Search for the rare decay $B \rightarrow K^* \nu \bar{\nu}$ with a recoil method at the $BABAR$ experiment.

Tesi presentata da

Elisa Manoni

per il conseguimento del titolo di Dottoressa di Ricerca
Contents

Introduction

1 Theoretical and experimental knowledge of \( b \rightarrow s \, \nu \overline{\nu} \) transitions
  1.1 Brief introduction to the Standard Model
  1.1.1 Electroweak interactions
  1.2 Effective Weak Hamiltonian formalism
  1.3 \( B \rightarrow K^{(*)} \nu \overline{\nu} \) and Standard Model predictions
  1.4 \( B \rightarrow K^{(*)} \nu \overline{\nu} \) and New Physics discoveries
    1.4.1 Non-Standard Model couplings
    1.4.2 New sources of missing energy
  1.5 Current experimental status on \( B \rightarrow K^{(*)} \nu \overline{\nu} \) decays

2 The \( BABAR \) detector
  2.1 Introduction
  2.2 The PEP-II \( B \)-factory
  2.3 The tracking system
    2.3.1 The Silicon Vertex Tracker: SVT
    2.3.2 The Drift Chamber: DCH
  2.4 The Čerenkov Light Detector: DIRC
  2.5 The Electromagnetic Calorimeter: EMC
  2.6 The Solenoid Magnet
  2.7 The Instrumented Flux Return: IFR
  2.8 The trigger system

3 Analysis method and samples
  3.1 Analysis method
3.2 Data and Monte Carlo samples ........................................... 55

4 The Semi-Exclusive $B$ reconstruction .................................. 59
   4.1 Semi-Exclusive reconstruction of hadronic $B$ final states .......... 59
   4.2 Selection of the hadronic $B$ sample .......................... 64
   4.3 Determination of the branching fraction normalization .......... 67
   4.4 Semileptonic $B$ reconstruction .................................. 71

5 Reconstruction of the $B^{0(+)} \rightarrow K^{*0(+)}\nu\bar{\nu}$ decay .......... 77
   5.1 Kaon and pion reconstruction and selection .......................... 77
   5.2 $K^*$ reconstruction and selection ................................ 78
      5.2.1 Studies on the $K^*$ vertex fit .......................... 78
   5.3 Signal side pre-selection ......................................... 81
   5.4 Selection variables ............................................. 83
      5.4.1 Definition of the selection variables .................... 83
      5.4.2 Model independence .................................. 91

6 Signal yield extraction .................................................. 93
   6.1 Cut and count analysis .......................................... 93
      6.1.1 Optimization procedure and selection results .......... 94
      6.1.2 Background estimation .............................. 99
      6.1.3 Background characterization .......................... 100
   6.2 Neural Network analysis ....................................... 101
      6.2.1 Neural Network philosophy .......................... 102
      6.2.2 Pre-selection and Neural Network optimization ........ 104
      6.2.3 Neural Network output fit .......................... 106
      6.2.4 Fit to data ........................................ 115

7 Cross-checks and systematic uncertainties ......................... 119
   7.1 Monte Carlo validation and control samples .................. 119
      7.1.1 $E_{\text{extra}}$ sideband ................................ 120
      7.1.2 Double hadronic tag sample .......................... 120
      7.1.3 Neural Network validation .......................... 125
   7.2 Evaluation of systematic uncertainties ........................ 128
Introduction

The Standard Model of particle physics is devoted to the description of the fundamental particle properties and their interactions. The experimental tests performed so far have shown an exceptionally good agreement with the theoretical prediction.

The electroweak sector of the Standard Model has been tested in several ways, one of which is the measurement of the matter and anti-matter symmetry breaking, also known as CP violation. A complex phase in the $3 \times 3$ unitary matrix, which is called Cabibbo-Kobayashi-Maskawa matrix [1], [2] and relates the mass and weak eigenstates, is known to be the responsible for the CP violation. The “golden channel” to prove this mechanism in $B$ meson decays is the $B \to J/\Psi K^0_s$ process, whose study allows to measure $\sin 2\beta$, $\beta$ being a combination of CKM matrix elements incorporating the CP violating term. This experimental challenge was the main motivation behind the construction of two experiments at the $B$ Factory: BABAR at the PEP-II collider in the United States and Belle at the KEK accelerator in Japan. The first evidence of CP violation in the $B$ meson system was measured in 2001 from both experiments. The result successfully proved the correctness of the theoretical prediction. The candidate contributed to the measurement of $\sin 2\beta$ using data collected by the BABAR detector up to 2006 [3], obtaining an uncertainty at the percent level.

There are, however, many parameters whose values are not predicted in the Standard Model and effects for which the same theory seems to be incomplete. For example, the asymmetry between matter and anti-matter in the Universe can not be explained by the CP violating mechanism known at the time being. The spontaneous symmetry breaking effect by which particles acquire finite mass, has not been confirmed by the discovery of the Higgs boson and also the mass hierarchy of the matter constituents and of the interaction mediator does not have a theoretical explanation. As a consequence it is believed that testing the Standard Model with sufficient high precision, or at sufficient high energy, could show hints of New Physics beyond the standard picture.
Two scenarios are usually conceived for this purpose: direct search for New Physics state at the Large Hadron Collider, which is starting to operate and aims to investigate the TeV scales; indirect search at machine running at lower energy but with a very clean environment, which allows very precise measurements in order to find potential discrepancies with the Standard Model prediction.

This is feasible, for example, by searching for decays that are expected to be forbidden or suppressed in the Standard Model but can receive enhancements due to new physics effects. The Flavor Changing Neutral Current transitions take part to this class of processes since they are forbidden at tree level in the Standard Model and can occur only through loop or box diagrams, where New Physics can enter. The large and clean $B\bar{B}$ sample, collected by the BABAR detector at the PEP-II $B$ Factory, represents a very fertile ground where to investigate them.

This thesis presents the first BABAR search for the $B \to K^*\nu\bar{\nu}$ decay, in which a $b$ quark decays to an $s$ quark by means of a neutral current, and is based on a sample of 454 million $B\bar{B}$ pairs. Regarding New Physics discoveries, the interest on the channel under investigation is twofold: non-standard couplings can participate in the $b \to s$ transition and, due to the undetectable nature of the neutrino pair in the final state, also non-standard massless particles can contribute. The experimental method used is called “recoil technique”: it consists on the complete reconstruction of one $B$ in the event in hadronic final states and on the search, in the opposite hemisphere, for a $K^*$ accompanied by missing energy produced by the neutrino pair. The ingredients needed for the branching fraction computation are extracted with two different strategies: a cut and count analysis and an extended maximum likelihood fit to the distribution of a Neural Network output. A Bayesian approach is then exploited to determine the final result.

The present thesis has the following structure:

- In Chapter 1 a brief review of the Standard Model is given, focusing on the study of rare electroweak decays as Flavor Changing Neutral Current transitions. The Standard Model predictions and possible New Physics contributions to the $B \to K^{(*)}\nu\bar{\nu}$ branching fractions, along with the current experimental knowledges, are also presented.

- In Chapter 2 the experimental apparatus, consisting of the PEP-II $B$ Factory and the BABAR detector, is described.
• In Chapter 3 the simulated and real data samples used to extract the experimental result are presented; also a brief overview of the analysis flow is given.

• In Chapter 4 the details of the $B$ reconstruction in hadronic final states are discussed, the strategy used to determine the branching fraction normalization, consisting on the number of correctly reconstructed hadronic $B$ decays, is described.

• Chapter 5 is devoted to the analysis of the technique adopted for the signal $B$ reconstruction, in particular the selection variables used to discriminate it from background processes are also shown.

• In Chapter 6 the two approaches exploited to compute the signal selection efficiency and the signal yield estimation, the cut and count analysis and the fit to the Neural Network output, are presented.

• Chapter 7 reports the details of the tests performed to validate the analysis procedure: the strategy of the systematics uncertainty evaluation is also discussed.

• In Chapter 8, after having introduced the statistical approach adopted, the experimental result is presented; examples on how it can be exploited to constrain parameters entering New Physic models are also discussed.

An appendix, describing the selection of the particles needed in the hadronic $B$ final state reconstruction, completes the thesis.
Introduction
Chapter 1

Theoretical and experimental knowledge of $b \rightarrow s \nu \bar{\nu}$ transitions

In the theoretical description of the fundamental interactions known as Standard Model (SM), the quark-to-quark transition $b \rightarrow s \nu \bar{\nu}$, characterized by the emission of a neutrino pair, is an example of Flavor Changing Neutral Current (FCNC) process. The decays of such class are prohibited at tree level in the SM, due to the unitarity of the $3 \times 3$ matrix which relates mass and electroweak eigenstates of the interacting quarks. In this framework, FCNC transitions can happen only through higher order diagrams: as a consequence they are particularly appealing since New Physics (NP) can enter loops or box diagrams and measurements of physical quantities, such as branching fractions (BF), can give confirmation of the SM predictions or constraint parameters which define NP models.

In Sec. 1.1 the basic concepts of the SM and the electroweak interaction are given. A particularly convenient approach to study FCNC decays is the Effective Weak Hamiltonian formalism, presented in Sec. 1.2, in which SM and NP contributions can be easily factorized. A review of SM expectations and NP effects related to the $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays are presented in Secs. 1.3 and 1.4, while the current experimental status is summarized in Sec. 1.5.

1.1 Brief introduction to the Standard Model

The Standard Model of particle physics describes the properties and the interactions between the elementary particles. These are classified in two categories: half-integer spin particles called fermions, which constitute the matter, and integer-spin bosons, that me-
**Theoretical and experimental knowledge of $b \to s \bar{\nu} \nu$ transitions**

Table 1.1: Properties of the elementary particles [4].

<table>
<thead>
<tr>
<th>Particle</th>
<th>Mass (MeV/$c^2$)</th>
<th>Charge (e units)</th>
<th>Spin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fermions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e$</td>
<td>0.511</td>
<td>-1</td>
<td>1/2</td>
</tr>
<tr>
<td>$\nu_e$</td>
<td>$\sim 0$</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>$\mu$</td>
<td>106</td>
<td>-1</td>
<td>1/2</td>
</tr>
<tr>
<td>$\nu_\mu$</td>
<td>$\sim 0$</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>$\tau$</td>
<td>1,777</td>
<td>-1</td>
<td>1/2</td>
</tr>
<tr>
<td>$\nu_\tau$</td>
<td>$\sim 0$</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>$d$</td>
<td>-</td>
<td>-1/3</td>
<td>1/2</td>
</tr>
<tr>
<td>$u$</td>
<td>-</td>
<td>2/3</td>
<td>1/2</td>
</tr>
<tr>
<td>$s$</td>
<td>-</td>
<td>-1/3</td>
<td>1/2</td>
</tr>
<tr>
<td>$c$</td>
<td>-</td>
<td>2/3</td>
<td>1/2</td>
</tr>
<tr>
<td>$b$</td>
<td>-</td>
<td>-1/3</td>
<td>1/2</td>
</tr>
<tr>
<td>$t$</td>
<td>-</td>
<td>2/3</td>
<td>1/2</td>
</tr>
<tr>
<td>Bosons</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$g$</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$W$</td>
<td>80,403</td>
<td>$\pm 1$</td>
<td>1</td>
</tr>
<tr>
<td>$Z^0$</td>
<td>91,188</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The fermion family is formed by *quarks* and *leptons*: the firsts are subject to both interactions described in the SM, the *strong* and the *electroweak* interactions, while leptons participate only in electroweak processes. The strong interaction is mediated by 8 spin-1 bosons called *gluons*, while $Z^0$, $W^\pm$ and $\gamma$ are responsible for the electroweak transitions. A summary of some elementary particle properties is given in Tab. 1.1, for each fermion a corresponding antiparticle, with opposite quantum numbers, exists.

The SM is a non abelian gauge theory [5] associated to the symmetry group:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y.$$  \hfill (1.1)

The $SU(3)_C$ term [6] accounts for the strong interaction theory of Quantum Chromodynamics (QCD) and $C$ is the color quantum number. The electromagnetic and weak interactions are unified under the electroweak interaction, whose gauge group is $SU(2)_L \otimes U(1)_Y$ [7]. The quantum number $L$ refers to the fact that charged-current weak force (i.e. those mediated by $W^\pm$) couples only with left-handed fermions, while $Y$ is the weak hypercharge defined, by means of the electric charge $Q$ and the third component of the weak
1.1 Brief introduction to the Standard Model

isospin $T_3$, as $Y = 2(Q - T_3)$.

The fermion fields are organized in multiplets $\Psi(x_\mu)$, where $x_\mu$ is the space-time coordinate. For the gauge invariance [8], $\Psi(x_\mu)$ must be unchanged under the following transformation:

$$\Psi(x_\mu) \rightarrow e^{i\alpha_a(x_\mu)T_a}\Psi(x_\mu), \quad (1.2)$$

where $\alpha_a$ define a set of real functions and $T_a$ are the generators of the Lie algebra associated to the gauge group. These correspond to $SU(3)_C$ for QCD and $SU(2)_L \otimes U(1)_Y$ for the electroweak sector. Besides the fermion fields, also the Lagrangian density ($\mathcal{L}$) must be invariant under the gauge transformation. The function $\mathcal{L}$ contains a kinetic term of the form:

$$\mathcal{L} \propto \bar{\Psi}i\gamma^\mu \partial_\mu \Psi. \quad (1.3)$$

To preserve the invariance, the derivative operator $\partial$ must be transformed into the covariant derivative $\mathcal{D}$ as:

$$\partial_\mu \rightarrow \mathcal{D}_\mu = \partial_\mu + igT^aA^a_\mu, \quad (1.4)$$

where $g$ is a universal coupling constant, determining the strength of the interaction, and the gauge fields $A^a_\mu$ transform as:

$$A^a_\mu \rightarrow A^a_\mu - \frac{1}{g}\partial_\mu \alpha_a(x). \quad (1.5)$$

The $A^a_\mu$ correspond to the set of gauge bosons that mediate the interaction: the first immediate consequence of requiring the gauge invariance is the presence of particles whose task is to carry the interaction matter is subject to.

The way in which the $T^a$ are represented determines the interaction properties of each fermion. In a given transformation group, fermions that do not take part to the interaction are represented as singlets; on the other hand fermion fields that are multiplets transform under the gauge representation. Concerning the QCD, the six quarks carry the color charge and are associated to fields that transform under $SU(3)_C$; in the same representation, the lepton fields are singlets and do not interact via QCD. In the electroweak group, the fermions are organized in left-handed doublets and right-handed singlets as shown in Tab. 1.2. The antiparticles are grouped in right-handed doublets and left-handed singlets.

The Standard Model Lagrangian density can be factorized as:

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{fermions}} + \mathcal{L}_{\text{Yang-Mills}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \quad (1.6)$$
describing respectively: the kinetic energy and the couplings to the gauge bosons of the fermions ($\mathcal{L}_{\text{fermions}}$); the kinetic and the self-coupling terms associated to the gauge bosons ($\mathcal{L}_{\text{Yang-Mills}}$), the spontaneous symmetry breaking terms by which the bosons ($\mathcal{L}_{\text{Higgs}}$) and the fermions ($\mathcal{L}_{\text{Yukawa}}$) acquire finite mass thanks to the Higgs mechanism [9]. The price to be paid to incorporate the mass terms and maintain the gauge invariance is the introduction of a scalar Higgs field with a non-zero vacuum expectation value. This, along with the coupling constants of the electroweak interaction, allows to predict the $W^\pm$ and $Z^0$ masses, while no expected values are given for the mass associated to the matter fields which needs to be determined experimentally.

In the following the electroweak interaction, which is responsible for $b \rightarrow s \nu \bar{\nu}$ transitions, will be discussed.

Table 1.2: Fermionic multiplets in the $SU(2)_L \otimes U(1)_Y$ gauge group. Leptons and quarks are divided according to the flavor family, the primed quarks refers to the electroweak eigenstates, not corresponding to the mass eigenstates, as discussed in Sec. 1.1.1.

<table>
<thead>
<tr>
<th>Leptons</th>
<th>Quarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_e$</td>
<td>$u'_L$</td>
</tr>
<tr>
<td>$e^-$</td>
<td>$d'_L$</td>
</tr>
<tr>
<td>$\mu^-$</td>
<td>$s'_L$</td>
</tr>
<tr>
<td>$\tau^-$</td>
<td>$t'_L$</td>
</tr>
</tbody>
</table>

### 1.1.1 Electroweak interactions

The Lagrangian density accounting for electroweak transitions ($\mathcal{L}_{\text{EW}}^{\text{int}}$) can be written as the sum of two contributions:

$$\mathcal{L}_{\text{EW}}^{\text{int}} = \mathcal{L}_{\text{CC}} + \mathcal{L}_{\text{NC}},$$

(1.7)

which refer to the $W^\pm$- and $Z^0$-mediated processes, respectively.

The charged current interaction has the form:

$$\mathcal{L}_{\text{CC}} = \frac{g_2}{2\sqrt{2}} (J^+_\mu W^{+\mu} + J^-_\mu W^{-\mu}),$$

(1.8)

where the positive current is defined as:

$$J^+_\mu = \bar{u}'_i \gamma_\mu (1 - \gamma^5) d'_i + \bar{d}'_i \gamma_\mu (1 - \gamma^5) e_i,$$

(1.9)
and $J_{\mu}^{-}$ is its hermitian conjugate. The $u'_i$ ($\nu_i$) are the up-sector quarks (leptons) of the doublets in Tab. 1.2, while $d'_i$ ($e_i$) belong to the down-sector, and $g_2$ is the $SU(2)_L$ coupling constant.

The neutral current interaction is defined as:

$$
\mathcal{L}_{NC} = - e J_{\mu}^{em} A_\mu + \frac{g_2}{2 \cos \theta_W} J_{\mu}^{0} Z_\mu ,
$$

(1.10)

here $e$ is the coupling constant of the electromagnetic interaction (QED) and $\theta_W$ the Weinberg angle which enters the mixing of the $Z^0$ and $\gamma$ bosons in the spontaneous symmetry breaking mechanism. The currents assume the following form:

$$
\begin{align*}
J_{\mu}^{em} &= \sum_f Q_f \bar{f} \gamma_\mu f , \\
J_{\mu}^{0} &= \sum_f \bar{f} \gamma_\mu (v_f a_f^5) f ,
\end{align*}
$$

(1.11)

where:

$$
v_f = T_3^f - 2 Q_f \sin^2 \theta_W , \quad a_f = T_3^f ,
$$

(1.12)

with $Q_f$ the charge of the left-handed fermion $f_L$. The charged current is purely $V - A$ (where $V$ stands for vector-like and $A$ for axial-like) while the neutral one contains both $V - A$ and $V + A$ terms.

The electroweak quark eigenstates that couple to the gauge bosons do not correspond to the mass eigenstates, and the transformation between them is defined by two unitary matrices $U_U$ and $U_D$

$$
\begin{pmatrix}
  u' \\
  c' \\
  t'
\end{pmatrix} = U_U
\begin{pmatrix}
  u \\
  c \\
  t
\end{pmatrix} , \quad \begin{pmatrix}
  d' \\
  s' \\
  b'
\end{pmatrix} = U_D
\begin{pmatrix}
  d \\
  s \\
  b
\end{pmatrix} .
$$

(1.13)

The product $V_{\text{CKM}} = U_D^T U_U$ defines a unitary non diagonal matrix called *Cabibbo-Kobayashi-Maskawa matrix* [1], [2]:

$$
V_{\text{CKM}} = \begin{pmatrix}
  V_{ud} & V_{us} & V_{ub} \\
  V_{cd} & V_{cs} & V_{cb} \\
  V_{td} & V_{ts} & V_{tb}
\end{pmatrix} ,
$$

(1.14)

and contains the quark mixing couplings. The nine elements $V_{ij}$ can be written in terms of three angles and a complex phase, which is responsible for the violation of the $CP$ symmetry in electroweak interactions. The theory does not make any prediction about the values of the $V_{ij}$, which are free parameters of the SM, but it establishes the unitarity of the CKM matrix which has two important implications.
Theoretical and experimental knowledge of $b \to s \nu \bar{\nu}$ transitions

The first is that, from the experimental point of view, the CKM elements can be measured independently and used to prove the unitarity of the matrix and the correctness of the Standard Model expectations. One of the constraints imposed by the unitarity is the following:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.15}$$

The equation terms above enter $B$ mesons transitions, for this reason the $B$ meson system is an optimal framework to test the electroweak sector of the SM. The unitary triangle in Fig. 1.1 is the graphical representation of Eq. 1.15: all the quantities defining it, except $V_{ud}$ and $V_{cd}$ that were measured by fixed target experiments [4], can be studied at the $B$ Factories [10]. Using the Wolfenstein parametrization [11] of the CKM matrix given by:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4), \tag{1.16}$$

the upper vertex of the unitary triangle is defined by the coordinates $(\bar{\rho}, \bar{\eta})$:

$$\bar{\rho} = \rho \left(1 - \frac{\lambda^2}{2}\right), \quad \bar{\eta} = \eta \left(1 - \frac{\lambda^2}{2}\right), \tag{1.17}$$

while the three angles in Fig. 1.1 are:

$$\alpha \equiv \arg \left[\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta \equiv \arg \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma \equiv \arg \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right] \equiv \pi - \alpha - \beta. \tag{1.18}$$

The magnitude of the $CP$ violation is related to the displacement of the $(\bar{\rho}, \bar{\eta})$ position from the real axis: as a consequence measurements of the unitary triangle angles are a proof of the SM prediction for the $CP$ violation. Other $CP$-violating and $CP$-conserving quantities are measured in $B$ meson decays and they are combined in the so-called Unitary triangle fits [12] [13]. The result of the fit with the approach proposed in Ref. [12] is shown in Fig. 1.2: as can be noticed, the intersection of the allowed regions for the physical quantities involved in the fit determines the apex of the unitary triangle. An eventual discrepancy between two or more experimental measurements would be a hint of some unpredicted effects in the SM picture.

A second implication of the unitarity of the CKM matrix, relevant to this work, is the absence of Flavor Changing Neutral Current at tree level in the Standard Model. This can be shown by considering the fermion-boson interaction terms of the Lagrangian (Eqs. 1.8
1.1 Brief introduction to the Standard Model

Figure 1.1: The unitary triangle (a), with all the sides rescaled to $|V_{cb} V_{cd}|$ (b).

Figure 1.2: Limits on $(\hat{\rho}, \hat{\eta})$ combining direct and indirect measurements of the angles. Every colored area represents the total uncertainty on the physical quantity.
Theoretical and experimental knowledge of $b \to s \nu \bar{\nu}$ transitions

and 1.10) and Eq. 1.13; the charged current interaction is described by:

$$
\left( \begin{array}{c} \bar{\nu} \\ \bar{c} \\ \bar{t} \end{array} \right) \gamma^\mu (1 - \gamma^5) \left( \begin{array}{c} d' \\ s' \\ b' \end{array} \right) = \left( \begin{array}{c} \bar{\nu} \\ \bar{c} \\ \bar{t} \end{array} \right) \gamma^\mu (1 - \gamma^5) U_D^\dagger U_D \left( \begin{array}{c} d \\ s \\ b \end{array} \right) 
$$

$$
= \left( \begin{array}{c} \bar{\nu} \\ \bar{c} \\ \bar{t} \end{array} \right) \gamma^\mu (1 - \gamma^5) V_{\text{CKM}} \left( \begin{array}{c} d \\ s \\ b \end{array} \right), \quad (1.19)
$$

and $V_{\text{CKM}}$ defines the probability for an upper-sector quark to decay to a down-sector quark. In the case of the neutral current, for example for the down sector, it is:

$$
\left( \begin{array}{c} \bar{d} \\ \bar{s} \\ \bar{b} \end{array} \right) \gamma^\mu (v_f - a_f \gamma^5) \left( \begin{array}{c} d' \\ s' \\ b' \end{array} \right) = \left( \begin{array}{c} \bar{d} \\ \bar{s} \\ \bar{b} \end{array} \right) \gamma^\mu (v_f - a_f \gamma^5) U_D^\dagger U_D \left( \begin{array}{c} d \\ s \\ b \end{array} \right), \quad (1.20)
$$

and, since $U_D$ is unitary by definition, transitions like $b \to s$ are not allowed at tree level but can occur only through higher order diagrams (as will be shown in the next section).

Historically, Glashow, Iliopoulos, and Maiani [14] postulated the existence of the charm quark and extended the formalism of the Cabibbo mixing angle [1] to the second quark family, by introducing a $2 \times 2$ unitary matrix which governs the transitions between the first two quark families. In this scheme, called Glashow-Iliopoulos-Maiani (GIM) mechanism, the FCNC are forbidden at tree level, with the same argument used for the CKM matrix, that can be regarded as the $3 \times 3$ extension of the previously mentioned $2 \times 2$ unitary matrix. As will be shown in the next section, the disparity of the quark masses breaks the GIM mechanism and makes the FCNC interactions, otherwise forbidden at every order, happen through second order diagrams.

The fact that loops and boxes enter the FCNC transitions makes such processes particularly interesting since New Physics particles can contribute, i.e. by the presence of non-standard particles in the loops. A representation that easily allows to distinguish between SM and NP contributions in weak decays is the Effective Weak Hamiltonian formalism, discussed in the next section.

### 1.2 Effective Weak Hamiltonian formalism

Within the Standard Model, FCNC decays are governed by:

- the CKM matrix elements which describe the weak charged current interaction;
• the GIM mechanism [14] which prohibits the existence of FCNC at tree level, while predicts such transitions at higher order if the intrinsic quark masses are degenerate;

• the QCD asymptotic freedom [15], by which the strong interaction contribution to weak decays is calculated;

• local operators parametrizing the FCNC diagrams.

The formulation that accounts for all the previous elements is called Effective Weak Hamiltonian formalism and its basic principle is that every hadron weak decay can be described by the effective Hamiltonian:

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \sum_i V_{\text{CKM}}^i C_i(\mu) Q_i, \quad (1.21)$$

$G_F$ being the Fermi constant, $V_{\text{CKM}}^i$ the CKM elements, $Q_i$ the local operators that enter the decay and $C_i(\mu)$ the Wilson Coefficients [16], which depend on the energy scale $\mu$ and, along with the CKM elements, describe the strength with which the associated operator enters $\mathcal{H}_{\text{eff}}$. The previous expansion, called Operator Product Expansion (OPE), is a sum of effective vertices multiplied by effective couplings.

In this framework, physical observables can be expressed as a function of Wilson coefficients and local operators. An example is the decay amplitude for a given exclusive decay of the $B$ meson to a final state $F$:

$$A(B \to F) = \langle F | \mathcal{H}_{\text{eff}} | B \rangle = \frac{G_F}{\sqrt{2}} \sum_i V_{\text{CKM}}^i C_i(\mu) \langle F | Q_i(\mu) | B \rangle, \quad (1.22)$$

$\langle F | Q_i(\mu) | B \rangle$ being the hadronic matrix elements of the local operators.

The power of the OPE is to allow the separation of long- and short-distance contributions: the first is incorporated in the matrix elements $\langle Q_i(\mu) \rangle$ and should be calculated with a non-perturbative method, the latter is included in the perturbative computation of the coefficients $C_i(\mu)$. The energy scale $\mu$ defines the arbitrary border line between long- and short-distance effects and it is usually chosen as the mass of the decaying quark. Both the couplings and the matrix elements depend on the scale $\mu$ and on the renormalization scheme, while the final result should be independent from such choices.

The $C_i(\mu)$ encompass the dependence on the mass of all the heavy particles in the effective vertices such as the top quark, the $W$ and $Z^0$ bosons, the Higgs and any other heavy particles predicted by NP models, i.e. supersymmetric partners of the standard
particles, Kaluza-Klein modes in extra dimension models [17], etc. The dependence on the internal particle masses is made explicit when evaluating the box and penguin diagrams. At this stage also the QCD effects should be calculated and they define the $C_i$ $\mu$-dependence. In $B$ decays $\mu = m_b \ll m_{W,Z}$ and the smallness of the QCD coupling constant should be compensated by the large logarithms $\ln(m_W/\mu)$. The adopted renormalization scheme defines the resummation of terms like $\alpha_s^n(\ln(m_W/\mu))^n$ at all $\alpha_s$ orders. In this way the so-called renormalization group improved perturbative expansion for the Wilson coefficients are determined [18].

Also a proper evaluation of the matrix elements $\langle Q_i(\mu) \rangle$ is needed in order to calculate a $\mu$- and renormalization scheme-independent decay amplitude. Since non-perturbative model are exploited, it is very challenging to find a proper cancellation of the unwanted dependencies. As a consequence, considering the decay amplitude of an exclusive decay, the main theoretical uncertainties is associated to the matrix elements $\langle F_i|Q_i(\mu)|B \rangle$.

The theoretical knowledge improves when considering inclusive decays. They are defined as the class of all the decay modes which share the same quark level electroweak transition but differ for the QCD contributions due to the spectator quark. Given a set of such final states $X$, the amplitude is expressed as:

$$A(B \to X) = \frac{G_F}{\sqrt{2}} \sum_i V_{C_i}^i C_i(\mu) \langle x | Q_i(\mu) | B \rangle.$$  \hspace{1cm} (1.23)

The resulting branching ratio can be written in a form known as Heavy Quark Expansion (HQE) [19]:

$$B(B \to X) = B(b \to q) + O\left(\frac{1}{m^2_b}\right),$$  \hspace{1cm} (1.24)

where in the first term the $B$ decay is modeled by the $b$ quark decay in the spectator model and the effective operators can be computed in perturbation theory with the consequent exact cancellation of the $\mu$ and renormalization scheme dependence. The next to leading order terms are suppressed at least by the inverse of $m_b$ at the second power. This peculiarity makes inclusive decays easier to treat from the theoretical point of view, even if they are more difficult to measure with respect to the inclusive processes. For example, the presence of a high momentum photon, makes the $B \to X_s \gamma$ decays experimentally accessible, while it is much more challenging to search for $B \to X_s \nu \bar{\nu}$ since, other than combinations of kaons and pions, nothing else is reconstructed and the measurement suffers from large background contamination.
In order to compute the decay amplitudes in Eqs. 1.22 and 1.23 and properly account for effects at different energy scales, several steps are needed:

- the Wilson coefficients $C_i(\mu_W)$ are evaluated at the energy scale $\mu_W = m_W$ to the desired $\alpha_s$ order. Since the logarithmic terms $\ln(\mu_W/m_W)$ are small, a perturbation calculation can be performed. This step amounts to the matching of the full theory onto the effective theory, where in the first all the particles appear as dynamical degrees of freedom while the second is a five quark effective theory in which the top, $W$, and $Z^0$ fields, along with those associated to heavy NP quanta with mass greater than $m_W$, are integrated out. Practically, this is accomplished by calculating the amplitude and the operator matrix elements in the full theory and finding the value of $C(\mu_W)$ for which the full theory amplitude is equal to the effective one.

- all the logarithmic terms are resummed and reliable expressions of the Wilson coefficients are found, by means of the renormalization group methods [20], at this stage the energy scale $\mu_W$ is decreased to the appropriate low energy scale $\mu$ and the coefficients can be factorized as:

$$C(\mu_i) = \eta^i_{\text{QCD}}(\mu, \mu_W)C_i(\mu_W). \quad (1.25)$$

When calculating the box and loops diagrams, the $C(\mu_W)$ turn out to depend also on $x_i = m_i/m_W$ where the index $i$ refers to internal quark;

- finally the hadronic matrix elements are computed by means of a non-perturbative theory and normalized to the $\mu$ scale.

At the end of this procedure, the decay amplitude, i.e. for an exclusive decay, can be rearranged as:

$$\mathcal{A}(B \rightarrow F) = \langle F|\mathcal{H}_{\text{eff}}|B\rangle = \sum_i B_i V^i_{\text{CKM}}\eta^i_{\text{QCD}}(\mu, \mu_W)F_i(x_i). \quad (1.26)$$

The $B_i$ are non-perturbative factors related to the hadronic matrix elements of the local operators and, as discussed before, carry the largest theoretical uncertainty; $\eta^i_{\text{QCD}}(\mu, \mu_W)$ hold the short distance information and the evolution from the high scale $\mu_W$ to the low scale $\mu$; the $F_i(x_i)$ arise from the evaluation of the loop and box diagrams which gives the Wilson coefficients at high energy scale. Such functions reduce, in the SM, to the Inami-Lim functions [21].
The formalism of Eq. 1.26 can be used to understand why the quark mass disparity is necessary to the presence of FCNC in the GIM mechanism. The CKM factors and the Inami-Lim functions enter the FCNC $B$ decay amplitudes as:

$$
\sum_{i = u, c, t} V^*_{ib}V_{id}F(x_i) \text{ or } \sum_{i, j = u, c, t} V^*_{ib}V^*_{jd}F(x_i, x_j). \tag{1.27}
$$

The explicit expression of the first term is:

$$
V^*_{ub}V_{ud}F(x_u) + V^*_{cb}V_{cd}F(x_c) + V^*_{tb}V_{td}F(x_t). \tag{1.28}
$$

If the quarks of a given charge would have the same mass, the previous would take the form:

$$
(V^*_{ub}V_{ud} + V^*_{cb}V_{cd} + V^*_{tb}V_{td})F(x), \tag{1.29}
$$

that is identically zero because of the unitarity of the CKM matrix ($V^*_{ub}V_{ud} + V^*_{cb}V_{cd} + V^*_{tb}V_{td} = 0$). The fact that in nature the mass symmetry is broken violates the GIM mechanism and allows FCNC decays at higher orders. For small $x_i \ (i \neq t)$ the Inami-Lim functions are hardly suppressed [22], while the top quark contribution does not suffer from such effect. This makes the $B$ decays particularly appealing since the top quark is the main mediator in this case. In $D$ meson decays, for example, lighter quarks are dominant in the loops and the GIM suppression is much more effective.

