td>$\log[</td>
<td>P_{D_s,v}/</td>
</tr>
<tr>
<td>$\chi^2_{vtx}(B^0)$</td>
<td>number of tracks which can be added to the $B^0$ vertex without decreasing the vertex fit quality</td>
</tr>
<tr>
<td>$B0_isoPlus_nGood$</td>
<td>$P_{B,sn}$</td>
</tr>
<tr>
<td>$\log[P_{D_s,v}/</td>
<td>\vec{p}_{D_s^+}</td>
</tr>
<tr>
<td>$P_{B,sn}$</td>
<td>the energy of the $3\pi$ system</td>
</tr>
</tbody>
</table>

culated through the partial reconstruction, also variables related to the $3\pi$ dynamics are used, like the di-pion masses which are related to intermediate resonances in the $D_s^+$ decays. The visible $B^0$ mass and $3\pi$ energy are discriminating too, since in the signal decay two neutrinos are missing, while in $D_s^+$ background events only neutral particles from $D_s^+$ decay are not detected. The $3\pi$ mass has been not used, since its distribution will be used after cutting on the BDT output to look for the characteristic shape due to the $a_1$ resonance in signal decays.

The normalized distributions of the input variables are illustrated in Fig. 4.11, 4.12 and 4.13.

The training sample used for signal is the signal MC after the first cut-based selection. The training sample of background is a data sample which satisfies, besides the first cut-based selection, the following requirements:

1. $1800 \text{ MeV} < \sqrt{m_{D_s,vn}^2} < 2200 \text{ MeV}$; 
2. $M(\tau^+) < 1800$. 


4.2 Isolation of signal

The cuts on $m^2_{D_{s,vn}}$ are needed to select a region dominated by $D_s^{(*)+}$ decays. The signal starts to contribute significantly at $\sqrt{m^2_{D_{s,vn}}} > \sim 2200 \text{ MeV}/c^2$. The distribution of this reconstructed mass is shown in Fig. 4.14 for data and in Fig. 4.15 for signal MC.

The BDT is trained on half the events in each sample, and tested on the other half to check for overtraining (see below).

4.2.2.3 BDT results

In Fig. 4.16 the BDT output distributions for signal and background are plotted superimposing training and test samples. This is a check to see whether there is overtraining, which happens when an excessive tuning on the training sample is obtained during the training phase. As a result, the classification algorithm will have an excellent performance on the training sample, but a bad performance on the test sample, that should be independent from the training one. When training and testing BDT outputs are similar, this indicates that overtraining has not taken place. The output distributions for both training and test samples are similar within the errors, so no overtraining has occurred.

Figure 4.17 shows the ROC curve, that is the background rejection versus the signal efficiency obtained through the BDT. The larger the area below this curve, the better the background rejection for a given signal efficiency.

The TMVA tool also calculates the signal and background efficiencies, the signal purity $S/(S+B)$ and the signal significance (defined as $S/\sqrt{S+B}$) as function of the BDT output cut. Figure 4.18 shows a scan of efficiencies, purity and signal significance as function of the BDT. From this plot, one can say that after applying the cut $\text{BDT} > 0.1$ it is possible to get a sample with a purity greater than 60%.

The $3\pi$ mass after $\text{BDT} > 0.1$ is shown in Fig. 4.19. The gaussian shape due to the resonance decay $a_1^+(1260) \rightarrow \pi^+\pi^-\pi^+$ is evident, although the mean value is shifted to the left, due to the presence of $\tau^+ \rightarrow \pi^+\pi^-\pi^+\pi^0$ decays. The shape of the distribution is also in good agreement with the $3\pi$ mass distribution of the signal MC (see Fig. 4.20).

Since after the BDT cut 602 events survive and the purity is about 60%. This means that about 360 signal events have been isolated on a data sample of $1 \text{ fb}^{-1}$. 
Study of the data sample composition and isolation of signal

Figure 4.11: Input variables used in BDT (1).
4.2 Isolation of signal

![Graphs showing input variables used in BDT (2).]

Figure 4.12: Input variables used in BDT (2).
Study of the data sample composition and isolation of signal

Figure 4.13: Input variables used in BDT (3).
4.2 Isolation of signal

Figure 4.14: $m_{D,vn}^2$ distribution for data after first cut-based selection.