In the same way as the top quark, also NP particles can enter the loops and another virtue of the OPE is the inclusion of NP effects, that can show up in the different terms of the effective Hamiltonian: new complex phase in the CKM matrix elements $V_i$, new contribution to the second order diagrams not proportional to the CKM elements entering the SM terms, new local operators, new contributions to the box and penguin diagrams modifying the Wilson coefficients. These effects would change the magnitude of the $CP$ violation and some decay amplitudes with respect to the Standard Model expectations. Since also the SM physics is described by the OPE formalism, a comparison between such predictions and experimental measurements can distinguish between the expected SM part and NP terms.

In the next sections SM and NP predictions for $B \rightarrow K^{(*)}\nu\bar{\nu}$ are discussed (Secs. 1.3 and 1.4) along with the current experimental status (Sec. 1.5).
1.3 \( B \to K^{(*)}\nu\bar{\nu} \) and Standard Model predictions

The SM Feynman diagrams for the quark level \( b \to s \nu\bar{\nu} \) transition are shown in Fig 1.3: the decay might happen through a \( W \)-mediated box and a \( Z^0 \)-penguin diagram. The SM effective Hamiltonian takes the form:

\[
\mathcal{H}_{\text{eff}} = -\frac{G_F}{\sqrt{2}} V_{ts}^* V_{tb} C_L^\nu Q_L^\nu,
\]

(1.30)

\( Q_L^\nu \) being the SM basis operators, that in the formalism of Ref. [23] is:

\[
Q_L^\nu = \frac{e^2}{4\pi^2} \bar{s}L\gamma_\mu b_L \gamma^\mu (1 - \gamma^5)\nu.
\]

(1.31)

The SM Wilson coefficient \( C_L^\nu \) depends on \( x_t = m_t^2/m_W^2 \) through the following relation:

\[
C_L^\nu_{\text{SM}} = \frac{4B(x_t) - C(x_t)}{\sin^2 \theta_W}.
\]

(1.32)

The two contributions on the numerator are the Inami-Lim functions [21] [22] accounting for the box \((B(x_t))\) and the penguin diagrams \((C(x_t))\) of Fig. 1.3. At the leading order they can be written as:

\[
B(x_t) = \frac{1}{4} \left[ \frac{x_t}{1 - x_t} + \frac{x_t \ln x_t}{(1 - x_t)^2} \right],
\]

\[
C(x_t) = \frac{x_t}{8} \left[ \frac{x_t - 6}{x_t - 1} + \frac{3x_t + 2}{(x_t - 1)^2} \ln x_t \right].
\]

(1.33)

For \( m_t = 174 \text{ GeV}/c^2 \) and \( m_W = 80.4 \text{ GeV}/c^2 \), \( C_L^\nu_{\text{SM}} \sim -6.6 \). In addition, in the SM no contributions from the right-handed operator \( Q_R^\nu \) are expected and the corresponding Wilson coefficient is \( C_R^\nu_{\text{SM}} \sim 0 \). The measurement of a physical quantity, such as the decay branching fraction that depends on the Wilson coefficients, allows to test the goodness of the SM prediction.
In Ref. [23] the differential decay rates for $B \to K^{(*)}\nu \bar{\nu}$ as a function of the $\nu \bar{\nu}$ invariant mass $m_{\nu \bar{\nu}}$ are computed, having defined $m_{\nu \bar{\nu}}^2 = (m_B^2 + m_H^2 - 2m_B E_H)$, where $H = K, K^*$, and $s = m_{\nu \bar{\nu}}^2/m_B^2$.

For the $B \to K\nu \bar{\nu}$ decay the dilepton spectrum is:

$$
\frac{d\Gamma(B \to K\nu \bar{\nu})}{ds} = \frac{G_F^2 \alpha^2 m_B^5}{256\pi^5} |V_{ts}V_{tb}|^2 \lambda^{3/2}_K(s) f_+(s) |C_L^\nu + C_R^\nu|^2.
$$

(1.34)

with $\lambda_K(s) = 1 + r_H^2 + s^2 - 2s - 2r_H - 2sr_H, r_H = m_H^2/m_B^2$ and $f_+(s)$ the form factor, as determined in Ref. [24]. Integrating over all the $s$-range, the branching fraction is:

$$
\mathcal{B}(B \to K\nu \bar{\nu}) = (3.8_{-0.6}^{+1.2}) \times 10^{-6} \left| \frac{C_L^\nu + C_R^\nu}{C_L^\nu|_{\text{SM}}} \right|^2,
$$

(1.35)

with uncertainties coming from the form factor. As already mentioned, in the Standard Model $C_R^\nu \sim 0$ and the expected branching fraction for $B \to K\nu \bar{\nu}$ is:

$$
\mathcal{B}(B \to K\nu \bar{\nu}) = (3.8_{-0.6}^{+1.2}) \times 10^{-6}.
$$

(1.36)

The $B \to K^*\nu \bar{\nu}$ decay rate depends on both $C_L^\nu + C_R^\nu$ and $C_L^\nu - C_R^\nu$ as:

$$
\frac{d\Gamma(B \to K^*\nu \bar{\nu})}{ds} = \frac{G_F^2 \alpha^2 m_B^5}{1024\pi^5} |V_{ts}V_{tb}|^2 \lambda^{1/2}_{K^*}(s) \left\{ \frac{8s\lambda_{K^*}(s)V^2\left(1 + \sqrt{r_{K^*}}\right)^2}{\left(1 + \sqrt{r_{K^*}}\right)^2} |C_L^\nu + C_R^\nu|^2 
\right. 
+ \frac{1}{r_{K^*}} \left[ (1 + \sqrt{r_{K^*}})^2 (\lambda_{K^*}(s) + 12r_{K^*}A_1(s)) + \frac{\lambda_{K^*}^2(s)A_2^2(s)}{1 + \sqrt{r_{K^*}}} \right] 
- 2\lambda_{K^*}(s)(1 - r_{K^*} - s)A_1(s)A_2(s) \left| C_L^\nu - C_R^\nu \right|^2 \right\}.
$$

(1.37)

with $V(s), A_1(s)$, and $A_2(s)$ form factors governing the process [24]. The integrated BF results to be:

$$
\mathcal{B}(B \to K^*\nu \bar{\nu}) = (2.4^{+1.0}_{-0.5}) \times 10^{-6} \left| \frac{C_L^\nu + C_R^\nu}{C_L^\nu|_{\text{SM}}} \right|^2 + (1.1^{+0.3}_{-0.2}) \times 10^{-5} \left| \frac{C_L^\nu - C_R^\nu}{C_L^\nu|_{\text{SM}}} \right|^2,
$$

(1.38)

with a SM expectation value of:

$$
\mathcal{B}(B \to K^*\nu \bar{\nu}) = (1.3^{+0.4}_{-0.3}) \times 10^{-5}.
$$

(1.39)

The experimental reaches of the $B \to K^{(*)}\nu \bar{\nu}$ searches start to be close to the SM prediction of Eqs. 1.36 and 1.39. As will be discussed in Sec. 1.5, all the current measurements are in agreement with the SM.
1.4 \( B \to K^{(*)}\nu\overline{\nu} \) and New Physics discoveries

The NP can enter the \( b \to s \nu\overline{\nu} \) transition in several ways. New particles can mediate the interaction and enter the loop and the box diagrams, as examples non-SM \( Z^0 \) couplings and the effect of the Kaluza-Klein particles in the Extra Dimensions framework will be treated. Moreover, since the final state includes two neutrinos which cannot be reconstructed from the experimental apparatus, other sources of invisible contributions can enhance the decay branching fractions, such as light dark matter or unparticles.

1.4.1 Non-Standard Model couplings

In generic extensions of the SM [23], particles \( X \) with mass \( m_X > m_{Z^0} \), can mediate weak transitions. Three kinds of operators can enhance the \( b \to s\nu\overline{\nu} \) decay rate:

- **four fermion operators** [25] that can enter by tree level or one loop diagrams and have dimension six in the new physics scale \( m_X \); as a consequence their Wilson coefficients are expected to be suppressed by a factor \( 1/m_X^2 \);

- **magnetic operators** with dimension five and related to Wilson coefficients suppressed by a factor \( 1/m_X \);

- **new FCNC \( Z^0 \) couplings** with non-standard contributions to the SM operator (Eq. 1.10) \( \overline{b}_L\gamma^\mu s_L Z_\mu \) and new coupling with the right handed component \( \overline{b}_R\gamma^\mu s_R Z_\mu \) with a dependence of the decay rates on \( C^*_R \); they are generated by replacing particles in the loop of Fig. 1.3 with their supersymmetric partners and have dimension four, their Wilson coefficients are not suppressed by the NP scale \( m_X \) and they turn out to be particularly interesting with respect to the other two effects.

The enhancements on the BF, both for \( B \to K\nu\overline{\nu} \) and \( B \to K^*\nu\overline{\nu} \), can be of up to a factor 10 with respect to Eqs. 1.36 and 1.39.

The Wilson coefficient \( C^*_L \) is modified also in **Extra Dimension** models [26]. Some extensions of the SM predict the existence of extra dimensions which allows the unification of the gravitational and the gauge interactions at weak scales [17]. Excited states of low energy gravitons would manifest as the so-called Kaluza-Klein particles, whose mass depends on the compactification radius \( R \) that defines the extra dimension. Such states can mediate the \( B \to K^{(*)}\nu\overline{\nu} \) transitions leading to a dependence on the compactification
radius of the Wilson coefficients [26]. In fact, to the numerator of Eq. 1.32, the following term should be added:

\begin{equation}
C(x_t, R) = \frac{x_t(7 - x_t)}{16(x_t - 1)} - \frac{\pi m_W R x_t}{16(x_t - 1)^2} [3(1 + x_t) J(R, -1/2) + (x_t - 7) J(R, 1/2)] ,
\end{equation}

with:

\begin{equation}
J(R, \alpha) = \int_0^1 dy \ y^\alpha \left[ \coth(\pi m_W R \sqrt{y}) - x_t^{1+\alpha} \coth(\pi m_t R \sqrt{y}) \right].
\end{equation}

The dependence of the $B \rightarrow K^{*} \nu \bar{\nu}$ branching fractions on the compactification radius is shown in Fig. 1.4 [26], and a measurement of $B(B \rightarrow K^{(*)} \nu \bar{\nu})$ can be used to constrain the allowed physical range for $R$.

### 1.4.2 New sources of missing energy

From the experimental point of view, the search for $B \rightarrow K^{(*)} \nu \bar{\nu}$ consists on reconstructing a strange meson and look for missing energy associated to the undetected neutrino pair. If a model independent analysis is performed (i.e. without exploiting the kinematics of the $\nu \bar{\nu}$ pair) the result can be interpreted as the search for a $B$ final state with a kaon and missing or invisible energy. In New Physics models, several mechanisms predict the existence of undetected particles, other than neutrinos, that can contribute to $B \rightarrow K^{(*)} + invisible$. In the following, two examples are given: light dark matter candidates and unparticles.

In Ref. [27], the presence of pair of Weakly Interacting Massive Particles (WIMPs) in a $B$ final state together with a strange meson is discussed. The analysis is performed in the singlet scalar WIMPs model framework [28], with the assumption that the mass of the dark matter scalar field $S$ is smaller than the SM Higgs mass. The diagram that
contribute to $B \to K^{(*)} + \text{invisible}$ by WIMPs production is shown in Fig. 1.5. The $B \to K^{(*)}\nu\bar{\nu}$ decay rates as a function of the $SS$ invariant mass $\hat{s}$ can be written as:

$$
\frac{d\Gamma(B \to K^{(*)}\nu\bar{\nu})}{d\hat{s}} = \frac{x_t^2 C_{DM}^2 f_{0}(\hat{s})^2}{512\pi^3} \frac{I_{K}(\hat{s}, m_S) m_B^2 (m_B^2 - m_K^2)^2}{m_B^2 (m_b - m_s)^2},
$$

$$
\frac{d\Gamma(B \to K^{(*)}\nu\bar{\nu})}{d\hat{s}} = \frac{x_t^2 C_{DM}^2 A_0(\hat{s})^2}{512\pi^3} \frac{I_{K^{(*)}}(\hat{s}, m_S) h(\hat{s})}{m_B}.
$$

(1.42)

The Wilson coefficient that enter the Eq. 1.42 is defined as:

$$
C_{DM} = \frac{\lambda}{m_h^2} \frac{3g_W^2 V_{ts} V_{tb}}{32\pi^2} x_t,
$$

(1.43)

The form factor $f_{0}(\hat{s})$ is computed in Ref. [24], while the functions $I_{K^{(*)}}(\hat{s}, m_s)$ and $h(\hat{s})$, which express the allowed phase space, have the following form:

$$
I_{K}(\hat{s}, m_s) = \left[\hat{s}^2 - 2\hat{s}(m_B^2 + m_K^2) + (m_B^2 - m_K^2)^2\right]^{1/2}[1 - 4m_s^2/\hat{s}]^{1/2},
$$

$$
h(\hat{s}) = \left(1 + \frac{m_{K^*}^2}{m_B^2} - \frac{\hat{s}}{m_B^2}\right)^2 - \frac{4m_{K^*}^2}{m_B^2},
$$

$$
I_{K^{(*)}}(\hat{s}, m_s) = \left[\hat{s}^2 + (m_B^2 - m_{K^*}^2)^2 - 2\hat{s}(m_B^2 + m_{K^*}^2)\right]^{1/2}[1 - 4m_s^2/\hat{s}]^{1/2}.
$$

(1.44)

For light WIMP candidates ($m_S \sim$ few MeV/c$^2$) and proper values of the coupling between the Higgs bosons and the scalar fields, the $B \to K^{(*)} + \text{invisible}$ branching fractions can be enlarged up to a factor 50 with respect to the SM final states with neutrinos.

Other NP sources that can experimentally be detected as missing energy are the Unparticles [29]. In this scenario, an effective field theory with a non trivial scale-invariant sector is postulated. At very high energy scales, the theory comprises both the SM fields and the massless Banks-Zacks (BZ) fields which are related to the scale invariant sector. The interaction between the two is mediated by massive particles ($\mathcal{U}$), that define an
energy scale \( (M_{U}) \). Below \( M_{U} \), non renormalizable couplings between \( U \) and both the SM and the Banks-Zacks fields are present and take the form:

\[
\frac{1}{M_{U}^{d}} O_{SM} O_{BZ}, \tag{1.45}
\]

where \( O_{SM} (O_{BZ}) \) is the operator with mass dimension \( d_{SM} (d_{BZ}) \) related to the SM (BZ) fields. By renormalizing the couplings, scale-invariance in the BZ sector emerges at an energy scale \( A_{U} \). In the effective theory, at low energies, the BZ fields match onto unparticle operators and Eq. 1.45 becomes:

\[
\frac{C_{U} A_{U}^{d_{BZ}-d_{U}}}{M_{U}^{d_{U}} O_{SM} O_{U}}, \tag{1.46}
\]

d\(_{U}\) being the scaling dimension of the unparticle operator \( O_{U} \).

Given the massless nature of the unparticles, they might be detectable in missing energy and momentum distributions: as a consequence they can contribute to final states with invisible particles. The way in which the unparticle operators enter the \( b \to s \) \( U \) effective Hamiltonian depends on the way the \( U \)-fields are assumed to be.

Scalar unparticle operators contributing to the \( B \to K^{(*)} U \) process lead to the following expressions for the decay rates, as a function of the kaon energy [30]:

\[
\frac{d\Gamma^{U}(B \to K U)}{dE_{K}} = \frac{1}{8\pi^{2}m_{B} A_{U}^{2d_{U}}} |C_{S}|^{2} \sqrt{E_{K}^{2} - m_{K}^{2}} \left( m_{B}^{2} + m_{K}^{2} - 2m_{B}E_{K}\right)^{d_{U}-2} \times \left[ f_{+}(m_{B}^{2} - m_{K}^{2}) + f_{-}(m_{B}^{2} + 2m_{K}^{2} - 2m_{B}E_{K}) \right]^{2},
\]

\[
\frac{d\Gamma^{U}(B \to K^{*} U)}{dE_{K^{*}}} = \frac{m_{B} A_{U}^{4d_{U}} |C_{P}|^{2} A_{0}^{2} \left(E_{K^{*}}^{2} - m_{K^{*}}^{2}\right)^{2} \left( m_{B}^{2} + m_{K^{*}}^{2} - 2m_{B}E_{K^{*}}\right)^{d_{U}-2}}{2\pi^{2} A_{U}^{2d_{U}}} \tag{1.47}
\]

where:

\[
A_{d_{U}} = \frac{16\pi^{5/2}}{(2\pi)^{2d_{U}} \Gamma(d_{U} + 1/2)} \frac{\Gamma(d_{U} + 1/2)}{\Gamma(d_{U} - 1)\Gamma(2d_{U})}. \tag{1.48}
\]

Also the existence of vector unparticle fields is postulated. In this case, the differential
1.4 $B \rightarrow K^{(*)}\nu\bar{\nu}$ and New Physics discoveries

decay rates are:

$$\frac{d\Gamma_{U}(B \rightarrow KU)}{dE_{K}} = \frac{1}{8\pi^{2}m_{B}} \frac{A_{dU}}{\Lambda^{2}q^{2}} \left|C_{V}\right|^{2} |f_{+}|^{2} \left(m_{B}^{2} + m_{K}^{2} - 2m_{B}E_{K}\right)^{qU-2} \sqrt{E_{K}^{2} - m_{K}^{2}} \times \left\{ - \left( m_{B}^{2} + m_{K}^{2} + 2m_{B}E_{K} \right) + \frac{(m_{B}^{2} - m_{K}^{2})^{2}}{(m_{B}^{2} + m_{K}^{2} - 2m_{B}E_{K})} \right\} ,$$

$$\frac{d\Gamma_{U}(B \rightarrow K^{*}U)}{dE_{K^{*}}} = \frac{1}{8\pi^{2}m_{B}} (q^{2})^{qU-2} \sqrt{E_{K^{*}}^{2} - m_{K^{*}}^{2}} \frac{A_{dU}}{\left(\Lambda^{qU-1}\right)^{2}} \left\{ 8|C_{V}|^{2} m_{B}^{2} \left( E_{K^{*}}^{2} - m_{K^{*}}^{2} \right) \left( m_{B} + m_{K^{*}} \right)^{2} + |C_{A}|^{2} \right\} \frac{1}{m_{K^{*}}^{2} \left( m_{B} + m_{K^{*}} \right)^{2} q^{2}}$$

$$\left[ (m_{B} + m_{K^{*}})^{4} (3m_{K^{*}}^{4} + 2m^{2}m_{K^{*}}^{2} - 6m_{B}m_{K^{*}}^{2}E_{K^{*}} + m_{B}^{2}E_{K^{*}}^{2})A_{1}^{2} + 4m_{B}^{4} \left( E_{K^{*}}^{2} - m_{K^{*}}^{2} \right) A_{2}^{2} \right] + 4(m_{B} + m_{K^{*}})^{2} (m_{B}E_{K^{*}} - m_{K^{*}})(m_{K^{*}}^{2} - E_{K^{*}}^{2})m_{B}^{2}A_{1}A_{2} \right\} . \quad (1.49)$$

Model independent measurements of the $B \rightarrow K^{(*)}\nu\bar{\nu}$ branching fractions allow constraining of the couplings $C_{S}-C_{P}$ and $C_{V}-C_{A}$ for scalar and vector unparticle fields, respectively, or the scaling dimension of the unparticle operator $d_{U}$. An example is given in Figs. 1.6 and 1.7, where the branching fraction of $B \rightarrow K^{(*)}\nu\bar{\nu}$ as a function of $d_{U}$, for different values of $A_{dU}$ is shown [30].

![Figure 1.6](image1.png)

Figure 1.6: $B \rightarrow K + invisible$ branching fraction as a function of $d_{U}$ for various values of $A_{dU}$: scalar operator contribution on the right and vector operator contribution on the left. The following values for the couplings are chosen: $C_{S} = 2 \times 10^{-3}$ and $C_{V} = 10^{-5}$. The black line shows the experimental measurement [31], [32].
Theoretical and experimental knowledge of $b \rightarrow s \nu \bar{\nu}$ transitions

![Graph](image)

Figure 1.7: $B \rightarrow K^* + \text{invisible}$ branching fraction as a function of $d_U$ for various values of $\Lambda_U$: scalar operator contribution on the right and vector operator contribution on the left. The following values for the couplings are chosen: $C_P = 2 \times 10^{-3}$ and $C_V = C_A = 10^{-5}$. The black line shows the experimental measurement [31].

1.5 Current experimental status on $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays

Search for $B \rightarrow K^{(*)}\nu\bar{\nu}$ have been performed from two $B$ Factory experiment: Belle in Japan and BABAR.

The Belle experiment has published a paper in which the searches for both $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^{*}\nu\bar{\nu}$ are presented, using 535 million $B\bar{B}$ pairs [31]. The bottom line of the work is the following: a hadronic $B$ decay is reconstructed on one side of the $B\bar{B}$ event, on the other side a $K$ or a $K^*$ associated to missing energy deposited in the electromagnetic calorimeter and not associated to the two $B$’s is searched for. Simulated data samples allow to estimate the expected number of background events. Comparing such a yield to the data events surviving the selection, an estimate of the signal event sample is extracted. A frequentist approach is used to extract the following upper limits at 90% confidence level $^1$:

\[
\begin{align*}
\mathcal{B}(B \rightarrow K^+\nu\bar{\nu}) &< 1.4 \times 10^{-5} \\
\mathcal{B}(B \rightarrow K^0\nu\bar{\nu}) &< 1.6 \times 10^{-4} \\
\mathcal{B}(B \rightarrow K^{*+}\nu\bar{\nu}) &< 1.4 \times 10^{-4} \\
\mathcal{B}(B \rightarrow K^{*0}\nu\bar{\nu}) &< 3.4 \times 10^{-4} .
\end{align*}
\]  

$^1$Charge conjugation is implied throughout this thesis, unless explicitly stated.
1.5 Current experimental status on $B \to K^{(*)}\nu\bar{\nu}$ decays

The BABAR experiment has searched for the charged $B \to K \nu\bar{\nu}$ channel [33]. One $B$ in the event is reconstructed in a semileptonic decay, associated to a charged kaon plus missing energy on the other side. Also in this case the expected background, estimated from simulated events, is compared to the final data sample and a frequentist approach is exploited to set the following upper limit at 90\% confidence level:

$$B(B \to K^+\nu\bar{\nu}) < 4.2 \times 10^{-5}. \quad (1.51)$$

The analyzed sample includes 351 million $B\bar{B}$ events. No search for the $B \to K^0_S \nu\bar{\nu}$ has been published by BABAR. The work presented in this thesis is the first measurement of the $B \to K^{*}\nu\bar{\nu}$ branching fraction of the experiment.

The previous upper limits are consistent with the SM expectation and no evidence for signal has been measured.
Theoretical and experimental knowledge of $b \rightarrow s \nu \overline{\nu}$ transitions
Chapter 2

The $\textit{Babar}$ detector

2.1 Introduction

Along with the Belle experiment running in Japan, $\textit{Babar}$ started taking data in 1999 with two main goals: measure $CP$ violation in $B$ decays and probe the correctness of the SM prediction. Up to April 2008, $\textit{Babar}$ has collected data at the PEP-II $B$ Factory, that is an asymmetric high-luminosity $e^+e^-$ collider operating at the center-of-mass (CM) energy of the $\Upsilon(4S)$ mass ($10.58 \text{ GeV}/c^2$). This particle, a $b\bar{b}$ bound state, decays in $B\bar{B}$ pairs both charged and neutral, with a decay rate approximately equals to one. The large and clean $B\bar{B}$ sample allowed significant advances in many areas of flavor physics by studying the $B$, $D$, and $\tau$ systems. In fact, other then testing $CP$ violation in $B$ meson system, several measurements related to rare processes, $CP$ violation in $D$ decays and $\tau$ physics have been produced, confirming the validity of the SM in describing the flavor physics. Moreover, search for discrepancies between SM prediction and experimental results have helped in constraining NP parameters.

$\textit{Babar}$ [34] is a multipurpose detector optimized for $CP$ violation studies. It was designed to meet stringent requirements imposed by the physics program. The detector needs to satisfy the following conditions:

- have the maximum possible acceptance in the CM frame: as a consequence the detector must be asymmetric to cover the forward region in which the decays products are boosted due to the different beam energies;

- accommodate the machine components close to the interaction region to record a luminosity as high as possible;
• have an excellent vertex resolution, essential for $CP$ measurements, especially along the $z$-axis almost parallel to the $B$ mesons direction;

• keep at the minimum the material in the active region to reduce multiple Coulomb scattering and particle-matter interaction effects that degrade the measurements of the track parameters and energy;

• have a good tracking efficiency in the momentum range $[60 \text{ MeV}/c, 4 \text{ GeV}/c]$;

• discriminate between $e$, $\mu$, $\pi$, $K$, and proton over a wide kinematic range;

• have excellent angular and energy resolutions to detect $\gamma$ and $\pi^0$ with energy in the range $[20 \text{ MeV}, 5 \text{ GeV}]$;

• identify neutral hadrons.

A view of the longitudinal section of the $BABAR$ detector is shown in Fig. 2.1 and the major subsystems are indicated. The detector has a cylindrical geometry in the inner zone and a hexagonal shape in the outermost; the central body is called “barrel” and it is closed by a backward and a forward “endcap”. The covered polar angle is 300 mrad in the forward and 400 mrad in the opposite direction (these are defined with respect to the high energy beam). The $BABAR$ coordinate system has the $z$-axis along the boost direction, the $y$-axis is vertical and the $x$-axis is horizontal and goes toward the external part of the PEP-II ring.

The inner device is a Silicon Vertex Tracker (SVT) that provides precise position determination on charged particles and tracking information for low-energy charged tracks. The main role in the momentum determination is played by a Drift Chamber (DCH) filled with a helium-based gas, that minimize multiple scattering effects; it participate in the Particle Identification (PID) through energy loss measurements. For this purpose a Detector of Internally Reflected Čerenkov light (DIRC) is designed. In the barrel and in the forward endcap a cesium iodide Electromagnetic Calorimeter (EMC) is placed to supply electron and neutral particle identification. All the subsystems listed above are surrounded by a 1.5 $T$ magnetic field supplied by a superconductive solenoid. The yoke for the flux return of the magnetic field is highly segmented and equipped with Resistive Plate Chambers (RPCs) and Limited Streamers Tubes (LSTs) for muon and long-living neutral hadron detection. A trigger system is used to distinguish collisions producing
events of interest from those which constitutes the background. In the next sections a
detailed description of the PEP-II $B$ Factory and of the $B$ABAR subsystems is given.

![BABAR Detector Longitudinal Section](image)

Figure 2.1: $B$ABAR detector longitudinal section.

## 2.2 The PEP-II $B$-factory

The PEP-II $B$ Factory (Fig. 2.2) is located at the Stanford Linear Accelerator Center
(SLAC) laboratories in California. It is an $e^+e^-$ asymmetric collider running at a center-
of-mass energy of 10.58 GeV corresponding to the $\Upsilon(4S)$ resonance. The energy of the
electron beam in the High Energy Ring (HER) is 9.0 GeV while the energy of the positron
beam in the Low Energy Ring (LER) is 3.1 GeV, resulting in a Lorentz boost of $\beta \gamma = 0.56$
of the $\Upsilon(4S)$. The two $B$ mesons in which the $\Upsilon(4S)$ decays are almost at rest in the CM
frame. The $\Upsilon(4S)$ Lorentz boost and the knowledge of the $B$ meson momentum in the
CM makes it possible to reconstruct the decay vertices of the two $B$’s and to determine
their relative decay time $\Delta t$, necessary to study CP asymmetry as a function of time. If
the laboratory frame matched the CM frame, the spatial difference between the two decay
points would be of the order of 30 $\mu$m and thus very challenging to measure; due to the
Lorentz boost, the spatial separation is about 250 $\mu$m.
The $B$ decays relevant for $CP$ violation studies have branching fractions typically smaller than $10^{-4}$, as a consequence to perform precise measurements a high luminosity machine is needed. PEP-II was designed to reach an instant luminosity of $3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ while the average operational luminosity has exceeded four times this value, reaching the record value of $12.07 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$. Design and typical values of some parameters of the machine are listed in Tab. 2.1.

Table 2.1: PEP-II parameters: design and typical values.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Design</th>
<th>Typical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy HER/LER (GeV)</td>
<td>9.0/3.1</td>
<td>9.0/3.1</td>
</tr>
<tr>
<td>Current HER/LER (A)</td>
<td>0.75/2.15</td>
<td>1.48/2.50</td>
</tr>
<tr>
<td># of bunches</td>
<td>1,658</td>
<td>553-829</td>
</tr>
<tr>
<td>bunch time separation (ns)</td>
<td>4.2</td>
<td>6.3-10.5</td>
</tr>
<tr>
<td>$\sigma_{Lx}$ (\mu m)</td>
<td>110</td>
<td>120</td>
</tr>
<tr>
<td>$\sigma_{Ly}$ (\mu m)</td>
<td>3.3</td>
<td>5.6</td>
</tr>
<tr>
<td>$\sigma_{Lz}$ (\mu m)</td>
<td>9,000</td>
<td>9,000</td>
</tr>
<tr>
<td>Luminosity ($10^{33}$ cm$^{-2}$s$^{-1}$)</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Daily average integrated luminosity (pb$^{-1}$/d)</td>
<td>135</td>
<td>700</td>
</tr>
</tbody>
</table>

Most of the data were collected at the $\Upsilon(4S)$ peak energy (on-peak data): Tab. 2.2 shows the active process cross-section breakdown at peak energy. Events are also collected 40 MeV below the $\Upsilon(4S)$ mass in order to study light ($u, d, s$) and charm-quark pair production (“continuum production”) and they are referred to as off-peak data. At the end of the data taking, the experiment has also collected data at the $\Upsilon(3S)$ and $\Upsilon(2S)$
resonances. The total integrated luminosity from the beginning to the end of the data taking (October 1999 - April 2008) is shown in Fig. 2.3.

Figure 2.3: Total integrated luminosity as delivered by PEP-II (blue) and collected by BABAR (red) during the whole data-taking period (October 1999 - April 2008). Recorded luminosities at the $\Upsilon(4S)$, $\Upsilon(3S)$, $\Upsilon(2S)$ and off-peak data are shown in magenta, purple, yellow, and green, respectively.

Table 2.2: Cross sections of active process at $\sqrt{s}=10.58$ GeV. Bhabha cross section is an effective cross section, within experimental acceptance.

<table>
<thead>
<tr>
<th>$e^+e^- \rightarrow$</th>
<th>Cross section (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b\bar{b}$</td>
<td>1.05</td>
</tr>
<tr>
<td>$c\bar{c}$</td>
<td>1.30</td>
</tr>
<tr>
<td>$s\bar{s}$</td>
<td>0.35</td>
</tr>
<tr>
<td>$u\bar{u}$</td>
<td>1.39</td>
</tr>
<tr>
<td>$d\bar{d}$</td>
<td>0.35</td>
</tr>
<tr>
<td>$\tau^+\tau^-$</td>
<td>0.94</td>
</tr>
<tr>
<td>$\mu^+\mu^-$</td>
<td>1.16</td>
</tr>
<tr>
<td>$e^+e^-$</td>
<td>$\simeq$ 40</td>
</tr>
</tbody>
</table>

PEP-II measures radiative Bhabha scattering to provide a luminosity fast monitoring. BABAR derives the absolute luminosity off-line by studying other QED processes, mainly $e^+e^- \rightarrow \mu^+\mu^-$: the uncertainty on the integrated luminosity determination is about 1.5%
and the main contribution arises from uncertainties in the Monte Carlo (MC) generator and the simulation of the detector.

The high luminosity environment produces large backgrounds, such as synchrotron radiation, electromagnetic showers generated by beam-beam collisions, and interaction between beam particles and residual gas in the ring. This affects the detector performances and lifetime and monitoring and protection systems have been developed to avoid severe damages.

2.3 The tracking system

The tracking of charged particles is committed to the SVT and the DCH. They should give an efficient detection of charged tracks and measurements of their momentum and direction. This information is important for the extrapolation of the track to the DIRC, the EMC, and the IFR. At transverse momentum ($p_T$) lower than 120 MeV/$c$, the SVT information dominate, while at higher momentum the DCH measurement becomes more important.

2.3.1 The Silicon Vertex Tracker: SVT

The separation between the decay vertices of the two $B$'s is crucial for many $BABAR$ analyses dedicated to $CP$ measurements. The vertex determination is the main task of the SVT. The proximity of the silicon tracker to the vertex production allows for a very precise measurement of points on the charged track trajectory in both longitudinal and transverse directions: the first information is important for measurements of the decay vertex distance while the second permits secondary vertex separation. Spatial resolutions of 80 $\mu$m in the $z$ direction and of 100 $\mu$m in the $x$-$y$ plane are required. The SVT should provide momentum measurement for particles with low $p_T$ that do not reach the DCH, with a tracking efficiency of at least 70% in the momentum range [50,200] MeV/$c$. It participates also in particle identification by giving the most precise measurement of track angles which is required to achieve the design resolution for the Čerenkov angle of high momentum tracks. The tracker should also have the maximum possible angular coverage and be very resistant to radiation.

The SVT is placed inside the support tube of the beam magnets, at a radius of 20 cm from the primary interaction region. It consists of five concentric cylindrical layers of
2.3 The tracking system

double-sided silicon strips (Figs. 2.4-2.5) that allow five position measurements of charged particles with polar angle in the region \(20.1^\circ < \theta < 150^\circ\). The strips are arranged in 340 wafers mounted on a carbon fiber frame. On the inner (outer) face of each wafer, strips and sensors run orthogonal (parallel) to the beam direction and measure the \(z\) (\(\phi\)) coordinate of the track. The wafers are organized in modules split into forward and backward sections. The read-out channels (150,000 in total) are located at the end of each wafer and the charge deposited by the charged particle is determined by a time-over-threshold measurement of the signal in each strip. The three innermost layers, containing six modules each, are placed close to the beam pipe and dominate the track position and angle measurements; the two outermost, composed of 16 and 18 modules respectively, are arch-shaped to minimize the silicon needed to cover the solid angle and to be placed as close as possible to the DCH, aiding the track matching between the two devices. The modules of the inner layers are tilted in \(\phi\) by \(5^\circ\) to provide a full azimuthal coverage; the modules of the outer layers can not be tilted due to the arch shape but are divided in two sublayers to have a suitable overlap and avoid gaps. The total active silicon area is 0.96 m\(^2\) and the geometrical acceptance is 90% of the solid angle in the CM frame.

![Figure 2.4: View of the transverse section of the SVT.](image)

The SVT has a software and hardware combined single-hit efficiency of 97% and this is computed, for each half-module, by comparing the number of associate hits to the number of tracks crossing the active area of the component. In Fig. 2.6 the hit reconstruction efficiency for \(\mu^+\mu^-\) events is shown. The spatial resolution of the SVT hits is determined by measuring the distance between the track trajectory and the hit for high-momentum tracks in two-tracks events; it is nearly 40 \(\mu\)m in all layers for all tracks angles, giving
a vertex resolution better than 70 $\mu$m. Also measurements of ionization loss $dE/dx$ are performed with up to ten measurements for every track, thanks to the time-over-threshold value reported by the SVT. For minimum ionizing particles the resolution in $dE/dx$ is approximately 14%, leading to a $2\sigma$ separation between kaons and pions up to a momentum of 500 MeV/$c$ and between kaons and protons beyond 1 GeV/$c$.

---

Figure 2.5: View of the longitudinal section of the SVT.

Figure 2.6: SVT hit reconstruction efficiency as measured on $e^+e^- \rightarrow \mu^+\mu^-$ events for a) forward half-modules and b) backward half-modules. The plot shows the probability of associating both a $\phi$ and a $z$ hit to a track passing through the active part of the detector. The vertical lines separate the different layers.
2.3 The tracking system

2.3.2 The Drift Chamber: DCH

The second part of the BABAR tracking system is the DCH. Its main purpose is the high-precision measurement of angles and momenta of charged particles, that is ensured by up to 40 measurements of space coordinate per track. It allows to determine impact parameters and directions of tracks, adding information to the SVT measurements near the impact point; it is also crucial in the extrapolation of the trajectories to outer subsystems. The DCH is the only tracking system that provides vertex information for particles decaying outside the SVT. It also participates in particle identification by performing dE/dx measurements, and it is complementary to the DIRC in the barrel region, while is the only device that can discriminate particles of different masses in the extreme backward and forward regions. The DCH should satisfy the following requirements: provide maximum solid angle coverage, perform good measurement of transverse momenta and transverse and longitudinal positions with a resolution of about 1 mm, be efficient in the reconstruction of tracks with momentum as low as 100 MeV/c, have small resolution in π/K separation up to few hundreds of MeV/c, degrade as less as possible the performance of the most external devices, and finally be able to operate in presence of high beam-generated background, whose expected rate is of about 5 kHz/cell in the innermost layers.

A schematic view of the BABAR DCH is shown in Fig. 2.7. It is a 280 cm-long cylinder with inner (outer) radius of 23.6 cm (80.9 cm). Its center is displaced by about 37 cm with respect to the interaction point (IP) in the forward direction. The polar angle covered by the active volume is $-0.92 < \cos \theta < 0.96$. The inner cylinder is made of 1 mm beryllium, and the outer 2 layers consist of carbon fiber glued on a Nomex core. The inner surface is kept as thin as possible to facilitate the matching of SVT and DCH tracks, to minimize the background from photon conversion and to maximize the track resolution for high momentum tracks. Also the material in the outer wall and in the forward direction is minimized to avoid degradation of DIRC and EMC performances. The region between the two cylinders is filled up with a Helium-Isobutane gas mixture in the proportion 80% : 20%. The device is bounded by the support tube at its inner radius and by the DIRC at its outer. The flat endplates are made of aluminum, the one in the forward region being thinner (12 mm) compared to the rear end-plate (24 mm), and all the electronics is placed on the rear end-plate.
The drift system consists of 7,104 hexagonal cells which are approximately 1.8 cm-wide and 1.2 cm-high, made up of one sense wire surrounded by six field wires. The sense wires are 20 \( \mu \)m gold-plated tungsten-rhenium, the field wires are 120 \( \mu \)m and 80 \( \mu \)m gold-plated aluminum. The low mass of the aluminum wires and the low density of the gas mixture reduce to minimum the multiple scattering inside the DCH, that represents less than 0.2% \( X_0 \) of the material. The total thickness of the DCH at normal incidence is 1.8% \( X_0 \). The cells are arranged in 10 super-layers of 4 cylindrical layers each. The wires of one superlayer have the same orientation; to make the \( z \) coordinate measurement possible, axial wire super-layers are alternated to super-layers with slightly rotated wires (stereo). The layout of the drift cells in the four innermost super-layers is shown in Fig. 2.8. The stereo angles vary between ±45 mrad and ±76 mrad. In the stereo super-layers a single wire correspond to different \( \phi \) angles and the \( z \) coordinate is determined by comparing the \( \phi \) measurement from axial wires and the measurements from rotated wires. The sense wires are at a positive high voltage, while the field wires are grounded: this configuration ensures an avalanche gain of approximately \( 5 \times 10^4 \) at the typical operational voltage of 1,960 V. Samples of \( e^+e^- \) and \( \mu^+\mu^- \) events are used to determine the precise relation between the measured drift time and the drift distance. For each signal, the drift distance is estimated by computing the distance of closest approach between the track and the wire.

The absolute efficiency for reconstructing tracks in the DCH is evaluated as the ratio of the number of reconstructed tracks in the DCH to the number of tracks detected in
2.3 The tracking system

The tracking system is composed of the BaBar Drift Cells (DCH), which are used to detect and measure the momentum of charged particles. The DCH consists of multiple superlayers, each containing drift cells arranged in a specific pattern. Figure 2.8 shows the schematic layout of the drift cells for the four innermost DCH superlayers.

At the design voltage of 1,960 V, the average efficiency for tracks with $p_T > 200$ MeV/c and $\theta > 500$ mrad is 98%. In Fig. 2.9 the track reconstruction efficiency is shown. The DCH resolution in the position determination is lower than 100 $\mu$m in the transverse plane and about 1 mm in the z direction. The DCH inner radius defines a minimum reconstruction and momentum measurement threshold of about 100 MeV/c. The achieved weighted-average resolution on the single hit is about 125 $\mu$m, to be compared with the 140 $\mu$m design value. The DCH spatial resolution as a function of the drift distance in layer 18 is shown in Fig. 2.10. The accuracy of the momentum measurements (combining SVT and DCH information) as a function of $p_T$ is well approximated by the following function:

$$\frac{\sigma_{p_T}}{p_T} = (0.13 \pm 0.01)\% \cdot p_T + (0.45 \pm 0.03)\%$$

where $p_T$ is given in GeV/c. Measurements of $dE/dx$ in the DCH are derived from the measurements of the total charge collected in each drift cell and the specific energy loss per track is computed as a truncated mean from the lowest 80% of the individual $dE/dx$ measurements. The resolution for Bhabha events is typically 7.5%. Concerning PID, a resolution of about 7% allows $\pi/K$ separation up to 700 MeV/c.
Figure 2.9: Track reconstruction efficiency in the DCH at operating voltages of 1,900 V and 1,960 V as a function of a) transverse momentum and b) polar angle.

Figure 2.10: DCH position resolution as a function of the drift distance in layer 18, for tracks on the left and right side of the sense wire. The data are averaged over all cells in the layer.
2.4 The Čerenkov Light Detector: DIRC

The PID is a crucial ingredient for BABAR physics since many analyses require the ability of fully or partially reconstructing one $B$ meson and determining the flavor of the other $B$ by means of a kaon or a lepton. The DIRC, that is the main PID system in BABAR, is a novel type of ring-imaging Čerenkov device. Ionization measurements by the DCH and the SVT allow for PID up to 700 MeV/$c$, so the DIRC is fundamental to identify particles with higher momentum. To minimize the performance degradation of the other devices, it should be thin and uniform in term of radiation length and small in volume. It should also provide fast signal response and be able to tolerate large background levels. Due to the boost, particles are produced mainly forward in the detector, therefore the DIRC photon detector is placed at the backward end. Its main components are shown in Fig. 2.11.

![Figure 2.11: The principal components of the DIRC detector.](image)

In total the DIRC only occupies 80 mm of radial space, including supports, with 17% radiation lengths of material. The subdetector consists of 144 fused silica quartz bars with rectangular cross section arranged in a 12-sided polygonal barrel; the bars are 1.7 cm-thick, 3.5 cm-wide, and 4.9 m-long. When a charged particle crosses a quartz bar, it generates Čerenkov photons at an angle $\theta_c$ with respect to its direction such that $\cos \theta_c = 1/(\beta n)$, where $\beta$ is the velocity of the particle and $n$ is the refractive index of quartz ($n = 1.473$). The Čerenkov light is produced within the bars that have also the task of internally reflecting the radiation and transporting it to one of the ends. The photons going forward are reflected by a mirror, so that all arrive at the instrumented rear end of the bar; they then exit the bar through a quartz window and emerge into a
The photon detection is committed to an array of 10,752 photomultiplier tubes (PMTs), located outside the return yoke of BABAR at about 1.2 m from the bar ends. The PMTs are arranged on the surface of a half torus with an inner and an outer radius of 1.2 m and 3 m, respectively. At the PMT surface the Čerenkov light pattern appears as a conic section with opening angle corresponding to the Čerenkov angle, modified by the refraction at the exit of the quartz window: given the known location of the PMT which observes a Čerenkov photon and the charged particle direction from the tracking system, the Čerenkov angle is determined. In addition, the arrival time of the signal provides an independent measurement of the propagation of the photon. This over-constrain on the angles and the signal timing are useful in dealing with ambiguities in the signal association and large background rates.

The expected number of photo-electrons \(N_{pe}\) is \(\sim 28\) for a \(\beta = 1\) particle entering normal to the surface at the center of a bar and increases by over a factor two in the forward and backward directions.

The time distribution of real Čerenkov photons from a single event is of the order of 50 ns-wide and during normal data taking they are accompanied by hundreds of random photons in a flat background distribution within the trigger acceptance window. The

---

\(^1\)Water has a refractive index similar to the quartz, so that internal reflection at the interface between the bar and the expansion region is reduced.
Čerenkov angle has to be determined in an ambiguity that can be up to 16-fold: the goal of the reconstruction program is to associate the correct track with the candidate PMT signal requiring the transit time of the photon from its creation in the bar to its detection at the PMT to be consistent with the measurement error of about 1.5 ns.

Dimuon events are used to determine the resolution on the Čerenkov angle and time measurements. For a single photon, the angular resolution is $10.2$ mrad and the time resolution is $1.7$ ns. The angle resolution per track is $2.5$ mrad and it corresponds to a separation of four standard deviations between kaons and pions at $3$ GeV/$c$. The efficiency for correctly identifying a charged kaon hitting a radiator bar and the probability of wrongly identifying a pion as a kaon (Fig. 2.13) are determined using $D^0$ decays kinematically selected from inclusive $D^*$ meson production: the mean efficiencies for kaon identification and pion misidentification are about 96% and 2%, respectively.

![Figure 2.13: Efficiency and misidentification probability for the selection of charged kaons as a function of the track momentum.](image)

2.5 The Electromagnetic Calorimeter: EMC

The *BABAR* electromagnetic calorimeter has been designed to measure electromagnetic showers with excellent energy and angular resolution over the range of energies $[20$ MeV, $9$ GeV]. The lower bound is set by the need for detecting $\pi^0 \rightarrow \gamma \gamma$ and $\eta \rightarrow \gamma \gamma$, whose photons in the final state may have very low energy; the upper bound allows to reconstruct Bhabha events and processes like $e^+ e^- \rightarrow \gamma \gamma$, used for luminosity monitoring and calibration. The EMC also participates in the electron detection that contributes to the determination...
of the $B$ meson flavor via semi-leptonic decays, and the reconstruction of $\tau$ decays and rare $D$ and $B$ processes. The most stringent requirements on energy resolution, namely of order of 1-2%, are posed by the need for measuring rare $B$ decays containing $\pi^0$ (e.g., $B^0 \rightarrow \pi^0 \pi^0$). Below 2 GeV, the $\pi^0$ mass resolution is dominated by energy resolution, while at higher energies, the angular resolution becomes dominant and it is required to be of the order of few mrad. The EMC information is also completing the IFR output on $\mu$ and $K_L^0$ identification.

To meet the requirements stated above, the calorimeter has been chosen to be hermetic and total-absorption. It is composed of a finely segmented array of thallium-doped caesium iodide (CsI(Tl)) crystals (Fig 2.14). CsI(Tl) has been adopted for its high light yield (50,000 photons/MeV) and the small Molière radius (3.8 cm), which provides the required energy and angular resolution; its radiation length of 1.86 cm guarantees complete shower containment at BABAR energies. The EMC consists of a cylindrical barrel and a conical forward, it has a full coverage in azimuth and extends in the polar angle from 15.8° to 141.8°, corresponding to a solid angle coverage of 90% in the CM frame. Overall the EMC extends from an inner radius of 91 cm to an outer radius of 136 cm. In the barrel, the crystals are arranged in 48 rings with 120 identical crystals each; the endcap holds 8 rings. Each crystals has a tapered trapezoidal cross section and ranges from 16 to 17.5 radiation length in thickness; the front faces are typically about 5 cm in each dimension. The readout system of each crystal is composed of two independent 1-cm$^2$ PIN photodiodes, glued to their rear faces and connected to low-noise preamplifiers which shape the signal with a short shaping time (400 ns) so to reduce soft beam-related photon backgrounds.

Figure 2.14: Longitudinal section of the top half of the EMC. Dimensions are in mm.

A typical electromagnetic shower spreads over many adjacent crystals, forming a **clus-**
ter of energy deposit. The clusters are identified by means of pattern recognition algorithms, which also allow to discriminate single clusters with one energy maximum from merged clusters with more than one local energy maximum, called *bumps*, and determine whether a bump is generated by a charged or a neutral particle. A cluster must contain at least one seed crystal with an energy above 10 MeV, surrounding crystals whose energy exceeds a threshold of 1 MeV, or crystals which are contiguous neighbors of a crystal with at least 3 MeV signal, are considered as part of the cluster. A bump is associated to a charged particle by projecting a track to the inner face of the calorimeter: the distance between the track impact point and the bump centroid is calculated and if it is consistent with the angle and momentum of the track, the bump is associated with this charged particle. Otherwise it is assumed to originate from a neutral particle. In one hadronic event 15.8 clusters are detected on average, and 10.2 are not associated to any charged particle. Currently, the beam-induced background contributes with about 1.4 neutral clusters with energy above 20 MeV.

The EMC calibration and monitoring is performed by using a neutron activated fluorocarbon fluid and a light pulser system: the first, activated by low-energy neutrons, produces a radiative isotope ($^{16}$N) originating a 6.1 MeV photon peak in each crystal, the second injects light into the rear of each crystal. For energy calibration and resolution determination also signal from data are used: they include $\pi^0$ decays, and $e^+e^- \rightarrow e^+e^-$, $e^+e^- \rightarrow \gamma\gamma$, $e^+e^- \rightarrow \mu^+\mu^-$ events.

The *BABAR* EMC is highly performing: it exceeds 96% efficiency for photon detection at energies above 20 MeV. The energy resolution is parametrized as:

$$\frac{\sigma_E}{E} = \frac{\sigma_1}{E^{1/4}(\text{GeV})} \oplus \sigma_2,$$

where $\sigma_1 = (2.32 \pm 0.30)\%$ and $\sigma_2 = (1.85 \pm 0.12)\%$, as determined from calibration studies. The first term in Eq. 2.2 is due to fluctuations in photon statistics, while the constant term includes several effects such as fluctuations in shower containment, non-uniformities, calibration uncertainties and electronic noise. Also the angular resolution is parametrized as a function of energy, in the following form:

$$\frac{\sigma_{\theta,\phi}}{E} = \frac{\sigma_3}{\sqrt{E(\text{GeV})}} \oplus \sigma_4,$$

where $\sigma_3 = (3.87 \pm 0.07)$ mrad and $\sigma_4 = (0.00 \pm 0.04)$ mrad. In Fig. 2.15 the energy and angular resolution of the EMC as a function of energy are shown.
The calorimeter is also used for electron-hadron separation in conjunction with the other subdetectors. The most discriminating variable to distinguish between the two is the ratio of the shower energy measured by the EMC to the track momentum from the tracking system. The electron identification efficiency is determined using radiative Bhabha and $e^+e^- \rightarrow e^+e^- e^+e^-$ events, while the pion misidentification probability is determined using pions from $K_S^0$ and $\tau$ lepton decays (Fig 2.16).

2.6 The Solenoid Magnet

The subsystems described above are enclosed in the BABAR magnet system. This is composed by a superconducting solenoid, a highly segmented steel flux return (Sec. 2.7) and a field-compensating, or “bucking”, coil. The system provides a 1.5 $T$ magnetic field which enables the measurement of charged particle momenta. The chosen field magnitude allows to achieve the desired momentum resolution and it is constant within the tracking system volume in order to simplify the tracking. The magnetic field mainly lies along the $z$-axis, approximatively corresponding to the direction of the electron beam, the azimuthal component is smaller than 1 mT everywhere in the tracking volume. The superconducting solenoid is made of Nb-Ti filaments, less than 40 $\mu$m in diameter. The thermal operating point of the magnet is 4.5 K and it is maintained by a thermo-syphon technique.

The bucking coil function is to reduce the leakage field into the PEP-II components and the DIRC photomultipliers: it consists of a 10-layer water-cooled copper coil.
Figure 2.16: Electron efficiency and pion misidentification probability as a function of the momentum (top) and polar angle (bottom).
Other than providing the magnetic field, the magnet system support the detector components and serves as a hadron absorber to help hadron/muon separation.

2.7 The Instrumented Flux Return: IFR

The steel flux return for the solenoidal magnetic field is finely segmented and instrumented in order to detect muons and long-lived neutral hadrons over a wide momentum range. An overview of the IFR is shown in Fig 2.17.

![Figure 2.17: Overview of the IFR: the barrel region (left) and the forward (FW) and backward (BW) doors (right) are shown. Dimensions are in mm.](image)

The main requirements for the IFR are: a large solid angle coverage, good efficiency and high background rejection power for muons with momentum down to 1 GeV/c.

Also the IFR, as all other \textit{BABAR} sub-detectors, has an asymmetric structure and covers a polar angle in the laboratory frame of \(17^\circ \leq \theta_{\text{lab}} \leq 150^\circ\). It is approximately 5.2 m-long and 6.5 m-high, weighting about 870 t, and it consists of an hexagonal barrel region closed in the forward and backward directions by two endcaps that can be opened to access the inner part of the \textit{BABAR} detector. Both the endcaps and the barrel region are layered and active detectors are inserted between the iron plates. The steel is segmented into 18 layers, with thickness varying from 2 cm for the innermost layers to 10 cm for the outermosts. The gaps between the iron plates range from 3.2 cm to 3.5 cm in width. In each endcap 18 layers of planar RPCs [35] are located; in addition two layers with cylindrically shaped RPCs are situated between the EMC and the superconducting solenoid to help with the detection of particles exiting the calorimeter. The layers of
the barrel region, initially filled with planar RPCs, are now equipped with LSTs [36]. Degradation effects that have affected both the barrel and the endcap regions, have led to the decision of replacing some components of the IFR. During 2002, most of the RPCs filling the forward endcap, were replaced with new and better built RPCs, while the remaining gaps were filled with leaded brass to reduce the contamination from hadron decays. In addition, a 2.5 mm steel plate has been placed behind the outermost layer to shield beam background and a belt of RPCs has been added in front of the forward door to improve the solid-angle coverage. Since the forward RPC’s began to show degradation in efficiency near the inner region closest to the beam axis, where the machine background is higher, the central RPCs have been switched to avalanche mode. This resulted in an expected higher strip occupancy, but led to the full recovery of efficiency in the inner region.

In the barrel, all the RPCs have been replaced by LSTs: the choice of using LSTs was made for maintenance simplicity and a guaranteed efficiency tested in two years of development. The upper and lower sextants of the barrel were upgraded during the summer 2004, while the other four sextants operate with LSTs since the end of the summer 2007.

![Section of a planar RPC](image)

Figure 2.18: Section of a planar RPC.

A scheme of a planar RPC section is shown in Fig 2.18. For the streamer RPCs, the gas mixture is made of 60.4% Argon, 34.8% Freon-134a and 4.8% Isobutane and the operating voltage is about 7,500 kV. In the avalanche RPCs the voltage is raised to about 9,200 kV and the gas has the following composition: 22.4% Argon, 72.6% Freon-134a, 4.4% Isobutane, and 0.6% of a sulfur and flourine combination. Iron layers keeping apart
RPC planes are chilled by a water system that keeps the temperature at \( \sim 20^\circ \) C. Signals produced by particles crossing the gas gap inside the RPCs are collected on both sides by using 1 cm wide thin strips (thickness \( \sim 40 \, \mu \text{m} \)). Strips are applied in two orthogonal directions on 200 \( \mu \text{m} \)-thick insulating planes, in order to have a bi-dimensional view. Each RPC module is about 125 cm long along the beam direction, with variable width in order to completely fill the gap. Groups of 3 RPCs, called chambers, are equipped with 96 \( \phi \)-strips placed along the \( z \)-axis that measure the \( \phi \) angle inside the barrel and 96 \( z \)-strips orthogonal to the beam direction. \( z \)-strips are subdivided into 3 panels of 32 strips with width ranging between 1.78 and 3.37 cm as a function of chamber radial position. This projective geometry allows a constant number of strips for all the various layers without deteriorating the detector resolution (each strip covers the same azimuthal angle).

LSTs consist of silver wires with a radius of 50 \( \mu \text{m} \) placed in the middle of cells with dimensions of \( 17 \times 17 \, \text{mm}^2 \). The cells are grouped in modules of 7- or 8-cells, whose outer structure is made of plastic material. A sketch of a 8-cells tube and a picture of the plastic endcap with high voltage and gas connectors are shown in Fig. 2.19. The internal surface of the cells is made of graphite with a resistivity between 0.2 and 1 \( M\Omega/\text{cm}^2 \). In each cell, plastic supports placed every 50 cm keep the wire straight. At the ends of the modules, plastic end-pieces host the gas inlets and the high-voltage pins. The gas mixture is Argon (3%), Isobutane (8%) and Carbon Dioxide (89%). Isobutane is used as a quencher, while carbon dioxide reduces the flammability of the gas mixture. The readout in the \( \phi \) coordinate is possible thanks to electronics located at one end of the tubes. The information on the longitudinal coordinate is obtained by strips placed perpendicularly to the wires. The acceptance of LST is limited by the ratio between active and total volume of the cells. For the dimensions chosen, the physical limit is about 95% in case of normal incidence.

RPCs and LSTs performances are measured by using cosmic data and the overall system reaches an efficiency of about 90\%. The LSTs have proved to work better than RPCs for the muon/pion discrimination.

Fig. 2.20 shows the efficiency map for the layer 10 of the IFR barrel and for layer 15 of the forward endcap.
2.7 The Instrumented Flux Return: IFR

Figure 2.19: Sketch of an 8-cell module (top) and picture of the plastic endcap with high voltage and gas connectors (bottom).

Figure 2.20: Efficiency map for layer 10 of the IFR barrel (top) and layer 15 of the forward endcap (bottom).
2.8 The trigger system

The BABAR trigger system is required to select the events of interest listed in Tab. 2.2, with high, stable, and well-known efficiency while rejecting background down to an overall event rate of 120 Hz. The trigger efficiency should exceed 99% for $B\bar{B}$ events and at least 95% for continuum production. Depending on the $\tau$ specific decay mode, the efficiency for $e^+e^- \rightarrow \tau^+\tau^-$ should be between 90% and 95%. It has to work under high background conditions and with dead or noisy channels.

The trigger system is implemented as a two-level hierarchy: the Level 1 (L1 hardware) and the Level 3 (L3 software). L1 is designed to have very high efficiency, known with a good accuracy, and to have an output rate of typically 1 KHz. L3 receives inputs from L1, perform a second stage rate reduction for the main physics sources and identifies and flags special event categories needed for luminosity determination and calibration purpose. At design luminosity, L3 has a signal rate of about 90 Hz (and 30 Hz for special event categories).

Event signatures are exploited to separate signal from background and the following properties are used in the L1 trigger as general event selection criteria: charged track multiplicity, calorimeter cluster multiplicity and event topology. The parameters used in the selection, that have associated thresholds, are: charged-track transverse momentum, energy of the calorimeter clusters, solid angle separation, and track-cluster match quality. Several trigger definitions are implemented and differ only by the values of thresholds. Specific trigger selection lines are combined in logical “or” to define the global selection for a given trigger level: each line is the result of a boolean operation on any combination of trigger objects, as indicated in Tab. 2.3. In the boolean combination a minimum angular separation may be also required in order to count more than one object (typically 90°). Dedicated trigger lines are used to select back-to-back topologies. The combined Level 1 trigger efficiency is above 99.9% for $B\bar{B}$ events and greater than 95% for the continuum.

Level 3 trigger exploits the L1 trigger information and the full event data for events that passed the L1 trigger; the output to mass storage is the full event and trigger data of events accepted by L3. L3 algorithms have all event information available and they operate by refining and improving the selection methods used by L1. It also includes a variety of filters to perform event classification and background reduction: the logging decision is based on two orthogonal filters, one relying exclusively on DCH data and the
Table 2.3: Some trigger objects for the Level 1 trigger.

<table>
<thead>
<tr>
<th>Object description</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short track reaching DCH super-layer 5</td>
<td>120 MeV/c</td>
</tr>
<tr>
<td>Long track reaching DCH super-layer 10</td>
<td>180 MeV/c</td>
</tr>
<tr>
<td>High $p_T$ track</td>
<td>800 MeV/c</td>
</tr>
<tr>
<td>All-θ MIP energy</td>
<td>100 MeV</td>
</tr>
<tr>
<td>All-θ intermediate energy</td>
<td>250 MeV</td>
</tr>
<tr>
<td>All-θ high energy</td>
<td>700 MeV</td>
</tr>
</tbody>
</table>

The DCH filters select events with one tight ($p_T > 600$ MeV/c) track or two loose ($p_T > 250$ MeV/c) tracks originating from the interaction point: track selection is based on the $x$-$y$ closest approach distance ($d_0$) to the IP and the corresponding $z$ coordinate ($z_0$). Tight (loose) tracks are required to satisfy a vertex condition defined as $|d_0| < 1.0$ cm ($|d_0| < 1.5$ cm) and $|z_0 - z_{IP}| < 7.0$ cm ($|z_0 - z_{IP}| < 10.0$ cm), where $z_{IP}$ is the $z$ coordinate of the interaction point. The EMC filters are dedicated to the selection of events with either high energy deposits ($E_{CM} > 350$ MeV) or high cluster multiplicity (at least 4 clusters): a high effective mass (greater than 1.5 GeV/$c^2$), calculated from the cluster energy sums and the energy weighted centroid positions of all clusters in the event assuming massless particles, is also required. A Bhabha veto filter is used to select one-prong (only a positron in the back part of the detector) and two-prong events (with both $e^+$ and $e^-$ detected) and it applies stringent criteria on EMC energy deposits relying on track momenta and $E/p$ values.

During a typical on-peak run with an average luminosity of $2.6 \times 10^{33}$ cm$^{-2}$s$^{-1}$, the physics events represent 13% of the total L3 output (with a rate of 16 Hz), while the calibration and diagnostic samples are 40% (with a rate 49 Hz): the total output rate is 122 Hz.
Chapter 3

Analysis method and samples

This chapter describes the data sample that has been used and the analysis procedure adopted to study the $B \to K^* \nu \overline{\nu}$ process. In Sec. 3.1 an overview of the analysis flow and the ingredients needed to obtain the physical result are presented. Details about the data and Monte Carlo (MC) samples, along with the simulation process, are given in Sec 3.2.

3.1 Analysis method

The aim of the present work is the search for the $B \to K^* \nu \overline{\nu}$ decay, a Flavor Changing Neutral Current process that, as discussed in Chap. 1, can be sensitive to New Physics effects.

From the experimental point of view, the signature that identifies the signal decay is given by a charged or neutral $K^*$ accompanied by missing energy that would be produced by the neutrino pair. As will be discussed in Chap. 5, the $K^*$ is reconstructed by means of a kaon and a pion with convenient charges. A measurement of the missing energy is feasible only if all the other particle in the event have been reconstructed and it can in principle be associated only to undetected neutrinos.

The $\Upsilon(4S)$ decays to pairs of $B$’s: one of them is used to search for the signal signature ($B_{\text{sig}}$) while the other meson has to be (semi-)completely reconstructed ($B_{\text{tag}}$). Two sets of final states are used in the reconstruction of the $B$ that produces the majority of the particles constituting the event:

- hadronic final states ($B_{\text{had}}$), consisting of $B \to D Y$ decays, where $D$ refers to a charm meson, and $Y$ represents a collection of hadrons with a total charge of $\pm 1$. The method used to identify the $B_{\text{had}}$, called semi-exclusive reconstruction, is
described in Chap. 4.

- semileptonic decays ($B_{sl}$), including $B \rightarrow D^{(*)} l \nu$ ($\ell = e, \mu$) processes.

In this analysis the hadronic decays are searched for and a complementary work has been
performed by using the semileptonic reconstruction [37]: the two datasets are completely
non overlapping and the results can be combined as independent measurements (Sec. 8.3).

Once a $B_{had}$ has been identified, one can look in the rest of the event, or recoil system,
to check the consistency with the signal decay. This technique is called recoil method.