Figure 4.15: $m_{D,vn}^2$ distribution for signal MC after first cut-based selection.
Study of the data sample composition and isolation of signal

Figure 4.16: Overtraining check.

Figure 4.17: ROC curve.
4.2 Isolation of signal

Figure 4.18: Signal and background efficiencies, signal purity and signal significance in function of the BDT cut.

Figure 4.19: The $3\pi$ mass after the BDT cut (and the first cut-based selection).
Figure 4.20: The $3\pi$ mass in the signal MC sample after the first cut-based selection.
Chapter 5

Conclusion and perspectives

In this thesis it has been shown that a sample of about 360 $B^0 \rightarrow D^*^-\tau^+ (\rightarrow \pi^+\pi^-\pi^-(\pi^0)\nu_\tau)\nu_\tau$ decays with a 60% purity can be isolated from a LHCb data sample of 1 fb$^{-1}$ integrated luminosity. Extending this study to the other 2 fb$^{-1}$ data taken at LHCb in 2012 will increase the sample by a factor 3, giving at the end more than 1000 signal events, resulting in a relative statistical uncertainty of the order of 3%.

In order to obtain a 3% of statistical error on $R_{D^*}$, a very precise measurement of the normalization channel $B^0 \rightarrow D^*^-\pi^+\pi^-\pi^+$ is required. The current relative uncertainty on this decays is at the 10% level.

The study presented in this thesis is subject to a number of approximations and assumptions regarding the estimations of efficiencies and consequently the evaluation of various background contributions in the data sample. A full MC simulation for the signal decay $B^0 \rightarrow D^*^-\tau^+ (\rightarrow \pi^+\pi^-\pi^-(\pi^0)\nu_\tau)\nu_\tau$ is needed in order to give a more accurate estimation of the signal efficiency. Another important sample, necessary to understand the performance on the first cut-based selection, is $B^0 \rightarrow D^{*-}D_{(s)}^{+(\ast*)\ast}$ decays, where the $D$ mesons give three pions and anything else in the final state, that are available at the moment only at generator-level. This MC is needed also to improve the training of the BDT, since at the moment a part of data sample, that can be contaminated by signal decays, is used as background sample for the BDT training.

It has been seen that events with misidentified kaons contribute to background. These events and the ways to suppress them must be studied deeper in order to reject them without affecting significantly signal efficiency.

Moreover, the $B^0 \rightarrow D^{**-}\tau^+\nu_\tau$ decay, where $D^{**-} \rightarrow D^*-N$, which has not been
considered in this study, is included in signal and should be precisely estimated to be subtracted in future studies.

The signal yield can be measured either by counting events surviving a cut on the BDT output, or by simultaneously fitting distributions such as the BDT output and the three-pion invariant mass. The former case will give higher signal purity, therefore systematic uncertainties due to background composition would generally be small and the measurement would likely be statistically dominated. In the latter case, the statistical error will be smaller, but the templates for the distributions to be used in the fit and their possible correlations will need to be obtained on simulated or data control samples. This could result in larger systematic errors due to data-MonteCarlo disagreements. The final strategy will depend on the outcome of detailed studies aimed at finding a tradeoff between these effects.

Last but not least, the first cut-based selection should be replaced by a more efficient multivariate technique. This requires a detailed study on large and inclusive simulated samples of B decays containing a $D^{*+}$ meson and three charged pions in the final state.
Bibliography


[15] A. Crivellin, C. Greub, A. Kokulu, Explaining $B \to D\tau\nu$, $B \to D^*\tau\nu$ and $B \to \tau\nu$ in a 2HDM of type III, Phys. Rev. D86 (2012) 054014, arXiv:1206.2634 [hep-ex]


[26] G. Wormser, *Observation of the decay $B^0 \to D^{*-}\tau^+\nu_\tau$ using $\tau^+ \to 3\pi$ channel*, LHCb Working Group on Semileptonic B Decays meeting, 30/07/2014


[30] RooFit v.14, Developed by Wouter Verkerke and David Kirkby, Copyright (C) 2000-2010 NIKHEF, University of California and Stanford University (http://roofit.sourceforge.net/license.txt)