A selection is then applied to the reconstructed $B_{had}$-$B_{sig}$ pairs in order to suppress
background contamination and extract the ingredients necessary to the branching fraction
determination. As described in Sec. 8.1, the branching fraction is given by:

$$\mathcal{B} = \frac{N_{sig}}{\varepsilon_{B_{sig}} \cdot N_{B_{had}}} \cdot \frac{B_{B,MC}}{\varepsilon_{B_{had}} \cdot K_{B,MC}},$$

where $N_{sig}$ is the estimate for the number of signal events, $\varepsilon_{B_{sig}}$ is the efficiency in the
signal side reconstruction and selection, $\varepsilon_{K_{had},MC}$ and $\varepsilon_{B_{had},MC}$ are the $B_{had}$ reconstruction
 efficiencies in events containing the signal process and with generic $B\bar{B}$ decays, respectively;
their ratio is used as a correction factor to account for differences on the $B_{had}$ reconstruction
among the two samples and is determined from MC simulations. The extraction of the $B_{had}$ yield and the correction factor deriving from its reconstruction are
discussed in detail in Sec. 4.3. $N_{sig}$ and $\varepsilon_{B_{sig}}$ have been estimated using two different
approaches: a cut and count analysis and an extended maximum likelihood fit to the
distribution of a Neural Network output.

In the cut and count analysis (Chap. 6), after a loose pre-selection, more stringent and
optimized requirements are applied. Control samples are used to extract the expected
background yield in data; this is subsequently compared to data events surviving the
selection to estimate $N_{sig}$. The efficiency $\varepsilon_{B_{sig}}$ is evaluated by counting the events surviving
the selection in a simulated sample that is generated as containing only events of interest
in the signal side (signal MC).

In the Neural Network analysis (Chap. 7), a loose selection is applied, without optimiz ing
the requirements. The most discriminating variables are used as input for Neural
Networks (one for each reconstructed $K^*$ decay mode); an extended maximum likelihood
fit to its distribution is performed, in order to extract $N_{sig}$; the efficiency $\varepsilon_{B_{sig}}$ is given by
3.2 Data and Monte Carlo samples

Data sample used in this analysis

For this analysis, data collected by BABAR at the $\Upsilon(4S)$ resonance in the period 1999-2007 have been used. The data set is referred to as “Run1-6” on-peak data and corresponds to an integrated luminosity of 413 fb$^{-1}$. Also 41 fb$^{-1}$ of off-peak data collected in the same period have been exploited for continuum modeling. Tab. 3.1 lists for each data taking period (Run) the corresponding years, the integrated luminosity, and the number of events collected.

Monte Carlo sample used in this analysis

Along with real data, large collections of simulated MC samples were generated. Typically a ratio MC : data of 3 : 1 is maintained for on-peak processes and 1 : 1 for off-peak. The simulated samples are used to optimize and test the analysis strategy. Before applying it to real data a comparison with the MC sample is required to estimate the accuracy of MC modeling. Potential discrepancy can be taken into account by systematic uncertainty studies.

The simulation process is divided in several steps. In the first one, a C++ software called GenFwkInt (Generators Framework Interface) [38], that includes the routine EvtGen (Event Generator) [39], generates four-vectors for physical processes: the energy spread

Cross-checks on the quality of the agreement between the data sample and the simulated one are described in Chap. 7. In the same chapter, the procedure adopted to evaluate the systematic uncertainties affecting the branching fraction normalization, the signal efficiency and the signal yield are presented.

In both approaches the procedures are tested and validated in the simulated MC samples and then applied to real data to extract the final results, presented in Chap. 8.

As anticipated, no evidence for signal is found neither in the charged $K^*$ channel nor in the neutral one; Upper Limits (ULs) on the branching fractions have been computed using a Bayesian approach and the results are finally used to constraint NP parameters, as discussed in the last chapter.
Table 3.1: Data sample used in this analysis, split by run period: years of data taking, integrated luminosity, and number of events.

<table>
<thead>
<tr>
<th>Run Period</th>
<th>Data Sample</th>
<th>Int. Lumi. (fb⁻¹)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>off-peak</td>
<td>2.615</td>
<td>32,556,616</td>
</tr>
<tr>
<td>Run 2 (2001-2002)</td>
<td>on-peak</td>
<td>61.076</td>
<td>917,764,240</td>
</tr>
<tr>
<td></td>
<td>off-peak</td>
<td>6.923</td>
<td>96,761,873</td>
</tr>
<tr>
<td>Run 3 (2002-2003)</td>
<td>on-peak</td>
<td>32.278</td>
<td>485,066,397</td>
</tr>
<tr>
<td></td>
<td>off-peak</td>
<td>2.468</td>
<td>34,190,901</td>
</tr>
<tr>
<td>Run 4 (2003-2004)</td>
<td>on-peak</td>
<td>100.282</td>
<td>1,482,896,628</td>
</tr>
<tr>
<td></td>
<td>off-peak</td>
<td>10.121</td>
<td>137,206,866</td>
</tr>
<tr>
<td>Run 5 (2005-2006)</td>
<td>on-peak</td>
<td>132.867</td>
<td>1,970,113,405</td>
</tr>
<tr>
<td></td>
<td>off-peak</td>
<td>14.485</td>
<td>195,179,612</td>
</tr>
<tr>
<td>Run 6 (2007)</td>
<td>on-peak</td>
<td>66.123</td>
<td>1,126,801,686</td>
</tr>
<tr>
<td></td>
<td>off-peak</td>
<td>4.613</td>
<td>79,032,530</td>
</tr>
</tbody>
</table>

allowed in the PEP-II collisions is simulated and used to compute the energy available to the generated particles. The next step consists on generating secondary interactions and the interaction with the detector volume; this is performed by using a BABAR GEANT4-based toolkit [40]. At this stage the particle position and their energy deposit in the detector material is computed by exploiting energy, charge and angular information. Signals that mimic the detector electronics are then generated and the L1 and L3 trigger requirements are applied. Finally the event reconstruction algorithms are applied to the generated samples.

Several physical processes can be simulated to produce a large variety of MC categories. The ones used on this analysis are:

- \( B^+ \to K^{*+}\nu\bar{\nu} \) and \( B^0 \to K^{*0}\nu\bar{\nu} \) signal MC: one \( B \) is forced to decay to \( K^{*}\nu\bar{\nu} \), while the other decays into one of all possible \( B \) final states. The \( K^{*+} \) is generated to decay to one of the allowed final states \( (K^{*+} \to K^{+}\pi^0, K^{0}\pi^+, K^{0}\gamma) \), while the \( K^{*0} \) is forced to decay to \( K^+\pi^- \) and this will be taken into account in the signal reconstruction efficiency computation (Secs. 6.1.1 and 6.2.3). The kinematic of the signal decay is described by a pure phase space model so that no assumption on the physical model is made.

- \( B^+B^- \) and \( B^0\bar{B}^0 \) generic MC: both \( B \)'s decay to any of the allowed final states;
3.2 Data and Monte Carlo samples

- $B^+B^-$ and $B^0\bar{B}^0$ “hadronic cocktail” MC: one $B$ decays generically while the other decays in any of the following states: $B^+ \to D^{(*)}\pi^+$, $B^0 \to D^{(*)}\pi^+$, $D^{(*)}\rho^+$, $D^{(*)}a_1^+$;

- $e^+e^- \to c\bar{c}$, $e^+e^- \to u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $e^+e^- \to \tau^+\tau^-$ MC: they consist of $e^+e^-$ continuum annihilations to $q\bar{q}$ or $\tau^+\tau^-$ final states, without production of the $\Upsilon(4S)$ resonance.

Tab. 3.2 lists the MC samples, the number of generated events, the equivalent integrated luminosity and the weight used to compute the expected numbers of events in the data sample. For the signal MC, the SM prediction for the branching fraction $(B(B \to K^*\nu\bar{\nu}) = 1.3 \times 10^{-5})$ [23] has been assumed. For the cocktail MC samples, the scale factors and the equivalent luminosities have not been computed since are not needed in analysis. Signal MC samples are used to estimate the signal reconstruction efficiency and to parametrize the Probability Density Function (PDF) for signal events in the fit to the Neural Network output. $B^+B^-$ and $B^0\bar{B}^0$ generic, $e^+e^- \to c\bar{c}$, $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $e^+e^- \to \tau^+\tau^-$ allow to characterize background contributions coming from resonant and non resonant $e^+e^-$ annihilation. The $B^+B^-$ and $B^0\bar{B}^0$ hadronic cocktail samples are used for validation studies in the Neural Network analysis.

Table 3.2: MC samples used in this analysis, number of generated events, and equivalent integrated luminosity.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Gen. Evt. ($\times 10^3$)</th>
<th>Equiv. Lumi. (fb$^{-1}$)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \to K^{*+}\nu\bar{\nu}$</td>
<td>7,767</td>
<td>504,350</td>
<td>7.60$\times 10^{-4}$</td>
</tr>
<tr>
<td>$B^0 \to K^{*0}\nu\bar{\nu}$</td>
<td>5,270</td>
<td>342,207</td>
<td>7.46$\times 10^{-4}$</td>
</tr>
<tr>
<td>$B^+B^-$ generic</td>
<td>686,354</td>
<td>1,248</td>
<td>0.3310</td>
</tr>
<tr>
<td>$B^0\bar{B}^0$ generic</td>
<td>662,726</td>
<td>1,273</td>
<td>0.3246</td>
</tr>
<tr>
<td>$B^+B^-$ hadronic cocktail</td>
<td>9,714</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$B^0\bar{B}^0$ hadronic cocktail</td>
<td>78,537</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$e^+e^- \to c\bar{c}$</td>
<td>1,087,216</td>
<td>836</td>
<td>0.4939</td>
</tr>
<tr>
<td>$e^+e^- \to u\bar{u}$, $d\bar{d}$, $s\bar{s}$</td>
<td>895,736</td>
<td>429</td>
<td>0.9637</td>
</tr>
<tr>
<td>$e^+e^- \to \tau^+\tau^-$</td>
<td>382,598</td>
<td>407</td>
<td>1.0148</td>
</tr>
</tbody>
</table>
Analysis method and samples
Chapter 4

The Semi-Exclusive $B$ reconstruction

In this chapter the principles of the semi-exclusive reconstruction of hadronic $B$ decays are discussed. The reconstruction chain starts from charged tracks and neutral clusters detected inside the $B$AR volume and used to identify the $B$ decay products (App. A). The $B_{\text{had}}$ candidates are then selected according to the criteria presented in Sec. 4.1 and 4.2. In Sec. 4.3 the extraction of the $B_{\text{had}}$ yield, used to normalize the branching fraction, is discussed. Finally some details about the reconstruction of the $B_{\text{tag}}$ in semileptonic final states are presented (Sec. 4.4).

4.1 Semi-Exclusive reconstruction of hadronic $B$ final states

Charmed $D$ and $D^*$ mesons, along with $K$ and $\pi$ combinations, are used to completely reconstruct the $B_{\text{had}}$ accompanying the $B_{\text{sig}}$. The aim of the semi-exclusive reconstruction of the $B_{\text{had}}$ is to have a high statistic sample with a background contamination as low as possible. The selected channels are of the type $B \rightarrow D Y$ where $D$ is a charmed meson ($D^0$, $D^+$, $D^{*0}$, $D^{*+}$) and $Y$ represents a collection of hadrons with total charge $\pm 1$ given by $Y = n_1 \pi^\pm + n_2 K^\pm + n_3 K^0_S + n_4 \pi^0$, where $n_1 + n_2 < 6$, $n_3 < 3$, and $n_4 < 3$. Using $D^0(D^+)$ and $D^{*0}(D^{*+})$ as seed for $B^-(\bar{B}^0)$ decays, about 1000 different decay chains are reconstructed. The branching fractions of the $D$ and $D^*$ decay modes used in this analysis are listed in Tab. 4.1.

The algorithm for the $B_{\text{had}}$ reconstruction is articulated in different steps. First a clean input list with charged kaons and pions ($\text{piKList}$), whose elements are called $h$, is defined; tracks identified as kaons are assigned the kaon mass while the others are
Table 4.1: Branching fractions of the $D$ and $D^*$ decay modes used in this analysis [4].

<table>
<thead>
<tr>
<th>Decay</th>
<th>Branching Fraction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^0 \to K\pi$</td>
<td>3.9</td>
</tr>
<tr>
<td>$D^0 \to K\pi\pi^0$</td>
<td>13.9</td>
</tr>
<tr>
<td>$D^0 \to K_S^0\pi\pi$</td>
<td>3.0</td>
</tr>
<tr>
<td>$D^0 \to K\pi\pi\pi$</td>
<td>8.1</td>
</tr>
<tr>
<td>$D^+ \to K\pi\pi$</td>
<td>9.2</td>
</tr>
<tr>
<td>$D^+ \to K\pi\pi\pi^0$</td>
<td>6.0</td>
</tr>
<tr>
<td>$D^+ \to K_S^0\pi$</td>
<td>1.5</td>
</tr>
<tr>
<td>$D^+ \to K_S^0\pi\pi^0$</td>
<td>6.8</td>
</tr>
<tr>
<td>$D^+ \to K_S^0\pi\pi\pi$</td>
<td>3.0</td>
</tr>
<tr>
<td>$D^{*0} \to D^{0}\pi^0$</td>
<td>61.9</td>
</tr>
<tr>
<td>$D^{*0} \to D^{0}\gamma$</td>
<td>38.1</td>
</tr>
<tr>
<td>$D^{*+} \to D^{0}\pi^+$</td>
<td>67.7</td>
</tr>
</tbody>
</table>

treated as pions. Pairs or quartets of charged mesons ($V^0 = h^+h^-$ and $W^0 = h^+h^-h^+h^-$ respectively) are constructed. The list of $B$ candidates is defined in the following steps:

- select a pre-seed ($D$ or $D^*$) and add first one charged of the \texttt{piKList} to form a seed that has the same charge of the $B$ meson and then pick elements from the $V^0$ and $W^0$ samples;
- study the $Y$ system looking for resonance, identify submodes and define subcategories according to the decay multiplicity and structure in order to distinguish the cleanest modes with respect to the ones with highest background contamination;
- determine mode by mode selection criteria to reject misreconstructed $B_{had}$, accounting for different background levels which depend on the number of charged tracks and $\pi^0$ defining the $Y$ system;
- rank submodes with respect to the a-priori purity defined as $S/\sqrt{(S+B)}$, $S$ ($B$) being the number of signal (background) events determined by a fit to the distribution of the kinematical variable $m_{ES}$, defined in the following; study then the a-priori purity as a function of the number of used modes to maximize the statistical significance of the sample;
- group submodes with similar purity;
- in case of multiple $B_{had}$ candidates in the event, select the best one.
4.1 Semi-Exclusive reconstruction of hadronic $B$ final states

The $B_{\text{had}}$ candidates identified with this procedure are then classified and selected, as described in the next section.

**Definition of the kinematical variables $m_{\text{ES}}$ and $\Delta E$**

The two most important variables in the $B_{\text{had}}$ candidate classification and selection are $m_{\text{ES}}$ and $\Delta E$. The beam energy-substituted mass ($m_{\text{ES}}$) is defined as

$$m_{\text{ES}} = \sqrt{E_{\text{beam}}^2 - |p_B^*|^2};$$

(4.1)

where $E_{\text{beam}}$ is the beam energy in the CM frame and $p_B^*$ is the momentum of the $B_{\text{had}}$ in the same frame. For correctly reconstructed $B_{\text{had}}$, the $m_{\text{ES}}$ distribution peaks at the nominal $B$ mass value. Since $|p_B^*| \ll \sqrt{s}/2$, the experimental resolution on $m_{\text{ES}}$ is dominated by beam energy fluctuations. To an excellent approximation, the shapes of the $m_{\text{ES}}$ distributions for $B$ meson reconstructed with charged tracks only are Gaussian. The presence of neutrals in the final states, in case their showers are not fully contained in the calorimeter, can introduce tails.

The energy difference $\Delta E$ is given by:

$$\Delta E = E_B^* - E_{\text{beam}},$$

(4.2)

$E_B^*$ being the energy of the $B_{\text{had}}$ in the CM frame. For correctly reconstructed $B_{\text{had}}$ candidates, the $\Delta E$ distribution peaks at zero. The resolution of this variable is affected by the detector momentum resolution and by the particle identification since a wrong mass assignment results in a shift in $\Delta E$. It depends essentially on the reconstructed $B$ mode and $\pi^0$ multiplicity and it varies from 20 to 40 MeV. Continuum and part of the $B\overline{B}$ background have a $\Delta E$ distribution that can be modeled with a polynomial distribution; instead $B\overline{B}$ background due to misidentification gives shifted Gaussian peaks.

Since the sources of experimental smearing on $m_{\text{ES}}$ and $\Delta E$ are independent (beam energy and detector momentum resolution, respectively), the two are basically uncorrelated.

The two kinematical variables are first used to classify $B_{\text{had}}$ candidates according to the region in the $m_{\text{ES}}-\Delta E$ plane in which they lie; $\Delta E$ then participates the background suppression and the assignment of the best candidate in the event, while $m_{\text{ES}}$ is exploited to determine the a-priori purity and rank submodes.

According to Fig. 4.1, several $m_{\text{ES}}$-$\Delta E$ regions are defined: candidates in the region (A) are considered as $B_{\text{had}}$ candidates and not as a seed since, adding one more pion, the
ΔE value would be pushed out of the allowed physical region; the sample in the region (B) can serve both as \( B_{\text{had}} \) candidates and as seed to be paired with additional pion(s); candidates in the region (C) are out of the allowed physics region but can enter it by adding elements from the piKList, as a consequence they are used as a seed; finally the sample populating the region (D) is not used.

![Diagram of mES-ΔE regions](image)

Figure 4.1: Sketch of the definition of the \( m_{\text{ES}-\Delta E} \) regions. A: used only as candidates; B: used both as candidates and as seeds; C: used as seeds but not as candidates; D: not used.

**Mode by mode background suppression**

The classification based on the \( m_{\text{ES}-\Delta E} \) plane allows to reconstruct all possible \( D(\ast)Y \) decay modes. Once this is done, a detailed study of the reconstructed \( Y \) system is performed by looking for resonances and studying the background contribution. In this way the cleanest mode are identified and the selection of the most efficient and pure among multiple candidates is possible.

An example is given by the mode \( B \to D^{\ast}\pi\pi^0 \): this can receive different contributions depending on the \( \pi\pi^0 \) invariant mass. A large contribution to the signal sample is given by events with \( m(\pi\pi^0) < 1.5 \text{ GeV}/c^2 \) due to the \( \rho \) resonance; another contribution comes from events in the region \( 2.4 < m(\pi\pi^0) < 2.5 \text{ GeV}/c^2 \) that is less clean and less populated than the first sample. Therefore two sub-modes are defined depending on whether \( m_{\pi\pi^0} \) is below or above 1.5 GeV/c², without requiring the sub-mode belonging to a precise
resonance structure (for this reason the $B_{\text{had}}$ reconstruction is called “semi-exclusive”). In this way the clean $B \to D^{*}\pi\pi^0$ sub-mode ($m_{\pi\pi^0} < 1.5$ GeV/$c^2$) has been separated from the low purity ones ($m_{\pi\pi^0} > 1.5$ GeV/$c^2$). Repeating this procedure for all the $D^{(*)}Y$ mode, 1,097 decay chains are reconstructed.

To clean up the sample, a preliminary requirement on $\Delta E$ is applied, the selection on this variable is then refined in a further stage of the analysis (Sec 4.2). The chosen $\Delta E$ window depends on the resolution determined from its distribution before the request of the best candidate in the event. This quantity essentially depends on the number of charged tracks and, above all, on the number of $\pi^0$'s in the $Y$ system (since the reconstructed $D$ is mass-constrained). For the modes without $\pi^0$'s a fit to a linear function that accounts for background contributions plus a Gaussian that describes the signal is performed and a $2\sigma$ symmetric window is taken. When neutral mesons such as $\pi^0$ and $K_S^0$ are present there is an increment of multiple candidates in the event and more detailed studies are required. They have lead to the following selection: $|\Delta E| < 45$ MeV for candidates without $\pi^0$'s and $K_S^0$'s, $|\Delta E| < 50$ MeV for candidates with up to one $\pi^0$ and two $K_S^0$'s and $-90 < \Delta E < 60$ MeV for all others.

Determinant of the a-priori purity and best candidate selection

The most relevant parameter at this stage of the $B_{\text{had}}$ selection is the purity that allows to classify the submodes, in order to choose a subsample of cleanest modes if needed, and select the best $B_{\text{had}}$ candidate in the event. The a-priori purity, given by $S/\sqrt{(S+B)}$, is computed by extracting $S$ and $B$ in a fit to the $m_{ES}$ distribution. The background component is parametrized by an Argus function [41], that provides a good parametrization of continuum ($c\bar{c}$ and $uds$) and $B\bar{B}$ background events. The signal component is fitted to a Crystal Ball function [42]. This has a tail in the lower $m_{ES}$ region that accounts for cases where the energy of the neutral candidates is not fully deposited in the EMC crystals and its shape depends on the reconstructed $B$ mode and in particular on the number of $\pi^0$'s. As an example, Fig. 4.2 shows the fitted shape on the MC sample for modes with no $\pi^0$, one $\pi^0$ an two $\pi^0$'s. The maximum total number of floating parameters in the $m_{ES}$ fits is 7. Two of them are for the ARGUS shape, while the remaining five parameters are for the Crystal Ball.

The a-priori purity and $\Delta E$ allow to select the best $B_{\text{had}}$ in case of multiple candidates in the same event. If there are multiple candidates in the same submode only the one with
the lowest $\Delta E$ is chosen and one candidate per submode is selected. If different submodes are involved $\Delta E$ cannot be used since modes with higher combinatorial background would be privileged with respect to the cleanest ones. An unbiased criterion for choosing a signal event is based on the a-priori purity computed in the signal region $m_{ES} > 5.27$ GeV: the candidate corresponding the highest value is retained.

In some cases, it is useful to optimize the composition of the modes to improve the analysis of the recoil. For this reason the decay modes are ranked according to their purity and are added to the sample of reconstructed $B$’s one at a time. At each addition of a mode the yield increases and the purity mostly decreases. The significance $S/\sqrt{(S + B)}$ is computed as a function of the number of added modes and the best composition can be chosen. An example for the $B^0 \to D^+ Y$ case is shown in Fig. 4.3.

### 4.2 Selection of the hadronic $B$ sample

The semi-exclusive reconstruction provides a large and clean sample in which the $B$ meson is reconstructed in a hadronic mode. In the previous section, it was discussed how the kinematic variables $m_{ES}$ and $\Delta E$ are used for a preliminary selection by defining a $\Delta E$ window and to rank the submodes by means of the a-priori purity.

Misreconstructed $B_{\text{had}}$ mesons can be classified as:

- **combinatorial $B_{\text{had}}$**: the reconstructed $B$ is not a real $B$ meson, but a combination
4.2 Selection of the hadronic $B$ sample

![Graph](image)

Figure 4.3: Dependence of the quality factor $S/\sqrt{S+B}$ as a function of the yield when adding modes for the $B_0 \rightarrow D^+Y$ case. Statistics corresponds to 80 fb$^{-1}$.

of tracks and neutrals that satisfies the requirements of the semi-exclusive reconstruction;

- $B_{\text{had}}$ from continuum production: for these events no $\Upsilon(4S)$ is produced but the $e^+e^-$ pair decays to $\tau^+\tau^-$, $u\bar{u}$, $d\bar{d}$ and $c\bar{c}$; few $\tau^+\tau^-$ events survive the semi-exclusive reconstruction selection, while a consistent amount of $q\bar{q}$ is retained;

To characterize the first class, generic $B^+B^-$ and $B_0\bar{B}_0$ Monte Carlo are generated, while to study the continuum $\tau^+\tau^-$, $uds$ and $c\bar{c}$ simulated samples are used.

The first step of the background suppression is the definition of a “signal region” in the $\Delta E$-$m_{ES}$ plane by the following requirements:

$$
-0.09 < \Delta E < 0.05 \text{GeV}, \\
5.270 < m_{ES} < 5.288 \text{GeV}/c^2.
$$

In Fig. 4.4 the two distributions for different MC contributions and data are plotted. All the comparison plots for data and Monte Carlo in the following contain each MC sample scaled to the dataset luminosity, with the data overlayed.

To suppress the continuum contribution the angular variable $\cos\theta_{B,T}^*$ is used. This is defined as the angle between the direction of the $B_{\text{had}}$ momentum and the thrust axis [43]
of all the signal-side particles. The thrust axis is the direction which maximizes the sum of the longitudinal momenta of the signal-side reconstructed particles. For jet-like events, coming from continuum background, the $\theta_{B,T}$ distribution peaks around $\pm 180^\circ$ while for $B\bar{B}$ events it has a flat shape. The requirement applied is:

$$|\cos \theta_{B,T}^*| < 0.95.$$  \hfill (4.4)

In Fig. 4.4 the $\cos \theta_{B,T}^*$ distribution for different Monte Carlo contributions and data is shown.

In Tab. 4.2 the number of events that survive the selection requirements on $m_{ES}$, $\Delta E$, and $\cos \theta_{B,T}^*$ is listed. The raw $B_{\text{had}}$ reconstruction efficiency ($e_{B_{\text{had}}}^{\text{raw}}$) is computed as the ratio between the number of events in which at least one $B_{\text{had}}$ has been reconstructed and the number of generated (produced) events in the simulated (data) sample.

Figure 4.4: $m_{ES}$ (left), $\Delta E$ (center), and $\cos \theta_{B,T}^*$ (right) distributions before the selection defined in Eqs. 4.3 and 4.4: sum of the charged and neutral signal Monte Carlo samples on the top, background simulated events (histogram) and data (dots) superimposed on the bottom. Concerning the MC components: $\tau^+\tau^-$ in dark gray, $uds$ in green, $c\bar{c}$ in blue, $B^0\bar{B}^0$ in light gray, and $B^+B^-$ in red.
4.3 Determination of the branching fraction normalization

The branching fraction, as anticipated in Sec. 3.1, is normalized to the number of correctly reconstructed $B_{\text{had}}$ ($N_{B_{\text{had}}}$). In principle the $B_{\text{had}}$ yield can be estimated by fitting the $m_{\text{ES}}$ distribution to the sum of a Argus and a Crystal Ball, but previous detailed studies have established that this does not yield satisfactory results, mainly due to a poor description of the non peaking $B\bar{B}$ contribution by means of an Argus function. The procedure adopted in this work consists on extracting the combinatorial and continuum shapes of different background sources from the MC sample, find an appropriate normalization by means of on-peak data and off-peak data that lie far from the $m_{\text{ES}}$ signal region defined in Eq. 4.3, and finally determine the $B_{\text{had}}$ signal yield by subtracting combinatorial and continuum contributions in the $m_{\text{ES}}$ data signal region.

In this stage of the analysis, the $B_{\text{had}}$ selection efficiency also is computed, both for $B^+B^-$ ($B^0\bar{B}^0$) generic MC and charged (neutral) signal MC samples: the ratio between the two gives a correction factor that will be used in the BF estimation (Sec. 8.1).

To distinguish between signal and combinatorial events in $B\bar{B}$ and signal MC samples, MC truth information are used; a $B_{\text{had}}$ is defined as truth-matched if it satisfies the following requirements:

- all reconstructed tracks on the tag side, except at most two, are associated to the products of the true $B_{\text{had}}$ candidate;

<table>
<thead>
<tr>
<th>Sample</th>
<th>$B_{\text{had}}$ Reco.</th>
<th>$\varepsilon_{B_{\text{had}}}^{\text{raw}}$ ($%$)</th>
<th>$m_{\text{ES}}$</th>
<th>$\Delta E$</th>
<th>$\cos \theta_{B,T}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \rightarrow K^+ \nu \bar{\nu}$ signal</td>
<td>32,236</td>
<td>74.6</td>
<td>23,329</td>
<td>22,709</td>
<td>21,176</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0 \nu \bar{\nu}$ signal</td>
<td>17,494</td>
<td>67.4</td>
<td>10,881</td>
<td>10,607</td>
<td>9,902</td>
</tr>
<tr>
<td>$B^+B^-$ generic decays</td>
<td>31,694,949</td>
<td>68.1</td>
<td>7,453,079</td>
<td>7,088,203</td>
<td>6,552,922</td>
</tr>
<tr>
<td>$B^0\bar{B}^0$ generic decays</td>
<td>28,112,657</td>
<td>68.0</td>
<td>5,777,079</td>
<td>5,495,105</td>
<td>5,059,558</td>
</tr>
<tr>
<td>$e^+e^- \rightarrow \ell\bar{\ell}$</td>
<td>37,049,726</td>
<td>61.2</td>
<td>5,185,759</td>
<td>4,768,042</td>
<td>3,064,528</td>
</tr>
<tr>
<td>$e^+e^- \rightarrow u\bar{u}, d\bar{d}, s\bar{s}$</td>
<td>18,015,748</td>
<td>58.1</td>
<td>2,650,015</td>
<td>2,423,116</td>
<td>1,514,643</td>
</tr>
<tr>
<td>$e^+e^- \rightarrow \tau^+\tau^-$</td>
<td>36,432</td>
<td>50.9</td>
<td>6,262</td>
<td>5,536</td>
<td></td>
</tr>
<tr>
<td>on-peak</td>
<td>60,114,924</td>
<td>62.6</td>
<td>10,043,090</td>
<td>9,356,151</td>
<td>7,072,914</td>
</tr>
<tr>
<td>off-peak</td>
<td>3,227,608</td>
<td>55.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.2: Number of events that pass the semi-exclusive reconstruction requirements and the selection on $m_{\text{ES}}$, $\Delta E$, and $\cos \theta_{B,T}$ for MC and data samples.
- $\Delta \theta_B \leq 0.07 \text{ rad}$, $\Delta \theta_B$ being the angle between the directions of true $B_{\text{had}}$ and reconstructed candidate.

In Fig. 4.5 the distributions of $m_{ES}$ and $\Delta \theta_B$ for events in which all tracks are matched in $B \overline{B}$ generic MC sample are shown; Fig. 4.6 shows the same for events in which one or two tracks are not matched. The peak in the $m_{ES}$ distributions in Fig. 4.6 (left) is mainly due to events in which two tracks are not matched: to satisfy the charge requirement, the two should have opposite charge and in some cases they can be the daughters of a neutral particle (i.e. $K^0_S$) originating from the $B_{\text{had}}$, that as a consequence is correctly reconstructed.

![Figure 4.5](image)

Figure 4.5: $m_{ES}$ (left) and $\Delta \theta_B$ (right) for events in which all tracks are matched; $B^+B^-$ on the top, $B^0\overline{B}^0$ on the bottom.

To evaluate $N_{B^+_{\text{had}}}$ and $N_{B^0_{\text{had}}}$, and the related reconstruction efficiencies, the combinatorial background from $B \overline{B}$, and the continuum contributions from $c\overline{c}$ and $uds$ should be estimated and subtracted (the $\tau^+\tau^-$ component has been found to be negligible). The full procedure consists of the following steps:

- separate the correctly reconstructed and the combinatorial component of $B \overline{B}$ using the MC truth criteria listed above;
4.3 Determination of the branching fraction normalization

Figure 4.6: $m_{ES}$ (left) and $\Delta\theta_B$ (right) for events in which one or two tracks are not matched; $B^+B^-$ on the top, $B^0\bar{B}^0$ on the bottom.

- determine the shape of $m_{ES}$ distributions from MC events for all background sources (combinatorial $BB$, $cc$, and $uds$); note that a requirement on the charge of the reconstructed $B_{had}$ is made, so events from $B^+B^-$ generic MC whose reconstructed $B_{had}$ is neutral contribute to the combinatorial part of $B^0\bar{B}^0$ sample (and vice-versa);

- extrapolate the expected fraction of continuum events in the $m_{ES}$ signal region by counting the on on-peak data, using the relation:

$$f_{\text{cont}} = \frac{L_{\text{on-peak}}}{L_{\text{off-peak}}} \cdot \frac{N_{\text{off-peak}}(5.20 \leq m_{ES} \leq 5.24 \text{ GeV}/c^2)}{N_{\text{on-peak}}(5.22 \leq m_{ES} \leq 5.26 \text{ GeV}/c^2)}$$

(4.5)

where the convenient $m_{ES}$ region are chosen, to take into account for different beam energies between on-peak and off-peak. The normalizations of $c\bar{c}$ and $uds$ are then separately computed accounting for the production cross section and the number of generated MC events, the following values have been found:

- charged tag:

$$f_{uds} = 35.47\%$$

$$f_{c\bar{c}} = 35.77\%$$
The Semi-Exclusive $B$ reconstruction

- neutral tag:

\[ f_{uds} = 26.80\% \]
\[ f_{\pi} = 33.64\% \]

- extract the relative normalization of $B^+B^-$ and $B^0\bar{B}^0$ by a $\chi^2$ fit to the $m_{ES}$ distribution in the on-peak $m_{ES}$ sideband defined as $m_{ES} \in [5.22, 5.26]$ GeV/$c^2$; the fits yields:

- charged tag:

\[ f_{B^+B^-} = (2.31 \pm 1.6)\% \]
\[ f_{B^0\bar{B}^0} = (26.45 \pm 1.6)\% \]

- neutral tag:

\[ f_{B^+B^-} = (27.70 \pm 2.40)\% \]
\[ f_{B^0\bar{B}^0} = (11.86 \pm 2.40)\% \]

where the uncertainty error on $f_{B^+B^-}$ comes from the fit and the one in $f_{B^0\bar{B}^0}$ (computed as $1 - f_{uds} - f_{\pi}$) from the error propagation.

- once the shapes and the normalization of all the background components are determined, compute $N_{B^+_\text{had}}$ and $N_{B^0_{\text{had}}}$ using a bin-by-bin subtraction of the fitted background histogram from the $m_{ES}$ data distribution in the signal region.

The peaking and combinatorial $m_{ES}$ distributions, as determined by the MC truth criteria, are reported in Fig. 4.7. A peaking component is present in the combinatorial shape of the $B^0\bar{B}^0$ sample and it could lead to an underestimation of the $B_{\text{had}}$ yield, a systematic error to account for this effect has been computed (Sec. 7.2.2). In Fig. 4.8 the fits to the $m_{ES}$ sideband to determine the relative fraction of charged and neutral $B\bar{B}$ combinatorial are shown. In Fig. 4.9 the fits to the whole $m_{ES}$ region and the $m_{ES}$ background subtracted histograms for charged and neutral $B_{\text{had}}$ are plotted.

This procedure leads to the following results:

- charged tag:

\[ N_{B^+_\text{had}} = 1,004,750 \pm 1,002 \text{(stat)} \]
\[ \varepsilon_{B^+_\text{had}} = (2.2121 \pm 0.0022) \times 10^{-3} \]
4.4 Semileptonic B reconstruction

- neutral tag:

\[ N_{B_{\text{had}}^0} = 610,111 \pm 781(\text{stat}) \]
\[ \varepsilon_{B_{\text{had}}^0} = (1.3432 \pm 0.0018) \times 10^{-3} \]

The MC truth requirements are applied also to the signal and the generic \( B\bar{B} \) samples. The first allow to compute the normalization to the signal efficiency; while the ratio of the \( B_{\text{had}} \) selection efficiencies between the two is used to correct the \( B_{\text{had}} \) yield (Sec. 8.1). The results on the MC samples are the following:

- charged tag:

\[ N^{K^* \nu \pi, \text{MC}}_{B_{\text{had}}^+} = 19,373 \pm 139(\text{stat}) \]
\[ \varepsilon^{K^* \nu \pi, \text{MC}}_{B_{\text{had}}^+} = (2.494 \pm 0.018) \times 10^{-3} \]
\[ N^{B\bar{B}, \text{MC}}_{B_{\text{had}}^+} = 3,397,479 \pm 1,843(\text{stat}) \]
\[ \varepsilon^{B\bar{B}, \text{MC}}_{B_{\text{had}}^+} = (2.4750 \pm 0.0013) \times 10^{-3} \]

- neutral tag:

\[ N^{K^* \nu \pi, \text{MC}}_{B_{\text{had}}^0} = 8,856 \pm 94(\text{stat}) \]
\[ \varepsilon^{K^* \nu \pi, \text{MC}}_{B_{\text{had}}^0} = (1.680 \pm 0.018) \times 10^{-3} \]
\[ N^{B\bar{B}, \text{MC}}_{B_{\text{had}}^0} = 1,960,847 \pm 1,400(\text{stat}) \]
\[ \varepsilon^{B\bar{B}, \text{MC}}_{B_{\text{had}}^0} = (1.4285 \pm 0.0010) \times 10^{-3} \]

4.4 Semileptonic B reconstruction

Other than hadronic modes, also a set of semileptonic final states have been used to reconstruct the tag \( B \) (\( B_{\text{sl}} \)). The search for the \( B \to K^* \nu \pi \) in the recoil of a semileptonic \( B \) decay has been performed in a related analysis [37].
Figure 4.7: $m_{ES}$ distributions for correctly reconstructed (left) and combinatorial (right) $B\bar{B}$ events as determined by the MC truth requirements; $B^+B^-$ on the top, $B^0\bar{B}^0$ on the bottom.

Figure 4.8: $m_{ES}$ fit in the $m_{ES}$ sideband to determine the relative fraction of charged and natural $B\bar{B}$ combinatorial components; charged $B_{\text{had}}$ on the left, neutral $B_{\text{had}}$ on the right; $c\tau$ in green, $uds$ in blue, $B^+B^-$ in red, and $B^0\bar{B}^0$ in magenta.
Figure 4.9: $m_{ES}$ background PDFs and data distributions in the whole $m_{ES}$ region (top) and background subtracted histograms (bottom); charged $B_{\text{had}}$ on the left, neutral $B_{\text{had}}$ on the right; in the top plots $c\bar{c}$ in green, $uds$ in blue, $B^+B^-$ in red, and $B^0\bar{B}^0$ in magenta.
The Semi-Exclusive B reconstruction

The decays used to identify the $B_{\text{sl}}$ are of the type $B \rightarrow D^{(*)} l \nu$, where $l$ is a muon or an electron. The charmed mesons are reconstructed and selected in the same modes described in App. A, except $D^+ \rightarrow K^-\pi^+\pi^+\pi^0$, $D^+ \rightarrow K_s^0\pi^+\pi^0$, $D^+ \rightarrow K_s^0\pi^+\pi^-\pi^+$, and $D^{*0} \rightarrow D^0\pi^0$. The charmed meson is paired to a lepton with a typical momentum in the CM frame greater than 1.35 GeV/c. The $Dl$ pair is subject to a vertex fit and a requirement on the $\chi^2$ vertex probability is applied.

The kinematic variables that play the main role in separating signal $B_{\text{sl}}$ candidates and misreconstructed tag $B$ are $\cos\theta_{B,Dl}$ and the $D$ invariant mass. The first is defined as the cosine of the angle between the direction of the $B$ and the vector resulting from the sum of the $D$ and $l$ momentum, calculated in the $\Upsilon(4S)$ rest frame. The $B$ momentum can not be calculated directly but is inferred by the beam energies in the CM frame, by assuming that the only unreconstructed particle is the neutrino. The $B$ energy is defined as:

$$E_B = \frac{1}{2} \sqrt{(P_e + P_p)^2},$$  \hspace{1cm} (4.6)

$P_e$ and $P_p$ being the electron and positron four-momenta respectively. The magnitude of the $B$ momentum is then calculated from the $B$ energy and mass as:

$$|\vec{p}_B| = \sqrt{E_B^2 - m_B^2}.$$  \hspace{1cm} (4.7)

The variable $\cos\theta_{B,Dl}$ is defined as:

$$\cos\theta_{B,Dl} = \frac{\vec{p}_B \cdot \vec{p}_{D1}}{|\vec{p}_B||\vec{p}_{D1}|},$$  \hspace{1cm} (4.8)

and assuming that the neutrino is the only unreconstructed particle, the former can be written as:

$$\cos\theta_{B,Dl} = \frac{2\vec{E}_B \vec{E}_{D1} - m_B^2 - m_{D1}^2}{2|\vec{p}_B||\vec{p}_{D1}|}.$$  \hspace{1cm} (4.9)

In the selection, $\cos\theta_{B,Dl}$ is required to lie in the region $[-2.5,1.1]$. The two unphysical boundaries account for reconstruction and resolution effects and unreconstructed particles other than neutrinos. Further selection requirements on the $D$ and $D^*$ invariant masses, on the lepton momentum in the CM frame and on the $D$ and $D^*$ $\chi^2$ vertex probability are applied.

The $B_{\text{sl}}$ reconstruction has an efficiency of $6.6 \times 10^{-3}$, combining both the neutral and the charged channels and the whole sample consists of about 40.9 million events. The higher tag efficiency with respect to the hadronic reconstruction is due to the higher
branching fractions of the $B \to D^{(*)} l \nu$ decays (16.7% [4]) compared to the hadronic $B$ modes and to the fact that the reconstruction, in the semileptonic case, is not completely exclusive due to the massless neutrino.
Chapter 5

Reconstruction of the $B^0(+) \rightarrow K^{*0}(+) \nu \bar{\nu}$ decay

Once the $B_{\text{tag}}$ has been selected, the data sample needs to be examined in order to identify a low multiplicity final state in the recoil system. The signal events are characterized by:

- a number of charged tracks equal to the one needed in $K^*$ reconstruction,
- few neutral objects and energy collected in the electromagnetic calorimeter,
- missing energy, that would be produced by the neutrino pair.

The $K^*$ candidates are reconstructed by means of charged and neutral kaons and pions as described in Secs. 5.1 and 5.2. The pre-selection strategy and the selection variables used in the following steps of the analysis are discussed in Secs. 5.3 and 5.4.

5.1 Kaon and pion reconstruction and selection

Charged tracks and neutral clusters used to reconstruct the signal side decays are subjected to a slightly different selection with respect to the one described in App. A, in order to maximize the signal side reconstruction efficiency and maintain a tolerable background contamination. A summary of the selection criteria is presented in Tab. 5.1.

Single EMC bumps not associated to any track are identified as clusters and assigned to the photon mass hypothesis. The variables exploited in the photon reconstruction are: the number of associated EMC clusters ($N(\text{EMC clus})$), the minimum energy deposit, LAT and $\Delta \alpha$ (both defined in Sec. A.1), and $\theta_{\text{clus}}$ which is the polar angle of the cluster in the laboratory frame.
Reconstruction of the $B^{0(+)} \rightarrow K^{*0(+)\nu\bar{\nu}}$ decay

Charged tracks identified as kaons are subject to requirements on the likelihood for each particle hypothesis ($L_K$, $L_{\pi}$, and $L_p$) and on the probability of being identified as an electron. Charged tracks not identified as kaons, electrons, muons or protons are assigned the pion mass hypothesis.

Neutral pions are reconstructed by means of two photons, constrained to originate from the same vertex. The vertexing algorithm is called Add4 and consists of a four-vector addition of the cluster candidates to re-assign the $\pi^0$ four-momentum. The sample is cleaned up by choosing proper windows for the $\pi^0$ reconstructed and reassigned mass.

Neutral kaons are reconstructed in the $K^0_S \rightarrow \pi^+\pi^-$ and $K^0_S \rightarrow \pi^0\pi^0$ decays. In the charged pion mode, the vertex fit algorithm TreeFitter is applied: this is designed to fit the entire decay tree simultaneously using a Kalman filter technique [44]. Before the vertex fit, requirements on the sum of the four-momenta of the two pions and on the $\pi^+\pi^-$ invariant mass are imposed; the latter is refined after the vertex fit and again in further steps of the selection (Secs. 6.1.1 and 6.2.2). When using the $\pi^0\pi^0$ channel for the $K^0_s$ reconstruction, Add4 is used as vertex fit and a signal region for the $\pi^0\pi^0$ invariant mass after the fit is defined. To maximize the $K^0_S \rightarrow \pi^0\pi^0$ reconstruction efficiency the selection is looser with respect to the $K^0_S \rightarrow \pi^+\pi^-$ case, that is cleaner due to the presence of two charged tracks and any neutral clusters.

5.2 $K^*$ reconstruction and selection

The $K^*$ candidates are reconstructed by kaon-pion pairs, passing the selection requirements listed in Tab. 5.1, in the following combinations: $K^{*+} \rightarrow K^+\pi^0$, $K^{*+} \rightarrow K^0_S\pi^+$, and $K^{*0} \rightarrow K^+\pi^-$. The branching fraction of the reconstructed $K^*$ decay modes are listed in Tab 5.2. Accurate studies, described in Sec. 5.2.1, have been performed to choose the best vertexing algorithm, that has turned out to be Add4. Before and after the vertexing, the proper $m(K\pi)$ invariant mass region is selected (Tab. 5.3).

5.2.1 Studies on the $K^*$ vertex fit

The Add4 algorithm has been used to compute the $K^*$ vertex. Before choosing this final configuration, several algorithms have been tested.

In order to improve the vertex determination by applying a fit, TreeFitter has been used with the following constraints:
5.2 \(K^*\) reconstruction and selection

Table 5.1: Summary of the selection criteria used for the particle reconstruction in the signal side.

<table>
<thead>
<tr>
<th>Selection Criteria</th>
<th>(\gamma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clusters in the EMC</td>
<td>(N(\text{EMC clus}) &gt; 3)</td>
</tr>
<tr>
<td>Neutral energy</td>
<td>(E_\gamma &gt; 50\ \text{MeV})</td>
</tr>
<tr>
<td>LAT</td>
<td>(\text{LAT} &lt; 0.6)</td>
</tr>
<tr>
<td>unmatched clusters</td>
<td>(\Delta \alpha &gt; 0.08)</td>
</tr>
<tr>
<td>(\theta_{\text{clus}})</td>
<td>(0.32 &lt; \theta_{\text{clus}} &lt; 2.44\ \text{rad})</td>
</tr>
</tbody>
</table>

- the \(K^*\) daughters are forced to originate from the interaction point taking its error into account. The error is assigned as the RMS of the beam spot size, which is about 10 \(\mu\)m in the \(y\) direction, 200 \(\mu\)m in \(x\), and 8 mm along the \(z\)-axis,
- pre-fit and post-fit mass constraints are defined with the same values listed in Tab. 5.3.

In Tab. 5.4 the comparison between this configuration and the Add4, for the charged signal

Table 5.2: Branching fraction of the reconstructed \(K^*\) mode [4].

<table>
<thead>
<tr>
<th>(K^*) decay mode</th>
<th>BF</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K^{*0} \rightarrow K^+ \pi^-)</td>
<td>66.57%</td>
</tr>
<tr>
<td>(K^{*+} \rightarrow K^{0}_S (\pi^+ \pi^-)\pi^+)</td>
<td>33.30%</td>
</tr>
<tr>
<td>(K^{*+} \rightarrow K^{0}_S (\pi^0 \pi^0)\pi^+)</td>
<td>23.04%</td>
</tr>
<tr>
<td>(K^{*+} \rightarrow K^{0}_S (\pi^0 \pi^0)\pi^+)</td>
<td>10.22%</td>
</tr>
</tbody>
</table>
Reconstruction of the $B^0(+) \rightarrow K^{*0(+)\nu\bar{\nu}}$ decay

Table 5.3: Summary of the selection criteria used for the reconstruction of the $K^*$ candidates. $m_{K^*}^{PDG}$ refers to the PDG $K^*$ mass that is 0.892 GeV/$c^2$ for $K^{*+}$ and 0.896 GeV/$c^2$ for $K^{*0}$ [4].

<table>
<thead>
<tr>
<th>Selection Criteria</th>
<th>$K^{*\pm} \rightarrow K^{\pm}\pi^0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite particle reconstruction mode</td>
<td>$\pi^0 \rightarrow \gamma \gamma$</td>
</tr>
<tr>
<td>pre-fit mass constraint</td>
<td>$m(K^\pm\pi^0) = m_{K^*}^{PDG} \pm 0.150$ GeV/$c^2$</td>
</tr>
<tr>
<td>post-fit mass constraint</td>
<td>$m(K^\pm\pi^0) = m_{K^*}^{PDG} \pm 0.075$ GeV/$c^2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$K^{*\pm} \rightarrow K^0_{S}\pi^\pm$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite particle reconstruction mode</td>
</tr>
<tr>
<td>pre-fit mass constraint</td>
</tr>
<tr>
<td>post-fit mass constraint</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$K^{*0} \rightarrow K^{+}\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-fit mass constraint</td>
</tr>
<tr>
<td>post-fit mass constraint</td>
</tr>
</tbody>
</table>

Table 5.4: Comparison between Add4 and TreeFitter vertex fit algorithm for the $K^*$ reconstruction in terms of events with a successful fit. The significance is computed by reweighting the MC events to the expected luminosity and branching fraction.

<table>
<thead>
<tr>
<th></th>
<th>$B^+ \rightarrow K^{*+}\nu\bar{\nu}$</th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$S/\sqrt{B}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>generated events</td>
<td>14,804</td>
<td>100,000</td>
<td>100,000</td>
<td>100,000</td>
<td>100,000</td>
<td></td>
</tr>
<tr>
<td>Add4</td>
<td>822</td>
<td>1,014</td>
<td>754</td>
<td>185</td>
<td>162</td>
<td>2.18×10⁻³</td>
</tr>
<tr>
<td>TreeFitter</td>
<td>661</td>
<td>988</td>
<td>749</td>
<td>201</td>
<td>194</td>
<td>1.72×10⁻³</td>
</tr>
</tbody>
</table>

MC and small samples of background, is reported. As can be noticed, with TreeFitter there is a decrease of the significance, defined as $S/\sqrt{B}$, with respect to Add4 since the background is almost unchanged while the tighter $K^*$ selection reduces the signal sample. The $S/\sqrt{B}$ trend is almost unchanged when applying further selection criteria on the $K^*$ decay mode.

The second fit vertex algorithm tested is called VtxTagBtaSelFit and consists of the following steps:

- once the $B_{sig}$ has been reconstructed, a list of charged particles, $\gamma$, and neutral particles ($V_0$) not assigned to the signal side is defined;

- the tag $B$ is reconstructed starting from the previously defined list and a fit to its vertex is performed;
5.3 Signal side pre-selection

- if the fits is successful the event is accepted, otherwise it is rejected.

Using this approach, only few events (both for signal and background MC samples) fail the fit. This is due to the fact that, using the recoil technique, the $B_{\text{sig}}$ is inferred from the tag $B$ (whose decay products are completely reconstructed) and as a consequence the two $B$’s are highly “correlated”, i.e. the tag $B$ built from the list of particle not involved in the signal side is very well-defined and the vertex fit always converges.

5.3 Signal side pre-selection

The events in which a $B_{\text{had}}$ and at least a $K^*$ are reconstructed are subjected to requirements on track and neutral multiplicity:

- number of tracks in the signal side $\leq 6$
- number of clusters in the signal side $\leq 10$
- number of reconstructed $\pi^0$ in the signal side $\leq 6$
- number of reconstructed $K^0_S$ in the signal side $\leq 5$

Events satisfying the previous criteria constitutes the “reconstructed” sample. The next step of the selection consists on choosing the best $K^*$ candidate in each event, if more than one is present. In Fig. 5.1 the $K^*$ multiplicity for signal, background MC samples and data is shown. As best candidate, the one with minimum $\Delta m_{K^*}$ is chosen, where $\Delta m_{K^*}$ is defined by means of the reconstructed $K^*$ mass ($m_{K^*}$) and the nominal PDG value ($m_{K^*}^{PDG}$):

$$\Delta m_{K^*} = |m_{K^*}^{PDG} - m_{K^*}|.$$  \hspace{1cm} (5.1)

The sample is then split in four subsamples, according to the reconstructed $K^*$ and $K^0_S$ decay modes: $K^{*+} \rightarrow K^+\pi^0$, $K^{*+} \rightarrow K^0_S(\pi^+\pi^-)\pi^+$, $K^{*+} \rightarrow K^0_S(\pi^0\pi^0)\pi^+$, and $K^{*0} \rightarrow K^+\pi^-$. For each of them the number of reconstructed tracks should match the number of tracks necessary to the $K^*$ reconstruction (i.e. two reconstructed tracks for $K^{*0} \rightarrow K^+\pi^-$ events). Flavor correlation between the $B_{\text{had}}$ and the $B_{\text{sig}}$ is requested. The meaningful combinations, from the physics point of view, are listed in Tab. 5.5. They lead to the following selection: the charge of the $B_{\text{had}}$ is requested to be opposite to the $K^*$ charge.
Reconstruction of the $B^{0(+) \to K^{*0(+)\nu\bar{\nu}}}$ decay

Figure 5.1: $K^*$ multiplicity in the $m_{ES}-\Delta E-Cos\theta_{B,T}$ signal region: signal MC on the left, generic MC (histogram) and data (dots) on the right. Concerning the MC components: $\tau^+\tau^-$ in dark gray, $uds$ in green, $c\bar{c}$ in blue, $B^0\bar{B}^0$ in light gray, and $B^+B^-$ in red. Note that the events must have a minimum of one suitable $K^*$ to pass the event pre-selection.

Table 5.5: Allowed combinations of the $B_{\text{had}}$ and $B_{\text{sig}}$ decays.

<table>
<thead>
<tr>
<th>$B_{\text{had}}$</th>
<th>$B_{\text{sig}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y^-D^{(*)+}$</td>
<td>$B^0 \to K^{*0}(K^+\pi^-)\nu\bar{\nu}$</td>
</tr>
<tr>
<td>$Y^+D^{(*)-}$</td>
<td>$\bar{B}^0 \to K^{*0}(K^-\pi^+)\nu\bar{\nu}$</td>
</tr>
<tr>
<td>$Y^-D^{(*)0}$</td>
<td>$B^+ \to K^{*+}(K^+\pi^0, K^0\pi^+)\nu\bar{\nu}$</td>
</tr>
<tr>
<td>$Y^+\bar{D}^{(*)0}$</td>
<td>$B^- \to K^{*-}(K^-\pi^0, K^0\pi^-)\nu\bar{\nu}$</td>
</tr>
</tbody>
</table>

for the $K^{*+}$ modes, while for neutral $K^*$ the charge of the the $B_{\text{had}}$ should be zero and the charge of the daughter kaon should be opposite to the $B_{\text{had}}$ charge. The charge correlation requirements lead to a 20% signal reduction for the $K^{*0}$ sample and a negligible loss for the $K^{*+}$ channels; the background suppression goes from 40% for the charged channel and 70% for the neutral. Event losses in the neutral sample are due to the fact that the possibility of $B^0-\bar{B}^0$ mixing is neglected.

The summary of the pre-selection criteria on the charged multiplicity and the charge correlation is given in Tab. 5.6. In Tab. 5.7 the number of events satisfying the pre-selection for each reconstructed $K^*$ mode and for each MC sample is reported. The amount of signal MC events surviving the $K^{*+} \to K_S^0(\pi^0\pi^0)\pi^+$ selection is very low compared to the other $K^*$ channels and also the background contamination is higher. For this reason, the $K^{*+} \to K_S^0(\pi^0\pi^0)\pi^+$ has not been included in the final result and will not be treated in the next sections of this thesis.
5.4 Selection variables

Both in the Neural Network and in the cut and count analysis, variables related to the global event and to the signal side are exploited. In this section a description of such variables and their distributions for signal MC, background MC samples, and data are reported (Sec 5.4.1). The plots are made after the pre-selection requirements.

As anticipated in Sec. 3.2, the signal MC is generated according to a phase-space distribution. A given physics model, i.e. the Standard Model or a New Physics scenario, could modify the kinematics of the process and the final result with respect to the phase-space prediction. Since the aim of this work is to give a model independent results for the $B \to K^* \nu \overline{\nu}$ branching fraction, the model-independence of the selection variables should be tested. This is discussed in Sec. 5.4.2.

5.4.1 Definition of the selection variables

The variable related to the global event properties, used in the selection, is:

- $R_2$ [45] (Fig. 5.2): it is the ratio between zeroth and second order Fox-Wolfram moments and describes the sphericity of the event. It is defined in the range [0,1]
and for jet-like events has a distribution shifted to higher values with respect to $B\bar{B}$ decays.

The selection variables related to the signal side of the events are the following:

- $E_{\text{extra}}$ (Fig. 5.3): it is defined as:

$$E_{\text{extra}} = E_{\text{tot}} - E_{B_{\text{had}}} - E_{B_{\text{sig}}}, \quad (5.2)$$

where $E_{\text{tot}}$ is the total energy of the event, $E_{B_{\text{had}}}$ and $E_{B_{\text{sig}}}$ are the energy of tag and signal $B$ respectively.

Defining the energy on the signal side as $E_{\text{sigSide}} = E_{\text{tot}} - E_{B_{\text{had}}}$, Eq. 5.2 becomes:

$$E_{\text{extra}} = E_{\text{sigSide}} - E_{B_{\text{sig}}}$$

$$= (E_{\text{clusSig}} + E_{\text{trkSig}}) - (E_{\text{clus}B_{\text{sig}}} + E_{\text{trk}B_{\text{sig}}}), \quad (5.3)$$

having split the total energy of signal side and of $B_{\text{sig}}$ in cluster and track contributions. In the pre-selection stage the exact number of tracks used for $K^*$ reconstruction is required. This condition simplifies the $E_{\text{extra}}$ definition, since $E_{\text{tracks}B_{\text{sig}}} \equiv E_{\text{tracks}B_{\text{sig}}}$. Letting $\text{extraCluster}$ be a neutral on the signal side not involved in the $K^*$ reconstruction, the final definition of $E_{\text{extra}}$ becomes:

$$E_{\text{extra}} = E_{\text{clusSig}} - E_{\text{clus}B_{\text{sig}}}$$

$$= \sum_{\text{extraClus}} E(\text{extraCluster}). \quad (5.4)$$

Its distribution peaks at zero for the signal while in the background it peaks at higher values;

- $K^*$ mass (Fig. 5.4);

- $K^*_{S}^{0}$ mass (Fig. 5.5), for $K^*+ \rightarrow K^*_{S}^{0}\pi^+$;

- $E^*_{\text{miss}} + p^*_{\text{miss}}$ (Fig. 5.6): these two variables, here computed in the CM frame, are the time and spatial component of the missing four-momentum, defined as

$$P_{\text{miss}} = P_{B_{\text{sig}}} - P_{\text{tracks}B_{\text{sig}}} - P_{\text{clus}B_{\text{sig}}}, \quad (5.5)$$

$P_{B_{\text{sig}}}$ being the $B_{\text{sig}}$ four-momentum, $P_{\text{tracks}B_{\text{sig}}}$ and $P_{\text{clus}B_{\text{sig}}}$ the four-momenta sums of all the tracks and clusters assigned to the signal side. For signal events it characterizes the kinematics of the invisible $\nu\bar{\nu}$ pair. In the $\Upsilon(4S)$ rest frame, the spatial
5.4 Selection variables

and time component of $P_{\text{miss}}$ are given by:

$$|\vec{p}_{\text{miss}}^*| = |\vec{p}_{\text{had}}^* - \vec{p}_{K^*}| \simeq |\vec{p}_{K^*}|,$$

$$E_{\text{miss}}^* = E_{\text{had}} - E_{K^*} \simeq m_{\text{had}} - \sqrt{p_{K^*}^2 + m_{K^*}^2}, \quad (5.6)$$

having assumed as $B_{\text{had}}$ four-momentum $P_{\text{had}} = (m_B, 0)$, since in the $\Upsilon(4S)$ rest frame the two $B$’s are produced at rest. The sum $E_{\text{miss}}^* + |\vec{p}_{\text{miss}}^*|$ is defined as:

$$E_{\text{miss}}^* + |\vec{p}_{\text{miss}}^*| \simeq |\vec{p}_{K^*}^*| + m_B - \sqrt{p_{K^*}^2 + m_{K^*}^2}. \quad (5.7)$$

Deriving Eq. 5.7 with respect to $p_{K^*}^*$, it shows a minimum for $p_{K^*}^* = 0$, whose value is $E_{\text{miss}}^* + |\vec{p}_{\text{miss}}^*| = m_B - m_{K^*} \simeq 4.4$ GeV. The maximum value can be inferred by noticing that for $|\vec{p}_{\text{miss}}^*| \to \infty$, $E_{\text{miss}}^* + |\vec{p}_{\text{miss}}^*|$ tends to be smaller than $m_B$. In conclusion the allowed physical range is $[4.4,5.3]$ GeV. In background events the sum $E_{\text{miss}}^* + p_{\text{miss}}^*$ is shifted at lower values with respect to the physical region since, in principle, no undetectable particles should be produced in these events:

- $\cos \theta_{\text{miss}}^*$ (Fig. 5.7): another very discriminating variable related to the $\nu\pi$ system is the angle between the direction of the missing three-momentum and the beam axis; the distribution of $\cos \theta_{\text{miss}}^*$ is almost flat for the signal, while it peaks at 1 for background.

Angular properties of the event (i.e. the angle between the signal $K^*$ and its $K$ daughter) have been investigated, but they have not shown any discriminating power. By applying selection criteria on $m_{K^*}^{PDG}$ and $m_{K^0_S}^{PDG}$, events in which the strange mesons are fake are rejected. This could be due to several effects: particle misidentification (a kaon used to reconstruct a $K^*$ is not a true $K$), a $B_{\text{had}}$ not correctly reconstructed (as a consequence some particles from the tag side are used to reconstruct the signal side composite particles), or a random combination of tracks and clusters that satisfies the $K^*$ or $K^0_S$ selection. The result is a broad distribution for the meson mass and background events are rejected by cutting on the tails. Some details about the $K^*$ mis-reconstruction for the signal MC sample are given in Sec. 6.1.3.
Reconstruction of the $B^{0(+)} \rightarrow K^{*(0(+))}\nu\overline{\nu}$ decay

Figure 5.2: $R_2$ distribution after the pre-selection requirements: $K^{*+} \rightarrow K^{+}\pi^0$ on the left, $K^{*+} \rightarrow K^0_s\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right; signal MC on the top, generic MC (histogram) and data (dots) on the bottom. Concerning the MC components: $\tau^+\tau^-$ in dark gray, $uds$ in green, $c\bar{c}$ in blue, $B^0\overline{B}^0$ in light gray, and $B^+B^-$ in red.
Figure 5.3: $E_{\text{extra}}$ distribution after the pre-selection requirements: $K^{*+} \rightarrow K^{+}\pi^{0}$ on the left, $K^{*+} \rightarrow K^{0}\pi^{+}$ on the center, and $K^{*0} \rightarrow K^{+}\pi^{-}$ on the right; signal MC on the top, generic MC (histogram) and data (dots) on the bottom. The same legend as in Fig. 5.2 is adopted.
Figure 5.4: $K^*$ mass distribution after the pre-selection requirements: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0_S\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right; signal MC on the top, generic MC (histogram) and data (dots) on the bottom. The same legend as in Fig. 5.2 is adopted.

Figure 5.5: $K_S^0$ mass distribution after the pre-selection requirements for $K^{*+} \rightarrow K_S^0\pi^+$; signal MC on the left, generic MC (histogram) and data (dots) on the right. The same legend as in Fig. 5.2 is adopted.
5.4 Selection variables

Figure 5.6: $E_{\text{miss}}^* + p_{\text{miss}}^*$ distribution after the pre-selection requirements: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0_S\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right; signal MC on the top, generic MC (histogram) and data (dots) on the bottom. The same legend as in Fig. 5.2 is adopted.
Figure 5.7: \( \cos \theta_{\text{miss}} \) distribution after the pre-selection requirements: \( K^{*+} \rightarrow K^+\pi^0 \) on the left, \( K^{*+} \rightarrow K^0\pi^+ \) on the center, and \( K^{*0} \rightarrow K^+\pi^- \) on the right; signal MC on the top, generic MC (histogram) and data (dots) on the bottom. The same legend as in Fig. 5.2 is adopted.
5.4 Selection variables

5.4.2 Model independence

In Sec. 1.3 it has been discussed how the kinematics of the event can be described by means of the neutrino pair invariant mass defined as [23]:

\[ s = \frac{m_B^2 + m_{K^*}^2 - 2m_B E_{K^*}}{m_B^2}, \]  

(5.8)

where \( m_B \) is the \( B \) meson mass and \( m_{K^*} \) and \( E_{K^*} \) are the \( K^* \) candidate mass and energy, respectively. In Fig. 5.8 the \( s \) distribution obtained by using the MC truth information and the SM prediction as computed in Ref. [23] are plotted.

To extract the \( B \rightarrow K^*\nu\bar{\nu} \) branching fraction assuming a given physical model, the signal MC sample needs to be re-weighted. For a Standard Model-dependent results, the MC correction could be done by using the results of Ref. [23]. In order to obtain the agreement between the experimental and the SM distribution, two procedures are feasible:

- **Re-weighting**: the ratio between the \( s \) MC-truth value and the theoretical expectation defines a weight for each event,

- **Hit and Miss**: here the ratio of the two distributions determines a probability with which each event is accepted or rejected.

Once this is done, all the kinematic variables, such as \( K^* \) and \( K^* \) daughter momenta, are correctly (SM-) modeled and can be used in the event selection.

On the other hand, to extract a completely model-independent result the phase-space correction should not be applied but the \( s \)-related variables can not be involved in the selection. In this thesis a model-independent analysis will be described, and the effects of the signal MC correction will be taken into account as systematic effects on the signal efficiency (Sec 7.2.1).

To ensure the model-independence of the selection variables, it has been checked that they are not correlated to \( s \) (some examples are given in Fig. 5.9). Two signal side variables \( (p_{K^*}^* \) and \( E_{\text{miss}}^* \)) correlated to \( s \) and not used in the selection are shown in Fig. 5.10.
Reconstruction of the $B^{0(+)} \rightarrow K^{*0(+)} \nu \overline{\nu}$ decay

Figure 5.8: Distribution of the kinematical variable $s$, signal MC truth (dots) and SM prediction [23] (line).

Figure 5.9: Correlation between $s$ and the sum of missing energy and missing momentum in the CM frame (left) and between $s$ and the $K^*$ mass (right) for the $K^{*0} \rightarrow K^+ \pi^-$ channel.

Figure 5.10: Correlation between $s$ and the MC truth-$K^*$ momentum in the CM frame (left) and between $s$ and the missing energy in the same frame (right) for the $K^{*0} \rightarrow K^+ \pi^-$ channel.
Chapter 6

Signal yield extraction

For the branching fraction extraction, three ingredients are necessary: the normalization given by the number of correctly tagged events in data, the signal efficiency evaluated in the signal MC sample, and an estimation for the number of signal events in data. The last two quantities are computed by using different approaches: a cut and count analysis in which the expected background is estimated and then compared to the data sample, and a fit to the output of a Neural Network that directly estimates the signal yield with its uncertainty. The two strategies are described in detail in this chapter: the Neural Network fit turns out to be the most powerful method in terms of estimated Upper Limit (as will be discussed in Sec. 8.2) since it takes advantage of the full shapes of the selection variable distribution.

6.1 Cut and count analysis

In this section the details of the cut and count analysis are presented. A multivariate algorithm is used to optimize the selection criteria on the variables listed in the previous chapters and the effectiveness of the selection is checked with respect to Monte Carlo samples. Studies on the MC truth information to identify the sources of irreducible background are also presented. An estimation of the background yield, as expected on data, is performed. The final result of this approach is a comparison between the expected background with the final data sample: the two are used to estimate the Upper Limit, as described in Sec. 8.1.
6.1.1 Optimization procedure and selection results

The selection requirements are optimized separately for each $K^*$ decay mode, and the variables used are:

1. $R_2$
2. $\cos \theta^*_{B,T}$
3. $K^*$ mass
4. $K^0_S$ mass (if applicable)
5. $\cos \theta^*_{\text{miss}}$

$E^\text{extra}_{\text{miss}}$ and $E^*_\text{miss} + p^*_\text{miss}$ are not involved in the optimization, but used at the end to define a signal region.

In the optimization algorithm the criteria are required to maximize the *Punzi* figure of merit (FOM) [46], defined as:

$$P_{\text{fom}} = \frac{\varepsilon_{\text{sig}}}{\frac{n_\sigma}{2} + \sqrt{N_{\text{bkg}}}},$$  \hspace{1cm} (6.1)$$

$\varepsilon_{\text{sig}}$ being the signal efficiency, computed as the ratio between the number of generated signal MC events and the amount of signal MC surviving the cut that is in process of being optimized; $n_\sigma$ is the number of standard deviations corresponding to the desired significance ($n_\sigma=1.285$ to set an Upper Limit at 90% confidence level); $N_{\text{bkg}}$ is the number of background events passing the selection. The advantage of using $P_{\text{fom}}$ (with respect to $S/\sqrt{B+S}$ for example) is its independence from the unknown branching fraction of the signal process.

The optimization algorithm used is the following [47]: rectangular regions in the selection variables are defined by optimizing a chosen figure of merit (*Punzi* in this case). To find the signal box with the overall optimal FOM, all possible binary splits in all dimensions are tested. The sample is split in three subsamples: *training*, *validate*, and *test* that correspond to one half, one quarter and one quarter of the whole sample, respectively. The training sample is used for the first iteration that consists of two steps: the *shrinking* and the *expansion*. The first one attempts to shrink the signal box in each dimension to optimize the FOM; the most important parameter at this stage is the *peel*
6.1 Cut and count analysis

parameter, defined as the fraction of events that can be peeled off the signal box in each iteration. In the expansion stage, once the signal box has been defined, the bounds of the box are relaxed in order to increase the FOM. After that, the validate sample is used: events outside the signal box are removed and a second iteration of the algorithm is run, to determinate the final cuts.

The optimization algorithm has been run testing several peel parameters an the one corresponding to the maximum FOM in the training sample has been chosen. The selection has been finally applied to the test sample (one quarter of the initial dataset). On the cut flow and efficiency tables only the test samples have been considered and to obtain the expected yields on data the data-MC scale factors of Tab. 3.2 have been multiplied by 4.

A summary of the selection criteria for each channel is reported in Tab. 6.1. In Fig. 6.1 and 6.2 the distributions of $E_{\text{extra}}$ and $E_{\text{miss}}^* + p_{\text{miss}}^*$ after the cuts are plotted. Looking at them the following signal box has been defined:

\[
E_{\text{miss}}^* + p_{\text{miss}}^* > 4.5 \text{ GeV},
\]
\[
0 \leq E_{\text{extra}} < 0.3 \text{ GeV}.
\] (6.2)

Another strategy for the yield estimation would have been applying a loose selection, in order to have enough statistics, and perform a 2-dimensional fit in $E_{\text{extra}}$ and $E_{\text{miss}}^* + p_{\text{miss}}^*$; Fig. 6.3 shows that the two variables are highly correlated and it would be difficult to take into account for it in a 2-dimensional fit.

A collection of cut flow tables (Tabs. 6.2-6.4) summarizing the event selection results on the data and MC samples is presented. The last two columns quote the number of observed events ($N_{\text{obs}}$) and the total number of expected background events ($N_{\text{bkg}}^\text{exp}$), whose uncertainty is computed as the quadrature sum of the event counting uncertainties from single MC contributions.

The signal MC sample is used to estimate the signal reconstruction efficiency. This is defined as the number of events surviving the selection, normalized to the number of correctly reconstructed $B_{\text{had}}$ in the simulated signal sample (Sec. 4.3). The values of $\epsilon_{B_{\text{sig}}}$ for each $K^*$ mode are listed in Tab. 6.5. Note that for the $K^{*0} \rightarrow K^+\pi^-$ mode the ratio $N_{\text{sig}}^{K^{*+}\nu\nu\text{MC}} / N_{B_{\text{had}}}^{K^{*+}\nu\nu\text{MC}}$ should be multiplied for $B(K^{*0} \rightarrow K^+\pi^-) = 66.57\%$ to account for the fact that in the signal simulation the $K^{*0}$ is forced to decay to $K^+\pi^-$. 
Table 6.1: Optimized cuts for each $K^*$ decay mode.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^{*+} \rightarrow K^+\pi^0$</td>
<td>$0.03 &lt; R_2 &lt; 0.70$</td>
</tr>
<tr>
<td></td>
<td>$0.04 &lt;</td>
</tr>
<tr>
<td></td>
<td>$0.84 &lt; m_{K^*} &lt; 0.95 \text{ GeV}/c^2$</td>
</tr>
<tr>
<td></td>
<td>$-0.78 &lt; \cos\theta^{miss}_{*} &lt; 0.93$</td>
</tr>
<tr>
<td>$K^{*+} \rightarrow K^0\pi^+$</td>
<td>$0.0 &lt; R_2 &lt; 0.49$</td>
</tr>
<tr>
<td></td>
<td>$0.0 &lt;</td>
</tr>
<tr>
<td></td>
<td>$0.86 &lt; m_{K^*} &lt; 0.95 \text{ GeV}/c^2$</td>
</tr>
<tr>
<td></td>
<td>$0.49 &lt; m_{K^0} &lt; 0.50 \text{ GeV}/c^2$</td>
</tr>
<tr>
<td></td>
<td>$-0.82 &lt; \cos\theta^{miss}_{*} &lt; 0.82$</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^+\pi^-$</td>
<td>$0.06 &lt; R_2 &lt; 0.53$</td>
</tr>
<tr>
<td></td>
<td>$0.002 &lt;</td>
</tr>
<tr>
<td></td>
<td>$0.85 &lt; m_{K^*} &lt; 0.97 \text{ GeV}/c^2$</td>
</tr>
<tr>
<td></td>
<td>$-0.86 &lt; \cos\theta^{miss}_{*} &lt; 0.90$</td>
</tr>
</tbody>
</table>

Figure 6.1: $E_{miss}^* + p_{miss}^*$ distribution after the selection cuts (Tab. 6.1): $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right; signal MC on the top, generic MC (histogram) and data (dots) on the bottom. Concerning the MC components: $\pi^+\pi^-$ in dark gray, $uds$ in green, $c\bar{c}$ in blue, $B^0\bar{B^0}$ in light gray, and $B^+B^-$ in red.
6.1 Cut and count analysis

Figure 6.2: $E_{\text{extra}}$ distribution after the selection cuts (Tab. 6.1): $K^{*+} \rightarrow K^{+}\pi^0$ on the left, $K^{*+} \rightarrow K^{0}\pi^+$ on the center, and $K^{*0} \rightarrow K^{+}\pi^-$ on the right; signal on the top, generic MC (histogram) and data (dots) on the bottom. The same legend as in Fig. 6.1 is adopted.

Figure 6.3: Correlation between $E_{\text{extra}}$ and $E_{\text{miss}}^* + p_{\text{miss}}^*$: $K^{*0} \rightarrow K^{+}\pi^-$ on the left, $K^{*+} \rightarrow K^{+}\pi^0$ on the center, and $K^{*+} \rightarrow K_0^{*}\pi^+$ on the right.
The events surviving the selection in the data sample are found to be 20 for $K^{*+} \rightarrow K^+\pi^0$, 11 for $K^{*+} \rightarrow K^{0}_S\pi^+$, and 19 for $K^{*0} \rightarrow K^+\pi^-$. They are in agreement with the background expectation, as can be noticed by comparing the last two columns of Tabs. 6.2-6.4, and no evidence for signal can be claimed.

Table 6.2: Cut flow table for $K^{*+} \rightarrow K^+\pi^0$ decay mode.

<table>
<thead>
<tr>
<th></th>
<th>Signal MC</th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N^\text{exp}_{\text{bkg}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>1.144</td>
<td>243.6</td>
<td>31.2</td>
<td>310.6</td>
<td>322.8</td>
<td>0</td>
<td>908 ± 24</td>
<td>795</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta^*_{B,T}</td>
<td>$</td>
<td>0.941</td>
<td>193.0</td>
<td>23.7</td>
<td>70.1</td>
<td>67.5</td>
<td>0</td>
</tr>
<tr>
<td>$m_{K^*}$</td>
<td>0.865</td>
<td>157.2</td>
<td>19.1</td>
<td>57.8</td>
<td>51.1</td>
<td>0</td>
<td>285 ± 12</td>
<td>272</td>
</tr>
<tr>
<td>$\cos \theta^\text{miss}_*$</td>
<td>0.843</td>
<td>128.7</td>
<td>16.9</td>
<td>45.9</td>
<td>40.5</td>
<td>0</td>
<td>232 ± 10</td>
<td>227</td>
</tr>
<tr>
<td>$E^<em>_{\text{miss}} + p^</em>_{\text{miss}}$</td>
<td>0.755</td>
<td>25.8</td>
<td>1.6</td>
<td>4.4</td>
<td>0</td>
<td>0</td>
<td>32 ± 3</td>
<td>44</td>
</tr>
<tr>
<td>$E_{\text{extra}}$</td>
<td>0.597</td>
<td>11.2</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>13 ± 2</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 6.3: Cut flow table for $K^{*+} \rightarrow K^{0}_S\pi^+$ decay mode.

<table>
<thead>
<tr>
<th></th>
<th>Signal MC</th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N^\text{exp}_{\text{bkg}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>1.087</td>
<td>1,005.5</td>
<td>216.5</td>
<td>754.1</td>
<td>753.6</td>
<td>0</td>
<td>2,730 ± 39</td>
<td>2837</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta^*_{B,T}</td>
<td>$</td>
<td>0.977</td>
<td>864.8</td>
<td>180.5</td>
<td>315.6</td>
<td>277.6</td>
<td>0</td>
</tr>
<tr>
<td>$m_{K^*}$</td>
<td>0.785</td>
<td>577.9</td>
<td>117.2</td>
<td>210.4</td>
<td>171.5</td>
<td>0</td>
<td>1,077 ± 22</td>
<td>1219</td>
</tr>
<tr>
<td>$m_{K^{0}_S}$</td>
<td>0.633</td>
<td>198.6</td>
<td>35.4</td>
<td>68.6</td>
<td>47.2</td>
<td>0</td>
<td>350 ± 12</td>
<td>409</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta^\text{miss}_*$</td>
<td>0.600</td>
<td>149.6</td>
<td>28.9</td>
<td>44.4</td>
<td>27.9</td>
<td>0</td>
<td>251 ± 10</td>
</tr>
<tr>
<td>$E^<em>_{\text{miss}} + p^</em>_{\text{miss}}$</td>
<td>0.548</td>
<td>19.5</td>
<td>2.6</td>
<td>3.5</td>
<td>2.8</td>
<td>0</td>
<td>28 ± 3</td>
<td>27</td>
</tr>
<tr>
<td>$E_{\text{extra}}$</td>
<td>0.438</td>
<td>6.6</td>
<td>0.3</td>
<td>1.0</td>
<td>0</td>
<td>0</td>
<td>8 ± 2</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 6.4: Cut flow table for $K^{*0} \rightarrow K^+\pi^-$ decay mode.

<table>
<thead>
<tr>
<th></th>
<th>Signal MC</th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N^\text{exp}_{\text{bkg}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>1.880</td>
<td>80.8</td>
<td>340.5</td>
<td>324.0</td>
<td>163.8</td>
<td>0</td>
<td>909 ± 21</td>
<td>856</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta^*_{B,T}</td>
<td>$</td>
<td>1.659</td>
<td>61.9</td>
<td>286.6</td>
<td>88.4</td>
<td>48.2</td>
<td>0</td>
</tr>
<tr>
<td>$m_{K^*}$</td>
<td>1.537</td>
<td>48.0</td>
<td>236.6</td>
<td>70.6</td>
<td>39.5</td>
<td>0</td>
<td>394 ± 13</td>
<td>415</td>
</tr>
<tr>
<td>$\cos \theta^\text{miss}_*$</td>
<td>1.501</td>
<td>41.0</td>
<td>194.4</td>
<td>46.4</td>
<td>23.1</td>
<td>0</td>
<td>305 ± 18</td>
<td>335</td>
</tr>
<tr>
<td>$E^<em>_{\text{miss}} + p^</em>_{\text{miss}}$</td>
<td>1.328</td>
<td>4.6</td>
<td>35.0</td>
<td>1.5</td>
<td>1.9</td>
<td>0</td>
<td>43 ± 4</td>
<td>37</td>
</tr>
<tr>
<td>$E_{\text{extra}}$</td>
<td>1.026</td>
<td>1.3</td>
<td>18.5</td>
<td>1.0</td>
<td>1.0</td>
<td>0</td>
<td>22 ± 3</td>
<td>19</td>
</tr>
</tbody>
</table>
6.1 Cut and count analysis

Table 6.5: Cut and count analysis: value of the signal selection efficiency and its statistical error for each $K^*$ mode, along with the number of signal events surviving the selection ($N_{B_{\text{had}}}^{K^*+ \nu \pi, \text{MC}}$) and the number of correctly reconstructed $B_{\text{had}}$ in the signal MC sample ($N_{B_{\text{had}}}^{K^*+ \nu \pi, \text{MC}}$), as computed in Sec. 4.3.

<table>
<thead>
<tr>
<th>$K^* \rightarrow K^+ \pi^0$</th>
<th>$K^* \rightarrow K^0 S \pi^+$</th>
<th>$K^0 \rightarrow K^+ \pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{sig}}^{K^*+ \nu \pi, \text{MC}}$</td>
<td>784</td>
<td>576</td>
</tr>
<tr>
<td>$N_{B_{\text{had}}}^{K^*+ \nu \pi, \text{MC}}$</td>
<td>19,373</td>
<td>19,373</td>
</tr>
<tr>
<td>$\epsilon_{B_{\text{sig}}}$</td>
<td>$(4.0 \pm 0.1) \times 10^{-2}$</td>
<td>$(3.0 \pm 0.1) \times 10^{-2}$</td>
</tr>
</tbody>
</table>

6.1.2 Background estimation

The expected number of signal and background events in data could be extracted by weighting the MC events that pass the selection, using the factors listed in Tab. 3.2.

For the Upper Limit computation, an estimation of the background yield in the $E_{\text{extra}}$ signal region is needed. In principle, once the data-MC agreement has been tested, one could consider the $N_{\text{bkg}}^{\text{exp}}$ value after the selection (Tabs. 6.2-6.4) as the expected background in data.

An alternative strategy which takes advantage of the information that can be extracted from the data sample is used. Counting the events in the $E_{\text{extra}}$ MC sideband ($E_{\text{extra}} > 0.6$) and signal region ($E_{\text{extra}} < 0.3$), the following ratio is defined:

$$r_{\text{MC}} = \frac{N_{\text{bkg}}^{\text{exp}}(E_{\text{extra}} < 0.3 \text{ GeV})}{N_{\text{bkg}}^{\text{exp}}(E_{\text{extra}} > 0.6 \text{ GeV})},$$

where the numerator (denominator) refers to the expected background yield in the $E_{\text{extra}}$ signal region (sideband). Trusting the MC-data agreement, this equality follows:

$$r_{\text{data}} = \frac{N_{\text{bkg}}(E_{\text{extra}} < 0.3 \text{ GeV})}{N_{\text{bkg}}(E_{\text{extra}} > 0.6 \text{ GeV})} = r_{\text{MC}},$$

in which the numerator (denominator) refers to the observed background yield in data in the $E_{\text{extra}}$ signal region (sideband). As a result, the background estimation in the $E_{\text{extra}}$ signal region for the data sample is:

$$N_{\text{bkg}}(E_{\text{extra}} < 0.3 \text{ GeV}) = N_{\text{bkg}}(E_{\text{extra}} > 0.6 \text{ GeV}) \cdot \frac{N_{\text{bkg}}^{\text{exp}}(E_{\text{extra}} < 0.3 \text{ GeV})}{N_{\text{bkg}}^{\text{exp}}(E_{\text{extra}} > 0.6 \text{ GeV})} .$$

In Tab. 6.6 the values of the factors entering Eq. 6.5 and the final yield estimation, split by $K^*$ decay mode, are shown. Details on the $E_{\text{extra}}$ sideband studies are discussed
Signal yield extraction

in Sec. 7.1.1. A systematic uncertainty resulting from the validity of Eq. 6.5 has been incorporated in the final result (Sec. 7.2.3).

The background yields, as computed in this section, are used in the UL evaluation (Sec. 8.3).

Table 6.6: Final estimation of the background yield in data as defined in Eq. 6.5, for each decay mode; the errors are computed by propagating the uncertainties on \(B(E_{\text{extra}} > 0.6 \text{ GeV})_{\text{data}}, B(E_{\text{extra}} > 0.6 \text{ GeV})_{\text{MC}}, B(E_{\text{extra}} < 0.3 \text{ GeV})_{\text{MC}}\).

<table>
<thead>
<tr>
<th>(K^+) Decay Mode</th>
<th>(E_{\text{extra}} &gt; 0.6 \text{ GeV})</th>
<th>(E_{\text{extra}} &lt; 0.3 \text{ GeV})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N_{\text{bkg}}^{\text{exp}})</td>
<td>(N_{\text{bkg}})</td>
</tr>
<tr>
<td>(K^{*+} \rightarrow K^+\pi^0)</td>
<td>13</td>
<td>8.8 ± 3.0</td>
</tr>
<tr>
<td>(K^{*+} \rightarrow K^0_S\pi^+)</td>
<td>9</td>
<td>8.4 ± 2.9</td>
</tr>
<tr>
<td>(K^{*0} \rightarrow K^+\pi^-)</td>
<td>7</td>
<td>10.9 ± 3.3</td>
</tr>
</tbody>
</table>

6.1.3 Background characterization

Studies on the MC truth information have been performed in order to identify the general features of \(B\bar{B}\) MC events surviving the selection.

In Tab. 6.7 the decays contributing to the background sample are listed. For the three channels, the main contamination arises from semileptonic decays, in which the lepton and some hadrons coming both from the \(B_{\text{had}}\) and the \(D^{(*)}\) on the signal side have not been reconstructed. A study of the kinematic properties of the lepton coming from the tag side shows that in most of the cases this particle is outside the detector acceptance (Fig. 6.4), as a consequence such events can not be rejected by applying veto on the lepton identification. Few events belong to the hadronic \(B\) decay category: in these cases either a true \(K^*\) is produced and most of the particles are missing, or there is a low multiplicity final state in which a pion and a kaon are randomly paired.

The MC truth information of the signal sample has also been investigated in order to check the goodness of the \(K^*\) reconstruction. An event is considered as matched if there is a true \(K^*\) coming from a \(B\), the \(K^*\) daughters are truth-matched, and are produced from the same \(K^*\). In the \(K^{*+} \rightarrow K^+\pi^0\) mode, less than 6% of the events are not truth matched, this happen when a \(\pi^0\) is not matched or the \(\pi^0-K\) pair does not originate from the same parent. For \(K^{*+} \rightarrow K^0_S\pi^+\) decay, most of the unmatched events are due to fake \(K^0_S\). For \(K^{*0} \rightarrow K^+\pi^-\), in 6% of the cases \(K\) and \(\pi\) are paired randomly or come from
6.2 Neural Network analysis

another $K^{*0}$ in the event.

Figure 6.4: Module of the cosine of the angle between the $z$-axis and the generated lepton momentum in the laboratory frame for $B \rightarrow D^{(*)}\ell\nu$ events: $K^{*+} \rightarrow K^{+}\pi^{0}$ on the left, $K^{*+} \rightarrow K^{0}\pi^{+}$ on the center, and $K^{*0} \rightarrow K^{+}\pi^{-}$ on the right.

Table 6.7: MC truth studies for $B\overline{B}$ MC events surviving the event selection defined in Tab. 6.1.

<table>
<thead>
<tr>
<th></th>
<th>$K^{*+} \rightarrow K^{+}\pi^{0}$</th>
<th>$K^{*+} \rightarrow K^{0}\pi^{+}$</th>
<th>$K^{*0} \rightarrow K^{+}\pi^{-}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B \rightarrow D^{(*)} e \nu_{e}$ (n $\gamma$)</td>
<td>13</td>
<td>7</td>
<td>30</td>
</tr>
<tr>
<td>$B \rightarrow D^{(*)} \mu \nu_{\mu}$ (n $\gamma$)</td>
<td>2</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td>$B \rightarrow \tau \nu_{\tau}$ (n $\gamma$)</td>
<td>12</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>hadronic $B$ decays</td>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

6.2 Neural Network analysis

The discriminating power of the selection variable distributions is exploited by combining them in a Neural Network. This is a computational model whose functioning is inspired by the operational principles of biological neural systems, allowing to perform tasks that can not be faced by linear algorithms. The features of the Neural Network will be discussed in Sec. 6.2.1. For each event, given a set of variables, their values are combined to compute the Neural Network output ($NN_{out}$).

In this analysis $NN_{out}$ distributions are fitted to extract the signal and background yields ($N_{sig}$ and $N_{bkg}$, respectively). The selection variables are first subjected to loose requirements that allow to maintain a larger signal statistics with respect to the cut and count approach, the higher background contamination being compensated by the better discriminating power of the Neural Network fit. After the Neural Network has been run, only events whose $NN_{out}$ values lie in a defined signal range are retained for the
Signal yield extraction

The number of signal MC events surviving the selection allows to compute the signal efficiency. The MC samples are also used to parametrize the signal and background distributions of the fit variable. The resulting parameters are then fixed in the fit to the real data, in which only $N_{\text{sig}}$ and $N_{\text{bkg}}$ are allowed to float. The signal yield and efficiency, with their statistical and systematic errors, are used to estimate the branching fraction Upper Limits (Sec. 8.3). As it will be shown in Chap. 8, the Neural Network approach results to be much more performing with respect to the cut and count.

6.2.1 Neural Network philosophy

Artificial Neural Networks are non-linear computational models based on the human brain functioning. The cerebral neural system is formed by a huge number of interconnected neurons that communicate by means of electric signals. The features of the human brain, that Neural Network algorithms try to reproduce, are the high versatility of the neurons organization that allows to face a great variety of problems and the ability on learning from the experience. The artificial Neural Network consists of artificial neurons called perceptrons. In this analysis a network called Multi-Layer Perceptron is used: it is implemented in the TMultiLayerPerceptron ROOT class [48] and comprises sets of perceptrons organized in several layers. A sketch of the Neural Network structure is shown in Fig. 6.5.

![Neural Network structure](image)

Figure 6.5: Sketch of the Neural Network structure used in this analysis.

There are three types of neuron layers: the input layer, the hidden layer, and the output layer. The neurons of the first layer receive the inputs $x_i$ which consist of real numbers (i.e. the value of the selection variables). The $x_i$’s are then forwarded to the hidden layer(s). Here the input variables are linearly combined by means of weight variables
(w_{ij}) that simulate the biological neural connections:

\[ x_j = w_{0j} + \sum_i w_{ij} x_i. \] (6.6)

The \( x_j \) are used to compute the input of the next layer (hidden layer, if more than one is present, or the output layer) as:

\[ u_j = A(x_i) = A\left(w_{0j} + \sum_i w_{ij} x_i\right), \] (6.7)

where \( A(x) \) is the so-called sigmoid function, defined as \( A(x) = 1/(1 + e^{-x}) \), which represents a “smooth” threshold that controls the activation or the inhibition of a perceptron.

The final \( NN_{\text{out}} \) is given by the linear combination of the \( y_k \), computed as:

\[ y_k = w_{0k} + \sum_j w_{jk} u_j. \] (6.8)

The weights are initially set to random numbers and then updated during the learning process, whose main purpose is the optimization of the weight matrix. Usually the available sample is split in two: one subset (training sample) is used in the learning step, then the weights are fixed and the Neural Network is run on a second subset (test sample), which is used for the actual analysis. The weight optimization can be done by comparing the Neural Network output to an expected value (supervised learning) or by exploiting some dedicated input information (non-supervised learning). The supervised learning algorithms are the most used: given a set of examples \( p \), the Neural Network outputs \( NN_{\text{out}}^p \) are compared to the expected values \( NN_{\text{out}}^{p,\text{exp}} \). The weights are updated by minimizing the quantity:

\[ E = \sum_p \frac{1}{2} (NN_{\text{out}}^p - NN_{\text{out}}^{p,\text{exp}})^2. \] (6.9)

Several training algorithms can be adopted depending on the mathematical formulation used to update the weights. In this analysis a method with line search [49] is exploited and consists of the following steps:

1. define the function to be minimized as \( E(\vec{w}) \),

2. compute a direction \( \vec{s} \) from the gradient \( \nabla E \),

3. find a parameter \( \alpha_{\text{min}} \) which minimizes \( E(\vec{w} + \alpha \vec{s}) \) (this part of the algorithm is called line search),
4. update the weight matrix to the values $\bar{w} + \alpha_{\min}\bar{s}$

Several parameters of the Neural Network need to be set by the user:

- **number of hidden layers and number of hidden neurons**, chosen according to the complexity of the problem;

- **number of epochs** which is the number of times the learning algorithm is iterated: it needs to be high enough to allow an accurate weight determination without over-training the Neural Network (i.e. the parameters are computed ad-hoc for the training sample but do not work properly for the test sample);

- parameters related to the training algorithm; for the one in use here, two parameters are needed: $\tau$ which defines the precision of the determination of the line search and $\text{reset}$ that is the number of epochs between two resets of the optimal search direction.

In the following section details about the Neural Network architecture and parameters chosen in this analysis are discussed.

### 6.2.2 Pre-selection and Neural Network optimization

Three different Neural Networks are trained, one for each $K^*$ decay mode. The $K^+\pi^-$ and $K^+\pi^0$ Neural Networks have 6 input variables: $R_2$, $m_{K^*}$, $\cos\theta^*_\text{miss}$, $|\cos\theta^*_{B,T}|$, $E_{\text{extra}}$, and $E^*_\text{miss} + p^*_\text{miss}$; the $K^0_{S}\pi^+$ Neural Network inputs include also $m_{K^0_{S}}$. To suppress the background contamination and maximize the signal efficiency a raw pre-selection on the Neural Network inputs is applied:

- $R_2 < 0.7$

- $|\cos\theta^*_{B,T}| < 0.95$

- $0.84 < m_{K^*} < 0.96 \text{ GeV}/c^2$

- $0.490 < m_{K_{S}^0} < 0.509 \text{ GeV}/c^2$ (for the $K^{*+} \to K^0_{S}\pi^+$ channel)

- $|\cos\theta^*_{\text{miss}}| < 0.9$
Each Neural Network has one hidden layer and the number of hidden neurons are twice the inputs. The training algorithm parameters are tuned in order to have the smallest error on the weights that define the connections between neurons, to obtain a high $S/\sqrt{B}$, and not to over-train the network. The Neural Network optimization and test is performed on MC signal and background samples. Only half of the whole statistics is used to find the set of optimal weights. The architectures of the three trained Neural Networks are shown in Fig. 6.6. In Fig. 6.7 the shift on the Neural Network output for a small variation of each input is plotted: for all channels the most discriminating variables are $E_{\text{extra}}$ and $E^*_{\text{miss}} + p^*_{\text{miss}}$, as expected.

Figure 6.6: Neural Network architecture: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right. The thickness of the line between two perceptrons indicates the strength of the connecting weights.

Figure 6.7: Shift on the Neural Network output for a small variation of each input variable: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right. In the $K^{*+} \rightarrow K^+\pi^0$ and $K^{*0} \rightarrow K^+\pi^-$ plots: $E_{\text{extra}}$ (black), $E^*_{\text{miss}} + p^*_{\text{miss}}$ (red), $m_{K^*}$ (green), $R_2$ (dark blue), $|\cos \theta^*_{B,T}|$ (yellow), $\cos \theta^*_{\text{miss}}$ (pink). In the $K^{*+} \rightarrow K^0\pi^+$ plot: $E_{\text{extra}}$ (black), $E^*_{\text{miss}} + p^*_{\text{miss}}$ (red), $m_{K^*}$ (green), $m_{K^0}$ (dark blue), $R_2$ (yellow), $|\cos \theta^*_{B,T}|$ (pink), $\cos \theta^*_{\text{miss}}$ (light blue).
6.2.3 Neural Network output fit

In Fig. 6.8 the $NN_{out}$ distributions are plotted: the histograms represent the events of the training sample, the dots are the test sample. The network is trained in order to have $NN_{out}$ equaling 0 for the background and 1 for the signal. For the $K^{*0} \rightarrow K^+\pi^-$ channel, the Neural Network performance is optimal: the signal $NN_{out}$ distribution has a narrow peak near 1 and just a small tail below 0.7; also for the background it has a peaking shape around 0. For the other two channels the signal $NN_{out}$ distribution has a broader shape and to include all the signal events the range has to be enlarged up to 1.1 for $K^{*+} \rightarrow K^+\pi^0$ and 1.2 for $K^{*+} \rightarrow K^0_S\pi^+$.  

In Fig. 6.9 the Neural Network output in the fit ranges are shown: the lower bound is 0.6 for all the channels while the upper bound changes according to the signal and background endpoints. The plots shows a similar behavior between test and training samples: this allows the use of the full MC sample in the fits in which the parametrization of the $NN_{out}$ distribution is determined. The fit region has been chosen in order to maximize the signal efficiency while keeping the background contamination under control.

The values of the signal reconstruction efficiency, computed as the number of events surviving the selection, normalized to the number of correctly reconstructed $B_{had}$ in the simulate signal sample, are shown in Tab. 6.8. As in the cut and count analysis, for the $K^{*0} \rightarrow K^+\pi^-$ mode the ratio $N_{sig}^{K^{*+} \nu\pi,MC}/N_{B_{had}}^{K^{*+} \nu\pi,MC}$ should be multiplied for $B(K^{*0} \rightarrow K^+\pi^-) = 66.57\%$ to account for the fact that in the signal simulation the $K^{*0}$ is forced to decay to $K^+ \pi^-$.  

<table>
<thead>
<tr>
<th>$K^{*+} \rightarrow K^+\pi^0$</th>
<th>$K^{*+} \rightarrow K^0_S\pi^+$</th>
<th>$K^{*0} \rightarrow K^+\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{sig}^{K^{*+} \nu\pi,MC}$</td>
<td>1,140</td>
<td>2,200</td>
</tr>
<tr>
<td>$N_{B_{had}}^{K^{*+} \nu\pi,MC}$</td>
<td>19,373</td>
<td>8,856</td>
</tr>
<tr>
<td>$\epsilon_{B_{sig}}$</td>
<td>$(5.8 \pm 0.2) \times 10^{-2}$</td>
<td>$(16.6 \pm 0.4) \times 10^{-2}$</td>
</tr>
</tbody>
</table>

The estimation of the signal and background yields are performed by an extended
Figure 6.8: $NN_{\text{out}}$ distribution: $K^{*+} \rightarrow K^{+}\pi^{0}$ on the left, $K^{*+} \rightarrow K_{S}^{0}\pi^{+}$ on the center, and $K^{*0} \rightarrow K^{+}\pi^{-}$ on the right; background on the top and signal on the bottom. The histogram represents the training sample, the bars the test sample.
Figure 6.9: $NN_{\text{out}}$ distributions in the region chosen for the fit: $K^{**} \rightarrow K^+\pi^0$ on the left, $K^{**} \rightarrow K^0_S\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right; background on the top and signal on the bottom. The histogram represents the training sample, the bars the test sample.
maximum likelihood fit to the $NN_{\text{out}}$ distribution. The likelihood is defined as:

$$
\mathcal{L}(N_{\text{sig}}, N_{\text{bkg}}) = \frac{e^{-(N_{\text{sig}} + N_{\text{bkg}})}}{(N_{\text{sig}} + N_{\text{bkg}})!} \times \prod_{i=1}^{N} \left[ f_{\text{sig}}(NN_{\text{out}}, i | \vec{p}_{\text{sig}}) N_{\text{sig}} + f_{\text{bkg}}(NN_{\text{out}}, i | \vec{p}_{\text{bkg}}) N_{\text{bkg}} \right].
$$

(6.10)

The functions $f_{\text{sig}}(NN_{\text{out}}, i | \vec{p}_{\text{sig}})$ and $f_{\text{bkg}}(NN_{\text{out}}, i | \vec{p}_{\text{bkg}})$ represent the PDFs for the $NN_{\text{out}}$ distribution for signal and background samples, respectively, and depend on the set of parameters $\vec{p}_{\text{sig}}$ and $\vec{p}_{\text{bkg}}$; they will be discussed in the next two sections. The fitted signal and background yields correspond to a maximum of the likelihood and if this has a Gaussian distribution in the plane $N_{\text{sig}} - N_{\text{bkg}}$ the fitted values are the central values along the two axes of the Gaussian and the widths correspond to the fitted statistical errors. The choice of the fit functions has been validated by toy Monte Carlo studies discussed at the end of this section; in addition, different parametrizations have been tested in order to check that the signal and background yields were compatible with the MC nominal fit result.

**Fit to background events**

For the three $K^*$ channels, the probability density function that parametrizes the background is a “modified” Fermi function, defined as:

$$
f(x) = \frac{x + k_1}{1 + e^{k_2 x}},
$$

(6.11)

depending on the two parameters $k_1$ and $k_2$ which are free in the fit. In Tab. 6.9 the expected background yield on data that survives the $NN_{\text{out}}$ cut is listed. To extract the background PDF parameters a sample that is three times the expected amount, maintaining the proportion between the different components, is used. Since the $c\bar{c}$ and $uds$ MC samples that survive the selection are not enough to triple the expected amount, the lacking events are taken from the $m_{\text{ES}}$ sideband. This is done after checking that the $NN_{\text{out}}$ distribution for continuum events does not depend on the $m_{\text{ES}}$ region. The comparison between continuum events in the $m_{\text{ES}}$ sideband and signal region is shown in Figs. 6.10 and 6.11: the Neural Network shape seems to agree very well between the two $m_{\text{ES}}$ regions, both for $c\bar{c}$ and $uds$ and for all the $K^*$ channels (this is not evident for the $uds$ sample surviving the $K^{*0} \rightarrow K^+\pi^-$ selection due to a lack of statistics in the $m_{\text{ES}}$ signal region). The values of the fitted parameters and the correlation coefficients defining the background $NN_{\text{out}}$ PDF are listed in Tab. 6.10 and the fits are plotted in Fig. 6.12.
Table 6.9: Expected background yield in data after the cut on $NN_{\text{out}}$.

<table>
<thead>
<tr>
<th>$K^*$ Decay Mode</th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$\bar{c}\bar{\tau}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N_{\text{bkg}}^\text{exp}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^{*+} \rightarrow K^+\pi^0$</td>
<td>31</td>
<td>2</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>$K^{*+} \rightarrow K^0\pi^+$</td>
<td>24</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^+\pi^-$</td>
<td>6</td>
<td>48</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>67</td>
</tr>
</tbody>
</table>

Figure 6.10: $NN_{\text{out}}$ distributions, normalized to the unity, for $c\bar{c}$ events in the $m_{ES}$ sideband (histogram) and signal region (bars): $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right. On the top the full Neural Network output region is shown while on the bottom the distributions for $NN_{\text{out}} > 0.1$ are plotted.
6.2 Neural Network analysis

Figure 6.11: $NN_{\text{out}}$ distributions, normalized to the unity, for $uds$ events in the $m_{\text{ES}}$ sideband (histogram) and signal region (bars): $K^{*+} \rightarrow K^{+}\pi^0$ on the left, $K^{*+} \rightarrow K^0_S\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right. On the top the full Neural Network output region is shown while on the bottom the distributions for $NN_{\text{out}} > 0.1$ are plotted.

Table 6.10: Values of the background PDF parameters and their correlations ($\rho$) for each $K^*$ mode.

<table>
<thead>
<tr>
<th>$K^{*\pm} \rightarrow K^{\mp}\pi^\mp$</th>
<th>Background PDF Params.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_1 = -0.5 \pm 0.1$</td>
</tr>
<tr>
<td></td>
<td>$k_2 = 7.8 \pm 1.3$</td>
</tr>
<tr>
<td></td>
<td>$\rho(k_1, k_2) = -0.838$</td>
</tr>
<tr>
<td>$K^{*+} \rightarrow K^0_S\pi^+$</td>
<td>$k_1 = -0.58 \pm 0.02$</td>
</tr>
<tr>
<td></td>
<td>$k_2 = 9.1 \pm 0.9$</td>
</tr>
<tr>
<td></td>
<td>$\rho(k_1, k_2) = -0.636$</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^+\pi^-$</td>
<td>$k_1 = -0.4 \pm 0.2$</td>
</tr>
<tr>
<td></td>
<td>$k_2 = 1.1 \pm 2.1$</td>
</tr>
<tr>
<td></td>
<td>$\rho(k_1, k_2) = -0.945$</td>
</tr>
</tbody>
</table>
Figure 6.12: Fit to the $NN_{out}$ distribution for the background: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0_S\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right.

**Fit to signal events**

To determine the signal shape, the full signal MC sample retained after the selection has been used; it consists of 1,140 events for $K^{*+} \rightarrow K^+\pi^0$, 1,017 for $K^{*+} \rightarrow K^0_S\pi^+$, and 2,200 for $K^{*0} \rightarrow K^+\pi^-$. To parametrize the $NN_{out}$ distribution, for the last channel an exponential function has been used while for the $K^{*+}$ samples a Crystal Ball [42] shape has been adopted. The definitions of the signal fit functions are:

$$K^{*0} \rightarrow K^+\pi^- : f(x) = e^{-c/x},$$

$$K^{*+} \rightarrow K^+\pi^0 \text{ and } K^{*+} \rightarrow K^0_S\pi^+ : f(x) = \begin{cases} 
\left(\frac{a}{|a|}\right)^n e^{-a^2/2} x < -|a| \\
\left(\frac{1}{|a|}\right)^n e^{-a^2/2} (x - \frac{a}{|a|})^2 x > -|a| 
\end{cases}$$

(6.12)

Tab. 6.11 summarizes the values of the parameters that define the signal PDFs; for the charged modes the parameters $a$, $m$, and $s$ are highly correlated while $n$ is not. In the nominal fit, $n$ has been therefore fixed to the value obtained when floating all parameters: this gives more stability to the fit and smaller statistical errors. The fits to the signal samples are shown in Fig. 6.13.

Figure 6.13: Fit to the $NN_{out}$ distribution for the signal MC: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K^0_S\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right.
6.2 Neural Network analysis

Table 6.11: Values of the signal PDF parameters and their correlations (\(\rho\)) for each \(K^*\) mode.

<table>
<thead>
<tr>
<th>(K^{*+} \rightarrow K^+\pi^0)</th>
<th>Signal PDF Params.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a = 0.32 \pm 0.04)</td>
<td>(\rho(a, m) = -0.897)</td>
</tr>
<tr>
<td>(m = 0.961 \pm 0.007)</td>
<td>(\rho(a, s) = 0.890)</td>
</tr>
<tr>
<td>(s = 0.050 \pm 0.004)</td>
<td>(\rho(m, s) = -0.901)</td>
</tr>
<tr>
<td>(n = 42)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(K^{*+} \rightarrow K_S^0\pi^+)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a = 0.28 \pm 0.05)</td>
<td>(\rho(a, m) = -0.929)</td>
</tr>
<tr>
<td>(m = 0.99 \pm 0.01)</td>
<td>(\rho(a, s) = 0.913)</td>
</tr>
<tr>
<td>(s = 0.060 \pm 0.007)</td>
<td>(\rho(m, s) = -0.942)</td>
</tr>
<tr>
<td>(n = 111)</td>
<td></td>
</tr>
</tbody>
</table>

| \(K^{*0} \rightarrow K^+\pi^-\) | \(c = 11.53 \pm 0.28\) |

**Fit to data-like sample**

The global fit (signal+background) has been tested on a “cocktail” MC sample. This is composed by a number of background and signal events equal to the expected yield on data. In Tab. 6.9 the amount of each background sources for each channel is listed. For the signal a branching fraction equal to the SM expectation (\(1.3 \times 10^{-5}\)) is assumed. For the three channels the expected number of signal events is:

- \(K^{*+} \rightarrow K^+\pi^0\): \(N_{\text{sig}}^{\text{exp,SM}} = 1\)
- \(K^{*+} \rightarrow K_S^0\pi^+\): \(N_{\text{sig}}^{\text{exp,SM}} = 1\)
- \(K^{*0} \rightarrow K^+\pi^-\): \(N_{\text{sig}}^{\text{exp,SM}} = 2\)

Fig. 6.14 shows the cocktail MC sample and the fit function; in Tab. 6.12 the fitted yields are listed.

Table 6.12: Results of the fit on the signal+background MC cocktail for each \(K^*\) mode.

<table>
<thead>
<tr>
<th>(K^{*+} \rightarrow K^+\pi^0)</th>
<th>Yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_{\text{sig}} = -3 \pm 5)</td>
<td>(N_{\text{bkg}} = 46 \pm 9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(K^{*+} \rightarrow K_S^0\pi^+)</th>
<th>Yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_{\text{sig}} = 1 \pm 6)</td>
<td>(N_{\text{bkg}} = 35 \pm 8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(K^{*0} \rightarrow K^+\pi^-)</th>
<th>Yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_{\text{sig}} = -4 \pm 9)</td>
<td>(N_{\text{bkg}} = 73 \pm 12)</td>
</tr>
</tbody>
</table>
Figure 6.14: Fit to the $N N_{\text{out}}$ distribution for a signal+background cocktail MC as expected on data: $K^*+ \rightarrow K^+\pi^0$ on the left, $K^*+ \rightarrow K_S^0\pi^+$ on the center, and $K^*0 \rightarrow K^+\pi^-$ on the right. The dots represent the MC events, the blue, red and green are the total, the background and the signal PDFs, respectively.

**Fit validation**

To validate the fit strategy toy MC studies have been performed. They consist of numerically generating samples according to the $N N_{\text{out}}$ probability density function for the signal and background samples. For each channel, 1000 toy experiments have been performed; in each of them, signal and background yields are generated according to a Poisson distribution around the values expected in the full data statistics. These samples are fitted with the PDFs whose shapes are fixed from the MC fits, and the signal and background yields are the only floating parameters. The expected signal and background yields with their expected errors are listed in Tab. 6.13: they are the mean of the fitted yield and fitted error distributions (left side and middle plots in Figs. 6.15-6.17). A Gaussian fit to the pulls of the $N_{\text{sig}}$ and $N_{\text{bkg}}$ distributions is used to check the absence of a potential fit bias: this is ensured by the mean of the pull being compatible with zero and the width with one (right side plots in Figs. 6.15-6.17). This also proves that the likelihood function has a Gaussian distribution centered at $N_{\text{sig}}$ and whose width corresponds to the statistical error returned by the fit on the same parameter.

An additional test has been done to evaluate the behavior of the pull on $N_{\text{sig}}$ for different BF values. Embedded fits have been performed: they consist on fitting a MC sample in which both the signal and the background components are taken in the expected proportions. To increase the statistics, three times of the expected yields is considered. The test is done by assuming different branching fractions, from the Standard Model value to a BF that is 25 times larger. In Fig. 6.18 the pull on $N_{\text{sig}}$ as a function of the BF for the three channels is shown: the points are distributed around 0, assuming both negative
and positive values and excluding a systematic underestimation or overestimation of $N_{\text{sig}}$.

Table 6.13: Expected signal and background yields with their expected errors as extracted from toy MC studies.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^{*+} \rightarrow K^+\pi^0$</td>
<td>$N_{\text{sig}}^{\text{exp,toy}} = 1 \pm 5$</td>
</tr>
<tr>
<td></td>
<td>$N_{\text{bkg}}^{\text{exp,toy}} = 43 \pm 8$</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^0\pi^+$</td>
<td>$N_{\text{sig}}^{\text{exp,toy}} = 1 \pm 5$</td>
</tr>
<tr>
<td></td>
<td>$N_{\text{bkg}}^{\text{exp,toy}} = 35 \pm 8$</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^+\pi^-$</td>
<td>$N_{\text{sig}}^{\text{exp,toy}} = 2 \pm 9$</td>
</tr>
<tr>
<td></td>
<td>$N_{\text{bkg}}^{\text{exp,toy}} = 67 \pm 12$</td>
</tr>
</tbody>
</table>

Figure 6.15: Results of toy MC experiments for $K^{*+} \rightarrow K^+\pi^0$: signal yield (top left), signal yield error (top middle), pull on the signal yield (top right), background yield (bottom left), background yield error (bottom middle), pull on the background yield (bottom right).

### 6.2.4 Fit to data

In this section, the results of the Neural Network fit in the data sample are presented. The events that survive the signal selection are 44 for $K^{*+} \rightarrow K^+\pi^0$, 54 for $K^{*+} \rightarrow K^0\pi^+$, and 67 for $K^{*0} \rightarrow K^+\pi^-$. In Tab. 6.14 the fitted yields are quoted and the data distributions with the PDFs superimposed are shown in Fig. 6.19.

As in the cut and count analysis, also the Neural Network fit does not show any evidence for signal.
Figure 6.16: Results of toy MC experiments for $K^{*+} \rightarrow K^0\pi^+$: signal yield (top left), signal yield error (top middle), pull on the signal yield (top right), background yield (bottom left), background yield error (bottom middle), pull on the background yield (bottom right).

Figure 6.17: Results of toy MC experiments for $K^{*0} \rightarrow K^+\pi^-$: signal yield (top left), signal yield error (top middle), pull on the signal yield (top right), background yield (bottom left), background yield error (bottom middle), pull on the background yield (bottom right).
6.2 Neural Network analysis

Figure 6.18: Pull on $N_{\text{sig}}$ as a function of the branching fraction assumed for the signal decay: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K_S^0\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right.

Table 6.14: Results of the $NN_{\text{out}}$ fit on data for each $K^*$ mode.

<table>
<thead>
<tr>
<th>Yields</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^{*+} \rightarrow K^+\pi^0$</td>
<td>$N_{\text{sig}} = 4.7 \pm 6.4$</td>
</tr>
<tr>
<td></td>
<td>$N_{\text{bkg}} = 39.3 \pm 8.7$</td>
</tr>
<tr>
<td>$K^{*+} \rightarrow K_S^0\pi^+$</td>
<td>$N_{\text{sig}} = 2.7 \pm 6.9$</td>
</tr>
<tr>
<td></td>
<td>$N_{\text{bkg}} = 51.3 \pm 9.7$</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^+\pi^-$</td>
<td>$N_{\text{sig}} = -10.3 \pm 9.2$</td>
</tr>
<tr>
<td></td>
<td>$N_{\text{bkg}} = 77.3 \pm 13.1$</td>
</tr>
</tbody>
</table>

Figure 6.19: Fit to the $NN_{\text{out}}$ distribution on data: $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K_S^0\pi^+$ on the center, and $K^{*0} \rightarrow K^+\pi^-$ on the right.
Signal yield extraction
Chapter 7

Cross-checks and systematic uncertainties

Monte Carlo samples are used to establish and validate the analysis procedure. In the cut and count approach, the simulated data allow to optimize the selection requirements and to estimate the expected background. In the Neural Network fit the MC samples are used to choose the best signal window in $NN_{\text{out}}$, extract the $NN_{\text{out}}$ PDF parameters, and validate the fit. Once the procedure is tested, it is frozen and applied to the real data: as a consequence is essential to check that the simulated sample well reproduces the real one. To this purpose, several tests using different control samples have been performed (Sec. 7.1).

The studies listed above also allow to compute some of the systematic uncertainties, whose detailed description is given in Sec 7.2.

7.1 Monte Carlo validation and control samples

The $E_{\text{extra}}$ sideband, a window defined away from the signal region, is used to compare the selection variable distributions in data and MC and the expected and observed number of events, once some selection criteria are applied (Sec. 7.1.1); similarly, a sample in which both the $B$’s decay to hadronic final states is exploited (Sec. 7.1.2). Dedicated studies have been performed for the Neural Network fit; the data and MC $NN_{\text{out}}$ distributions are compared in two control samples: the $m_{\text{ES}}$ sideband and a “wrong charge” sample, in which the charge correlation between the $B_{\text{sig}}$ and the $B_{\text{had}}$ is incorrect. The Neural Network output on Monte Carlo $B\bar{B}$ sample is validated by comparing generic $B\bar{B}$ MC and “cocktail” MC, which consists on generating a generic B decay and a hadronic final
Cross-checks and systematic uncertainties

state on the other side of the event.

7.1.1 \(E_{\text{extra}}\) sideband

Plots of relevant variables for the analysis \((m_{ES}, \Delta E, R_2, \cos \theta_{B^* T}, m_{K^*}, m_{K^0}, \cos \theta_{\text{miss}}^*, E^*_{\text{miss}} + p^*_{\text{miss}}, \text{and} E_{\text{extra}})\) are shown in Figs. 7.1-7.9 in order to quantify the data-MC agreement in the \(E_{\text{extra}}\) sideband, region defined as:

\[
E_{\text{extra}} > 0.6 \text{ GeV}
\] (7.1)

The data distribution of some of the variables used in the selection is compared to the sum of properly normalized MC samples, before any requirement is applied. In Tabs. 7.1-7.3 the number of MC and data events which pass each selection requirement is shown.

A good agreement is shown both in the distribution shapes and in the number of expected and observed events, above all after the selection requirements.

To account for systematics effects, the same study has been performed separately for each data taking period: also in this case a good matching between simulated and real samples has been found.

Figure 7.1: comparison between data (dots) and MC (histograms) in the \(E_{\text{extra}}\) sideband: \(m_{ES}\) for \(K^{*+} \to K^+ \pi^0\) (left), \(K^{*+} \to K_s^0 \pi^+\) (center), and \(K^{*0} \to K^+ \pi^-\) (right); the real data (dots) are superimposed to the normalized sum of the different MC contributions: \(\tau^+ \tau^-\) (dark gray), \(uds\) (green), \(c\bar{c}\) (blue), \(B^0 \bar{B}^0\) (light gray), and \(B^+ B^-\) (red).

7.1.2 Double hadronic tag sample

The Double hadronic tag control sample is composed of events where both the \(B\) mesons are reconstructed in one of the \(B \to D^{(*)} X\) listed in Sec. 4.1. This sample is exploited to test the data-MC agreement. All the MC samples \((B^+ B^-, B^0 \bar{B}^0, c\bar{c}, uds, \tau^+ \tau^-)\) and
7.1 Monte Carlo validation and control samples

Figure 7.2: Comparison between data (dots) and MC (histograms) in the $E_{\text{extra}}$ sideband: $\Delta E$ for $K^{*+} \rightarrow K^+ \pi^0$ (left), $K^{*+} \rightarrow K^0_S \pi^+$ (center), and $K^{*0} \rightarrow K^+ \pi^-$ (right). The same legend as in Fig. 7.1 is adopted.

Figure 7.3: Comparison between data (dots) and MC (histograms) in the $E_{\text{extra}}$ sideband: $R_2$ for $K^{*+} \rightarrow K^+ \pi^0$ (left), $K^{*+} \rightarrow K^0_S \pi^+$ (center), and $K^{*0} \rightarrow K^+ \pi^-$ (right). The same legend as in Fig. 7.1 is adopted.

Figure 7.4: Comparison between data (dots) and MC (histograms) in the $E_{\text{extra}}$ sideband: $\cos \theta_{B,T}$ for $K^{*+} \rightarrow K^+ \pi^0$ (left), $K^{*+} \rightarrow K^0_S \pi^+$ (center), and $K^{*0} \rightarrow K^+ \pi^-$ (right). The same legend as in Fig. 7.1 is adopted.
Cross-checks and systematic uncertainties

Figure 7.5: Comparison between data (dots) and MC (histograms) in the $E_{\text{extra}}$ sideband: $K^*$ mass for $K^{*+} \rightarrow K^+\pi^0$ (left), $K^{*+} \rightarrow K_S^0\pi^+$ (center), and $K^{*0} \rightarrow K^+\pi^-$ (right). The same legend as in Fig. 7.1 is adopted.

Figure 7.6: Comparison between data (dots) and MC (histograms) in the $E_{\text{extra}}$ sideband: $K_S^0$ mass for $K^{*+} \rightarrow K_S^0\pi^+$. The same legend as in Fig. 7.1 is adopted.

Figure 7.7: Comparison between data (dots) and MC (histograms) in the $E_{\text{extra}}$ sideband: $\cos\theta_{\text{miss}}^*$ for $K^{*+} \rightarrow K^+\pi^0$ (left), $K^{*+} \rightarrow K_S^0\pi^+$ (center), and $K^{*0} \rightarrow K^+\pi^-$ (right). The same legend as in Fig. 7.1 is adopted.
Table 7.1: Expected background and observed data in the $E_{\text{extra}}$ sideband for $K^{*+} \rightarrow K^+\pi^0$ decay mode; $N_{\text{bkg}}^{\text{exp}}$ is the total number of background events in the MC sample, re-weighted to data luminosity.

<table>
<thead>
<tr>
<th>$R_2$</th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$\bar{c}\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N_{\text{bkg}}^{\text{exp}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>190.6</td>
<td>25.3</td>
<td>288.4</td>
<td>301.6</td>
<td>0</td>
<td>0</td>
<td>806 ± 29</td>
<td>684</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta_{R,T}</td>
<td>$</td>
<td>151.3</td>
<td>19.1</td>
<td>62.2</td>
<td>64.6</td>
<td>0</td>
</tr>
<tr>
<td>$m_{K^*+}$ (GeV/c²)</td>
<td>124.8</td>
<td>15.6</td>
<td>50.9</td>
<td>50.1</td>
<td>0</td>
<td>241 ± 15</td>
<td>214</td>
</tr>
<tr>
<td>$\cos \theta_{\text{miss}}$</td>
<td>99.3</td>
<td>13.3</td>
<td>40.0</td>
<td>39.5</td>
<td>0</td>
<td>192 ± 14</td>
<td>174</td>
</tr>
<tr>
<td>$E_{\text{miss}}^* + \vec{p}_{\text{miss}}^*$ (GeV)</td>
<td>7.3</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>9 ± 3</td>
<td>13</td>
</tr>
</tbody>
</table>
Cross-checks and systematic uncertainties

Table 7.2: Expected background and observed data in the $E_{\text{extra}}$ sideband for $K^{*+} \rightarrow K^0\pi^+$ decay mode; $N_{\text{bkg}}^{\text{exp}}$ is the total number of background events in the MC sample, re-weighted to data luminosity.

<table>
<thead>
<tr>
<th></th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N_{\text{bkg}}^{\text{exp}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>857.9</td>
<td>192.5</td>
<td>704.7</td>
<td>724.7</td>
<td>0</td>
<td>2,480 ± 50</td>
<td>2,588</td>
</tr>
<tr>
<td>$</td>
<td>\cos\theta^*_{B,T}</td>
<td>$</td>
<td>730.1</td>
<td>159.0</td>
<td>289.0</td>
<td>256.3</td>
<td>0</td>
</tr>
<tr>
<td>$m_{K^*}$ (GeV/$c^2$)</td>
<td>496.5</td>
<td>104.8</td>
<td>188.7</td>
<td>161.9</td>
<td>0</td>
<td>952 ± 31</td>
<td>1,065</td>
</tr>
<tr>
<td>$m_{K^0_S}$ (GeV/$c^2$)</td>
<td>172.8</td>
<td>50.5</td>
<td>60.7</td>
<td>45.3</td>
<td>0</td>
<td>309 ± 18</td>
<td>361</td>
</tr>
<tr>
<td>$\cos\theta^*_{\text{miss}}$</td>
<td>129.1</td>
<td>25.0</td>
<td>39.5</td>
<td>27.0</td>
<td>0</td>
<td>221 ± 15</td>
<td>274</td>
</tr>
<tr>
<td>$E_{\text{miss}}^* + p_{\text{miss}}^*$ (GeV)</td>
<td>2.0</td>
<td>7.5</td>
<td>0.5</td>
<td>1.0</td>
<td>0</td>
<td>11 ± 3</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 7.3: Expected background and observed data in the $E_{\text{extra}}$ sideband for $K^{*0} \rightarrow K^+\pi^-$ decay mode; $N_{\text{bkg}}^{\text{exp}}$ is the total number of background events in the MC sample, re-weighted to data luminosity.

<table>
<thead>
<tr>
<th></th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N_{\text{bkg}}^{\text{exp}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>73.8</td>
<td>275.5</td>
<td>310.1</td>
<td>156.1</td>
<td>0</td>
<td>816 ± 29</td>
<td>758</td>
</tr>
<tr>
<td>$</td>
<td>\cos\theta^*_{B,T}</td>
<td>$</td>
<td>55.9</td>
<td>230.4</td>
<td>81.0</td>
<td>46.3</td>
<td>0</td>
</tr>
<tr>
<td>$m_{K^*}$ (GeV/$c^2$)</td>
<td>44.0</td>
<td>189.2</td>
<td>65.2</td>
<td>37.6</td>
<td>0</td>
<td>360 ± 19</td>
<td>359</td>
</tr>
<tr>
<td>$\cos\theta^*_{\text{miss}}$</td>
<td>37.4</td>
<td>151.6</td>
<td>43.0</td>
<td>22.2</td>
<td>0</td>
<td>254 ± 16</td>
<td>286</td>
</tr>
<tr>
<td>$E_{\text{miss}}^* + p_{\text{miss}}^*$ (GeV)</td>
<td>2.0</td>
<td>7.5</td>
<td>0.5</td>
<td>1.0</td>
<td>0</td>
<td>11 ± 3</td>
<td>7</td>
</tr>
</tbody>
</table>
on-peak data have been used. The reconstruction acts as follows: a list of reconstructed $B \rightarrow D^{(*)}X$ is made and the best one (according to the Semi-Exclusive $B$ reconstruction criteria) is chosen as $B_{\text{had}}$; another $B (B_{\text{sig}})$ that does not share any track or neutral with the $B_{\text{had}}$ is then searched for.

Separate studies have been performed for charged and neutral reconstructed $B_{\text{had}}$. The following selection is applied:

1. charge of $B_{\text{had}} = \pm 1$ (or 0),
2. at least one $B_{\text{sig}}$ reconstructed: in case of multiple $B_{\text{sig}}$ candidates, the one with the highest a-priori purity is chosen,
3. $5.27 < m_{ES} < 5.285$ GeV/$c^2$ for $B_{\text{had}}$,
4. $-0.05 < \Delta E < 0.05$ GeV for $B_{\text{had}}$,
5. total charge of the event $Q_{\text{tot}} = 0$,
6. $R_2 < 0.5$,
7. $|\cos \theta^*_B| < 0.85$
8. $E_{\text{extra}} < 0.6$ GeV for the signal region
   $E_{\text{extra}} > 0.7$ GeV for the sideband

The comparison between MC and data events which are retained at each selection step is shown in Tabs. 7.4-7.5: in the cut and count analysis, such results are used to associate a systematic uncertainties on the background yield estimation to account for data-MC discrepancies (Sec. 7.2.3).

[h!tbp]

7.1.3 Neural Network validation

The validation of the Neural Network has been accomplished on different control samples. To check the data-MC agreement the $m_{ES}$ sideband, where $m_{ES} < 5.26$ GeV/$c^2$, and the “wrong charge” sample, that is obtained by reversing the cut on the flavor correlation (Tab. 5.6), have been used. To validate the Neural Network output shape for the $B\bar{B}$ sample, the generic $B\bar{B}$ and the cocktail $B\bar{B}$ are compared.
Cross-checks and systematic uncertainties

Table 7.4: Cut flow table for double hadronic tag sample for charged $B_{\text{had}}$; the MC is re-weighted to data luminosity.

<table>
<thead>
<tr>
<th></th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N_{\text{exp}}^{\text{bkg}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>no cuts</td>
<td>10,533,800</td>
<td>9,122,490</td>
<td>18,296,700</td>
<td>17,361,900</td>
<td>36,970</td>
<td>55,351,860±7,439</td>
<td>60,110,418</td>
</tr>
<tr>
<td>$Q_{B_{\text{had}}}$</td>
<td>7,543,480</td>
<td>4,275,700</td>
<td>12,425,000</td>
<td>12,563,200</td>
<td>29,936.9</td>
<td>36,837,316±6,069</td>
<td>40,216,800</td>
</tr>
<tr>
<td>$m_{ES_{B_{\text{had}}}}$</td>
<td>1,931,820</td>
<td>548,742</td>
<td>1,595,190</td>
<td>1,671,960</td>
<td>4,594.9</td>
<td>5,752,307±2,398</td>
<td>6,117,100</td>
</tr>
<tr>
<td>$\Delta E_{B_{\text{had}}}$</td>
<td>1,666,480</td>
<td>439,949</td>
<td>1,244,750</td>
<td>1,288,820</td>
<td>3,195.5</td>
<td>4,643,194±2,155</td>
<td>4,946,790</td>
</tr>
<tr>
<td>$n_{B_{\text{sig}}}$</td>
<td>1,939.9</td>
<td>26.9</td>
<td>186.7</td>
<td>260.2</td>
<td>0</td>
<td>2,414±49</td>
<td>2,555</td>
</tr>
<tr>
<td>$Q_{\text{tot}}$</td>
<td>1,870.4</td>
<td>8.8</td>
<td>138.3</td>
<td>196.6</td>
<td>0</td>
<td>2,214±47</td>
<td>2,259</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta_{B,T}^*</td>
<td>$</td>
<td>1,388.1</td>
<td>7.5</td>
<td>109.6</td>
<td>142.6</td>
<td>0</td>
</tr>
<tr>
<td>$R_2$</td>
<td>1,304.0</td>
<td>6.8</td>
<td>69.1</td>
<td>65.5</td>
<td>0</td>
<td>1,446±38</td>
<td>1,463</td>
</tr>
<tr>
<td>$E_{\text{extra}} &lt; 0.6$</td>
<td>875.8</td>
<td>3.9</td>
<td>36.6</td>
<td>33.7</td>
<td>0</td>
<td>950±31</td>
<td>834</td>
</tr>
<tr>
<td>$E_{\text{extra}} &gt; 0.7$</td>
<td>339.2</td>
<td>2.6</td>
<td>26.2</td>
<td>28.9</td>
<td>0</td>
<td>397±20</td>
<td>517</td>
</tr>
</tbody>
</table>
Table 7.5: Cut flow table for double hadronic tag sample for neutral $B_{\text{had}}$; the MC is re-weighted to data luminosity.

<table>
<thead>
<tr>
<th></th>
<th>$B^+B^-$</th>
<th>$B^0\bar{B}^0$</th>
<th>$c\bar{c}$</th>
<th>$uds$</th>
<th>$\tau^+\tau^-$</th>
<th>$N_{\text{exp}}^{\text{bkg}}$</th>
<th>$N_{\text{obs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>no cuts</td>
<td>10,533,800</td>
<td>9,122,490</td>
<td>18,296,700</td>
<td>17,361,900</td>
<td>36,970</td>
<td>55,351,860±7,439</td>
<td>60,110,418</td>
</tr>
<tr>
<td>$Q_{B_{\text{had}}}$</td>
<td>2,990,320</td>
<td>4,846,790</td>
<td>5,870,780</td>
<td>4,798,750</td>
<td>7,033.4</td>
<td>18,513,673±4,303</td>
<td>19,893,600</td>
</tr>
<tr>
<td>$m_{\text{ESB}_{\text{had}}}$</td>
<td>386,382</td>
<td>2,119,811</td>
<td>744,507</td>
<td>658,451</td>
<td>1,183.2</td>
<td>3,910,334±1,382</td>
<td>3,143,140</td>
</tr>
<tr>
<td>$\Delta E_{B_{\text{had}}}$</td>
<td>304,394</td>
<td>1,034,430</td>
<td>570,087</td>
<td>499,698</td>
<td>880.8</td>
<td>2,409,490±1,552</td>
<td>2,543,620</td>
</tr>
<tr>
<td>$n_{B_{\text{sig}}}$</td>
<td>22.2</td>
<td>767.9</td>
<td>90.0</td>
<td>69.4</td>
<td>0</td>
<td>940±31</td>
<td>1,106</td>
</tr>
<tr>
<td>$Q_{\text{tot}}$</td>
<td>6.6</td>
<td>683.5</td>
<td>34.1</td>
<td>31.8</td>
<td>0</td>
<td>756±27</td>
<td>840</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta_{B_{\text{had}}^+}^*$</td>
<td>4.0</td>
<td>521.9</td>
<td>25.2</td>
<td>27.0</td>
<td>0</td>
<td>578±24</td>
</tr>
<tr>
<td>$R_2$</td>
<td>3.6</td>
<td>504.4</td>
<td>17.3</td>
<td>12.5</td>
<td>0</td>
<td>538±23</td>
<td>581</td>
</tr>
<tr>
<td>$E_{\text{extra &lt; 0.6}}$</td>
<td>2.0</td>
<td>324.6</td>
<td>8.4</td>
<td>6.7</td>
<td>0</td>
<td>342±18</td>
<td>348</td>
</tr>
<tr>
<td>$E_{\text{extra &gt; 0.7}}$</td>
<td>1.6</td>
<td>142.2</td>
<td>8.4</td>
<td>5.8</td>
<td>0</td>
<td>158±13</td>
<td>191</td>
</tr>
</tbody>
</table>
Tab. 7.6 shows the number of events on data and MC (normalized to data expectation) for the $m_{ES}$ sideband and the “wrong charge” samples: in most of the cases the agreement is within 1 sigma in the fit region. In Figs. 7.10-7.11 the $NN_{out}$ shapes for data and MC and the ratio of the two distributions are plotted. In Fig. 7.12 the distributions of $NN_{out}$ for generic $B \bar{B}$ and cocktail superimposed are plotted and a good agreement between the two MC samples is shown.

Table 7.6: Number of events on data and MC (renormalized to data expectation) for the $m_{ES}$ sideband and the “wrong charge” samples, for each $K^*$ decay mode.

<table>
<thead>
<tr>
<th></th>
<th>Events ($NN_{out} &gt; 0.6$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^{++} \rightarrow K^+ \pi^0$</td>
<td></td>
</tr>
<tr>
<td>$m_{ES}$ data</td>
<td>863 ± 29</td>
</tr>
<tr>
<td>$m_{ES}$ MC</td>
<td>854 ± 23</td>
</tr>
<tr>
<td>wrong charge data</td>
<td>361 ± 19</td>
</tr>
<tr>
<td>wrong charge MC</td>
<td>388 ± 14</td>
</tr>
<tr>
<td>$K^{+0} \rightarrow K^0 \pi^+$</td>
<td></td>
</tr>
<tr>
<td>$m_{ES}$ data</td>
<td>4,025 ± 63</td>
</tr>
<tr>
<td>$m_{ES}$ MC</td>
<td>4,067 ± 33</td>
</tr>
<tr>
<td>wrong charge data</td>
<td>1,085 ± 32</td>
</tr>
<tr>
<td>wrong charge MC</td>
<td>1,062 ± 16</td>
</tr>
<tr>
<td>$K^{00} \rightarrow K^0 \pi^-$</td>
<td></td>
</tr>
<tr>
<td>$m_{ES}$ data</td>
<td>2,222 ± 47</td>
</tr>
<tr>
<td>$m_{ES}$ MC</td>
<td>2,463 ± 37</td>
</tr>
<tr>
<td>wrong charge data</td>
<td>1,270 ± 36</td>
</tr>
<tr>
<td>wrong charge MC</td>
<td>1,350 ± 26</td>
</tr>
</tbody>
</table>

### 7.2 Evaluation of systematic uncertainties

In Sec. 8.1 it will be shown that the branching fraction is obtained from:

$$B = \frac{N_{sig}}{\varepsilon_{B_{sig}} \cdot N_{B_{had}}} \cdot \frac{B_{B,MC}}{\varepsilon_{B_{had}}}$$ \quad (7.2)

Systematic uncertainties on the following quantities should therefore be accounted for:

- signal efficiency
- $B_{had}$ yield estimation
7.2 Evaluation of systematic uncertainties

Figure 7.10: $m_{ES}$ sideband: $NN_{out}$ shapes for data (bars) and MC (histogram) on the top and the ratio of the two distributions on the bottom; $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K_S^0\pi^+$ on the center and $K^{*0} \rightarrow K^+\pi^-$ on the right.

Figure 7.11: “wrong charge” sample: $NN_{out}$ shapes for data (bars) and MC (histogram) on the top and the ratio of the two distributions on the bottom; $K^{*+} \rightarrow K^+\pi^0$ on the left, $K^{*+} \rightarrow K_S^0\pi^+$ on the center and $K^{*0} \rightarrow K^+\pi^-$ on the right.
Cross-checks and systematic uncertainties

Figure 7.12: Comparison between $NN_{\text{out}}$ shapes for generic $B\overline{B}$ MC (histogram) and $B\overline{B}$ cocktail MC (bars) in the full $NN_{\text{out}}$ region (top) and for $NN_{\text{out}} > 0.1$ (bottom); $K^{*+} \rightarrow K^{+}\pi^0$ on the left, $K^{*+} \rightarrow K_S^{0}\pi^+$ on the center and $K^{*0} \rightarrow K^+\pi^-$ on the right.

- background yield estimation in the cut and count approach
- $N_{\text{sig}}$ estimation in the Neural Network fit.

Their evaluation is discussed in the following sections. A summary of the systematic contributions is given in Tab. 7.11. In the cut and count analysis the main systematic contribution arises from the background estimation, while for the Neural Network fit the uncertainty on $N_{\text{sig}}$ dominates.

### 7.2.1 Systematic uncertainties on signal efficiency

The signal efficiency, estimated from the signal MC sample, is affected by systematics mainly due to possible discrepancies between simulated and real data. The limited size of the MC sample and effects related to the kinematic modeling of the signal process are also considered.

#### MC statistics

The systematic uncertainty due to the limited signal MC sample that has been used, is given by the formula:

$$\frac{\delta \varepsilon_{\text{B}_{\text{sig}}}}{\varepsilon_{\text{B}_{\text{sig}}}} = \sqrt{\frac{1 - \varepsilon_{\text{B}_{\text{sig}}}}{\varepsilon_{\text{B}_{\text{sig}}} \cdot N_{B_{\text{had}}}}}$$

(7.3)
$N_{B_{\text{had}}}^{K^*\nu\tau}$ being the number of events with one correctly reconstructed $B_{\text{had}}$ in the signal MC sample.

**Selection variables**

The signal reconstruction efficiency is estimated using the MC signal sample and a discrepancy between real and simulated data should be accounted for. For this purpose, high statistic MC and data samples with a similar pattern with respect to the $B \rightarrow K^*\nu\tau$ signal would be needed. In this way, the comparison of the selection efficiency in MC and data would give a correction factor to $\varepsilon_{K^*\nu\tau}^{\text{sig}}$. Since such a sample is not available, the selection cut values (cut and count) or the selection variables distributions (Neural Network fit) are varied to “simulate” a data-MC disagreement.

- **Cut and count**

  The selection requirements are shifted and the variation in the signal efficiency is evaluated. In this way, if the selection variable $x$ ranges in $[a, b]$ in the MC sample, in the data it is assumed to lie between $a \pm \delta$ and $b \pm \delta$ with the same distribution with respect to the simulated data. Different strategies are adopted to evaluate the shift $\delta$ for the selection variables: when possible the width of the $x$ variable distribution in the MC is estimated and half of this quantity is assumed as $\delta$; alternatively the physical effects not properly simulated that could affect the $x$ distributions are evaluated to define a suitable $\delta$; ultimately, if the previous two are not feasible, a conservative estimation of the shift is adopted. Here are the details for each variable:

  - $R_2$: $\delta = 0.025$ is conservatively estimated;
  - $\cos \theta_{B,T}^*$: $\delta = 0.015$ is conservatively estimated;
  - $K^*$ mass and $K_S^0$ mass: the width of the distributions are close to the physical width of the two meson masses, so a shift on data could be due only to resolution effects that are estimated to be 5 MeV/$c^2$ for $K^*$ and 1 MeV/$c^2$ for $K_S^0$;
  - $\cos \theta_{\text{miss}}^*$: the distribution of $\theta_{\text{miss}}^*$ has a width of about 0.5 rad; propagating a shift of 0.25 rad on $\cos \theta_{\text{miss}}^*$, $\delta = 0.025$ is found;
  - $E_{\text{miss}}^* + p_{\text{miss}}^*$: $\delta$ is defined by propagating the half-width of $E_{\text{miss}}^*$ and $p_{\text{miss}}^*$ into their sum, as a result $\delta = 0.075$;
– $E_{\text{extra}}$: in this case the disagreement between data and MC could be mainly due to a different resolution in the $\gamma$ energy and a bad modeling of the $\gamma$ energy loss; in both cases a shift of some MeV in the $E_{\text{extra}}$ distribution is estimated and $\delta=10$ MeV is conservatively chosen.

In Tab. 7.7 a breakdown of each variable contribution to the systematic error on the selection efficiency is shown.

Table 7.7: Cut and count analysis: systematics on the selection efficiency due to the signal selection variable distributions for each $K^*$ decay mode. The total error is given by the sum in quadrature of each contribution.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$K^{*+} \rightarrow K^+\pi^0$</th>
<th>$K^{*+} \rightarrow K^0\pi^+$</th>
<th>$K^{*0} \rightarrow K^+\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>1.5%</td>
<td>4.2%</td>
<td>0.8%</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta_{B,T}^*</td>
<td>$</td>
<td>1.0%</td>
</tr>
<tr>
<td>$m_{K^*}$</td>
<td>1.0%</td>
<td>2.0%</td>
<td>2.9%</td>
</tr>
<tr>
<td>$m_{K^0_s}$</td>
<td>-</td>
<td>3.5%</td>
<td>-</td>
</tr>
<tr>
<td>$\cos b_{\text{miss}}$</td>
<td>1.5%</td>
<td>1.4%</td>
<td>2.3%</td>
</tr>
<tr>
<td>$E_{\text{miss}}^* + p_{\text{miss}}^*$</td>
<td>2.0%</td>
<td>2.8%</td>
<td>3.1%</td>
</tr>
<tr>
<td>$E_{\text{extra}}$</td>
<td>1.2%</td>
<td>1.4%</td>
<td>2.3%</td>
</tr>
<tr>
<td>Total</td>
<td>3.4%</td>
<td>7.0%</td>
<td>6.0%</td>
</tr>
</tbody>
</table>

- **Neural Network fit**

For the Neural Network fit the same philosophy adopted in the cut and count is used, but in this case the Neural Network input distributions are properly modified. The hypothesis is that the data-MC agreement could be so bad that in data the discriminating power of each input variable is softened (i.e. in data, the signal distribution is more “similar” to the background one with respect to the simulations). Separately for each variable, the distributions are shifted and smeared to obtain a “data-like” sample: this is done by shifting the value of each variable, event by event, by a number that is generated with a Gaussian distribution with a mean $m$ and a width $\sigma$ (Tab. 7.8). The Neural Network is then run on the “data-like” sample (one for each variable) and the requirement $NN_{\text{out}}>0.6$ is applied: the number of events surviving the selection gives the new efficiency. In Tab. 7.9 the effects due to the smearing of each variable distributions in the signal efficiency are listed.
Table 7.8: Values of the mean and the width of the Gaussian functions used to generate the shift in the Neural Network input variable distributions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$m$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>0.025</td>
<td>$2 \times 0.025$</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta_{B,T}^*</td>
<td>$</td>
</tr>
<tr>
<td>$m_{K^*}$</td>
<td>0</td>
<td>$4 \times 0.005$</td>
</tr>
<tr>
<td>$m_{K^0_S}$</td>
<td>0</td>
<td>$2 \times 0.001$</td>
</tr>
<tr>
<td>$\cos \theta_{\text{miss}}^*$</td>
<td>$5 \times 0.025$</td>
<td>$5 \times 0.025$</td>
</tr>
<tr>
<td>$E_{\text{miss}}^* + p_{\text{miss}}^*$</td>
<td>$-1 \times 0.075$</td>
<td>$2 \times 0.075$</td>
</tr>
<tr>
<td>$E_{\text{extra}}$</td>
<td>$5 \times 0.01$</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 7.9: Neural Network fit: systematics on the selection efficiency due to the signal selection variable distributions for each $K^*$ decay mode. The total error is given by the sum in quadrature of each contribution.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$K^{*+} \rightarrow K^{+}\pi^0$</th>
<th>$K^{*+} \rightarrow K^{0}_{S}\pi^+$</th>
<th>$K^{*0} \rightarrow K^{+}\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2$</td>
<td>1.8%</td>
<td>2.2%</td>
<td>0.9%</td>
</tr>
<tr>
<td>$</td>
<td>\cos \theta_{B,T}^*</td>
<td>$</td>
<td>1.9%</td>
</tr>
<tr>
<td>$m_{K^*}$</td>
<td>1.3%</td>
<td>2.2%</td>
<td>0.9%</td>
</tr>
<tr>
<td>$m_{K^0_S}$</td>
<td>-</td>
<td>1.6%</td>
<td>-</td>
</tr>
<tr>
<td>$\cos \theta_{\text{miss}}^*$</td>
<td>1.0%</td>
<td>1.1%</td>
<td>1.3%</td>
</tr>
<tr>
<td>$E_{\text{miss}}^* + p_{\text{miss}}^*$</td>
<td>3.0%</td>
<td>4.0%</td>
<td>1.7%</td>
</tr>
<tr>
<td>$E_{\text{extra}}$</td>
<td>3.1%</td>
<td>6.4%</td>
<td>2.1%</td>
</tr>
<tr>
<td>Total</td>
<td>5.3%</td>
<td>8.6%</td>
<td>3.8%</td>
</tr>
</tbody>
</table>
Cross-checks and systematic uncertainties

Tracking reconstruction

The differences in data and MC tracking efficiency are evaluated by using an $e^+e^- \to \tau^+\tau^-$ sample in which one $\tau$ decays leptonically and the other in a three-hadrons final state with neutrinos and neutrons. This constitutes a large control sample due to the high $e^+e^- \to \tau^+\tau^-$ cross section (0.94 nb) and the signal branching fraction to lepton plus 3 hadrons (11%). Moreover the momentum distribution of $\tau$ decay products is similar to the one of the $B$ daughters. Events with three to five tracks are selected, when an odd number of tracks is reconstructed, one has been missed or is fake. The comparison between the efficiency in reconstructing the right number of charged tracks in data and MC samples gives a relative systematic uncertainty to $\varepsilon_{B_{\text{sig}}}$: 0.3%, 1.0%, and 0.7% for $K^{*+} \to K^{+}\pi^0$, $K^{*+} \to K_s^0\pi^+$, and $K^{*0} \to K^+\pi^-$ respectively.

$K_s^0$ reconstruction

A systematic uncertainty due to the $K_s^0 \to \pi^+\pi^-$ reconstruction efficiency is estimated by using a control sample of $B$ decays to a final state with two charged tracks plus a $K_s^0$. Fits to the $K_s^0$ mass in bins of the $K_s^0$ transverse momentum, the transverse distance between the $B$ and the $K_s^0$ decay vertices, and the polar angle are performed to extract the number of $K_s^0$ both in data and MC. The discrepancy between the two bin-by-bin fitted values are combined to give a systematic error of 2.6% to the $K^{*+} \to K_s^0\pi^+$ signal efficiency.

$\pi^0$ reconstruction

$e^+e^- \to \tau^+\tau^-$ events are also used for the systematic associated to the $\pi^0$ reconstruction. In this case the ratio between the number of $\tau \to e \nu \overline{\nu}$ and the number of $\tau \to h\pi^0\pi^0\nu$ (where $h$ is a hadron) is computed both in data and MC, as a function of the $\pi^0$ energy. The comparison between the two leads to a correction factor of $0.984 \pm 0.030$: the central value and the error, multiplied to the estimated $\varepsilon_{B_{\text{sig}}}$ for $K^{*+} \to K^{+}\pi^0$, give the corrected efficiency and the associated systematic error, respectively.

$K$ particle identification

The difference in identifying a charged kaon between data and MC samples is evaluated with the following procedure: a kaon that has previously passed the particle identification requirement is rejected and one that was previously rejected is accepted, with probabilities
extracted from data and MC control samples. The comparison between the nominal and
the new efficiency gives a systematics of 1.6% and 1.5%, associated to the $K^{*+} \rightarrow K^+\pi^0$
and $K^{*0} \rightarrow K^+\pi^-$ modes, respectively.

**Model dependence: phase space re-weighting**

In Sec. 5.4.2 it has been anticipated that the signal MC is generated according to a phase
space model; to extract the final result no MC re-weighting related to a specific physics
model has been applied. The effect of assuming a given kinematic distribution is quantified
by applying a re-weighting according to the SM prediction [23], using the two methods
described in Sec. 5.4.2: the Hit and Miss and the Re-weighting. The results are listed in
Tab. 7.10. Different values between the two procedures are found: to be conservative, the
largest one is taken as uncertainty.

**Table 7.10: Systematic error on the selection efficiency due to the MC phase space re-
weighting.**

<table>
<thead>
<tr>
<th></th>
<th>Cut and Count</th>
<th>Neural Network Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Re-weighting</td>
<td>Hit and Miss</td>
</tr>
<tr>
<td>$K^{*+} \rightarrow K^+\pi^0$</td>
<td>6.9%</td>
<td>7.2%</td>
</tr>
<tr>
<td>$K^{*+} \rightarrow K^0_S\pi^+$</td>
<td>6.8%</td>
<td>6.6%</td>
</tr>
<tr>
<td>$K^{*0} \rightarrow K^+\pi^-$</td>
<td>6.7%</td>
<td>5.4%</td>
</tr>
</tbody>
</table>

**7.2.2 Systematic uncertainties on $B_{\text{had}}$ yield determination**

In order to distinguish between peaking and combinatorial $B_{\text{had}}$ candidates, requirements
on the truth matched tracks in the tag side and on the angle between reconstructed and
true $B_{\text{had}}$ momenta ($\Delta \theta_B$) have been made (Sec 4.3). Different values of the cut on $\Delta \theta_B$
may change the tag yield; its estimation is therefore repeated by varying the cut on $\Delta \theta_B$
of $\pm 0.01$ rad. The resulting systematic errors on $N_{B_{\text{had}}}$ are 0.15\% for the charged tag
and 2.1\% for the neutral one.

From the fit to the $m_{ES}$ data sideband, the relative normalization of the combinatorial
$B^+B^-$ and $B^0\overline{B}^0$ is estimated, in particular the $B^+B^-$ fraction; this fraction is different
with respect to what expected from the MC: $f_{B^+B^-,\text{MC}} = 17\%$ for the charged tag and
$f_{B^+B^-,\text{MC}} = 16\%$ for the neutral one. A systematic uncertainty (3.20\% for $N_{B^+_{\text{had}}}$ and
Cross-checks and systematic uncertainties

0.50% for \(N_{B^+_{\text{had}}}\) is computed by fixing the relative normalization of combinatorial \(B^+B^-\) and \(B^0\bar{B}^0\) events to the MC expectation.

To take into account for the peak in the \(m_{ES}\) \(B^0\bar{B}^0\) combinatorial distribution (Fig. 4.7 bottom right), the fraction of peaking events in the \(m_{ES}\) signal region that falls in the combinatorial category is estimated using the \(B^0\bar{B}^0\) MC sample. In Fig. 7.13 the fit to the \(m_{ES}\) \(B^0\bar{B}^0\) combinatorial distribution to the sum of a Gaussian and an Argus in the \(m_{ES}\) signal region is shown: the fraction of peaking events described from the Gaussian is 5%. By multiplying this fraction for the expected number of combinatorial \(B^0\bar{B}^0\) events in the data \(m_{ES}\) signal region a systematic error of 1.9% for \(N_{B^0_{\text{had}}}\) is found. The effect is negligible for the charged sample.

Fig. 4.9 (bottom left) shows that the background subtraction in the \(m_{ES}\) sideband leads to an overestimation of the background that can be propagated also in the signal region. A systematic error to \(N_{B^+_{\text{had}}}\) is assigned, using the following procedure: the absolute difference between the estimated background component and the data in each bin of the \(m_{ES}\) sideband (5.22 < \(m_{ES}\) < 5.26 GeV/\(c^2\)) is computed and summed over all bins, the expected underestimation on \(N_{B^+_{\text{had}}}\) in the signal region is estimated as such a deviation multiplied by the ratio between the expected background yield in the signal region and in the sideband. This correspond to a systematic uncertainty of 0.83% on the charged \(B_{\text{had}}\) yield. The effect is negligible for \(N_{B^0_{\text{had}}}\).

The decay of the \(B\) recoiling against the \(B_{\text{had}}\), may affect the tag efficiency and the performance of the tag reconstruction may be different in a low multiplicity events as \(B \rightarrow K^*\nu\bar{\nu}\), with respect to a generic \(B\) decay in the signal side. To account for it in the branching fraction estimation (as discussed in Sec. 8.1), a correction factor given by the ratio of the tag efficiency in signal and \(B\bar{B}\) MC (Sec. 4.3) is computed:

- charged tag:
  \[
  c_{B^+_{\text{had}}} = 1.008 \pm 0.007 \tag{7.4}
  \]
- neutral tag:
  \[
  c_{B^0_{\text{had}}} = 1.176 \pm 0.013. \tag{7.5}
  \]

The error on the correction factors gives a systematic uncertainties on the tag yield, to be summed in quadrature with the uncertainties due to changes in the parametrization and normalization of \(B\bar{B}\) combinatorial components.

The overall relative systematic error on the tad yield is:
7.2 Evaluation of systematic uncertainties

- charged tag: 3.4%
- neutral tag: 3.1%

Figure 7.13: Fit to the $m_{ES}$ distribution for the $B^0\bar{B}^0$ combinatorial MC sample to extract the fraction of the peaking component.

7.2.3 Systematic uncertainty on the background yield estimation in the cut and count analysis

In the background yield computation (Sec. 6.1.2), the ratio of events in the $E_{\text{extra}}$ sideband and signal region is assumed to be the same for data and MC. To assign a systematic, the two ratios are computed in the double tag sample:

$$r_{\text{MC,dt}} = \frac{B(E_{\text{extra}} < 0.6 \text{ GeV})_{\text{MC,dt}}}{B(E_{\text{extra}} > 0.7 \text{ GeV})_{\text{MC,dt}}} \quad (7.6)$$

- charged sample: 2.39
- neutral sample: 1.82

$$r_{\text{data,dt}} = \frac{B(E_{\text{extra}} < 0.6 \text{ GeV})_{\text{data,dt}}}{B(E_{\text{extra}} > 0.7 \text{ GeV})_{\text{data,dt}}} \quad (7.7)$$

- charged sample: 1.63
- neutral sample: 2.05

The value $1 - (r_{\text{data,dt}}/r_{\text{MC,dt}})$ defines a percentage error on the background yield estimation of 48% for the $K^{*+}$ channels and 18% for the $K^{*0}$, corresponding to an uncertainty of 9.0, 4.1, and 2.5 events for the $K^{*+} \rightarrow K^+\pi^0$, $K^{*+} \rightarrow K^0\pi^+$, and $K^{*0} \rightarrow K^+\pi^-$, respectively.
7.2.4 Systematic uncertainties on the signal yield estimation in the Neural Network fit

Systematics on the parametrization of the \( NN_{\text{out}} \) distribution

The parameters that define the signal and background PDFs for the Neural Network output are fixed to the values obtained from fits to the MC sample (Tabs. 6.11-6.10). A systematic error on the signal yield related to the statistical uncertainties on the parameters should be evaluated. For the three \( K^* \) mode, the background PDF is defined by \( k_1 \) and \( k_2 \) which are significantly correlated. To take into account the correlation, the following strategy is adopted: the covariance matrix \( V \) obtained from the MC fit is diagonalized by a transformation matrix \( M \). This transforms \( k_1 \) and \( k_2 \) and their errors in a set of independent parameters \((k'_1,k'_2; \sigma'_1, \sigma'_2)\). Varying \( k'_1 \) and \( k'_2 \) by their statistical error and going back to the \((k_1,k_2)\) basis, a new set of dependent parameters is defined. This is used in the data fit to compute the shift on \( N_{\text{sig}} \).

The signal PDFs for \( K^{*+} \rightarrow K^+ \pi^0 \) and \( K^{*+} \rightarrow K_S^0 \pi^+ \) are Crystal-Ball functions which depend on four parameters, one of them \((n_1)\) is fixed in the fit to the signal MC, while the three are floating. A fit to the signal MC, in which \( n_1 \) is floated, is performed; the signal shape is fixed according to this fit and the signal yield is extracted: the variation with respect to the nominal fit is found to be negligible. To evaluate the systematic error due to the statistical uncertainties on three free parameters of the Crystal-Ball, the same recipe used for the background PDFs is followed. The \( K^{*0} \rightarrow K^+ \pi^- \) signal PDF is defined by one parameter \( c \), that is varied by \( \pm \sigma_c \) in the data fit.

Systematics on the \( NN_{\text{out}} \) shape due to data-MC disagreement

The signal and background shapes of \( NN_{\text{out}} \), modeled from the MC, can differ from the data distribution.

For the background, data and MC distributions in the \( m_{ES} \) and flavor sidebands are compared, as shown on Sec. 7.1.3. A bin-by-bin ratio between the \( NN_{\text{out}} \) distributions in data and MC in the \( m_{ES} \) sideband is parametrized with a 1-order polynomial; this function is then multiplied for the background PDF function in the data fit and the change in \( N_{\text{sig}} \) is quoted as systematics.

For the signal, the “data-like” samples described in Sec. 7.2.1 is used. For each variable, the ratio of the \( NN_{\text{out}} \) distribution between “data-like” and MC samples is fitted with a
7.2 Evaluation of systematic uncertainties

1-order polynomial and the resulting function multiplies the signal PDF in the data fit. For each variable a shift on $N_{\text{sig}}$ is evaluated and the sum in quadrature of the effects due to each variable is taken as systematic uncertainties.

Table 7.11: Summary of systematic uncertainties on the signal efficiency, signal yield, background estimation, and normalization.

<table>
<thead>
<tr>
<th>$K^*$ mode</th>
<th>Cut and Count</th>
<th>Neural Network Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Signal efficiency (%)</td>
<td>Signal efficiency (%)</td>
</tr>
<tr>
<td>MC statistics</td>
<td>$K^+\pi^0$</td>
<td>$K^0\pi^+$</td>
</tr>
<tr>
<td>Selection variables</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>Tracking</td>
<td>0.3</td>
<td>1.0</td>
</tr>
<tr>
<td>$K^0_S$ reconstruction</td>
<td>–</td>
<td>2.5</td>
</tr>
<tr>
<td>$\pi^0$ reconstruction</td>
<td>3.0</td>
<td>–</td>
</tr>
<tr>
<td>Particle ID</td>
<td>1.5</td>
<td>–</td>
</tr>
<tr>
<td>Model dependence</td>
<td>6.7</td>
<td>6.8</td>
</tr>
<tr>
<td>Total</td>
<td>8.9</td>
<td>11.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Background yield (events)</th>
<th>$N_{\text{bkg}}$</th>
<th>$N_{\text{bkg}}$</th>
<th>$N_{\text{bkg}}$</th>
<th>$N_{\text{bkg}}$</th>
<th>$N_{\text{bkg}}$</th>
<th>$N_{\text{bkg}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal yield (events)</td>
<td>$N_{B_{\text{had}}}$</td>
<td>3.4</td>
<td>3.1</td>
<td>3.4</td>
<td>3.1</td>
<td>3.1</td>
</tr>
</tbody>
</table>
Cross-checks and systematic uncertainties
Chapter 8

Final result: theoretical interpretation and future prospects

In this chapter the physical results and its theoretical implications are discussed. In Sec. 8.1 the relation between the branching fraction and the signal yield is derived; the statistical method used to extract the upper limit is presented in Sec. 8.2. The experimental results are listed in Sec. 8.3: the combination between the semileptonic and hadronic tag analyses leads to the most stringent limit on the branching fraction under investigation reported to date.

As mentioned in Chapt. 1, several new physics models make predictions for $\mathcal{B}(B \to K^*\nu\overline{\nu})$. A study of some phenomenological constraints that can be set by exploiting the measurement presented in this thesis and the current knowledge of $B \to K^{(*)}\nu\overline{\nu}$ is documented in Sec. 8.4.

8.1 Branching fraction computation

In the data sample, let $N_{B\overline{B}}^{\text{data}}$, $N_{\text{gen}}^{\text{data}}$ and $N_{K^*\nu\overline{\nu}}^{\text{data}}$ be the total number of $B\overline{B}$ events, the number of $B\overline{B}$ events in which the $B_{\text{sig}}$ decayed generically and the number of events in which $B_{\text{sig}}$ decayed to the signal channel, respectively. The relation between the three quantities is the following:

$$N_{B\overline{B}}^{\text{data}} = N_{\text{gen}}^{\text{data}} + N_{K^*\nu\overline{\nu}}^{\text{data}}. \quad (8.1)$$

$N_{K^*\nu\overline{\nu}}^{\text{data}}$ could be expressed as a function of $\mathcal{B}(B \to K^*\nu\overline{\nu})$ as:

$$N_{K^*\nu\overline{\nu}}^{\text{data}} = 2 \cdot \mathcal{B}(B \to K^*\nu\overline{\nu}) \cdot N_{B\overline{B}}^{\text{data}}. \quad (8.2)$$
Eq. 8.1 becomes:

\[ N_{BB}^{\text{data}} = N_{\text{gen}}^{\text{data}} + 2 \cdot B(B \to K^* \nu \bar{\nu}) \cdot N_{BB}^{\text{data}} \]

\[ \simeq N_{\text{gen}}^{\text{data}}, \]

assuming \( B(B \to K^* \nu \bar{\nu}) \cdot N_{BB}^{\text{data}} \) negligible with respect to the first term. If \( \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \) and \( \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \) are the \( B_{\text{had}} \) selection efficiencies for a \( B \bar{B} \) and a signal events, the total number of \( B_{\text{had}} \) events in data is defined as:

\[ N_{B_{\text{had}}} = 2 \cdot N_{\text{gen}}^{\text{data}} \cdot \varepsilon_{B_{\text{had}}}^{BB} + 2 \cdot B(B \to K^* \nu \bar{\nu}) \cdot N_{BB}^{\text{data}} \cdot \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \]

\[ \simeq 2 \cdot N_{\text{gen}}^{\text{data}} \cdot \varepsilon_{B_{\text{had}}}^{BB}, \]

neglecting again the term proportional to \( B(B \to K^* \nu \bar{\nu}) \). In the light of the last equation, \( N_{\text{gen}}^{\text{data}} \) can be written as:

\[ N_{\text{gen}}^{\text{data}} = \frac{N_{B_{\text{had}}}^{BB}}{2 \cdot \varepsilon_{B_{\text{had}}}^{BB}}. \]  

Let \( N_{\text{sig}} \) and \( \varepsilon_{B_{\text{sig}}} \) be the measured number of signal events in data and the signal selection efficiency; the first one could be defined as:

\[ N_{\text{sig}} = \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \cdot \varepsilon_{B_{\text{sig}}} \cdot N_{\text{K}^* \nu \bar{\nu}}^{\text{data}} \]

\[ = \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \cdot \varepsilon_{B_{\text{sig}}} \cdot 2B(B \to K^* \nu \bar{\nu}) \cdot N_{BB}^{\text{data}} \cdot \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \]

\[ = \varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}} \cdot \varepsilon_{B_{\text{sig}}} \cdot 2B(B \to K^* \nu \bar{\nu}) \cdot \frac{N_{B_{\text{had}}}^{BB}}{2 \cdot \varepsilon_{B_{\text{had}}}^{BB}}, \]

having used Eqs. 8.2, 8.3, and 8.5. Note that, since the BF is normalized with respect to the \( B_{\text{had}} \) yield, \( \varepsilon_{B_{\text{sig}}} \) is given by the ratio of signal MC events that survive the signal and tag selection and the number of \( B_{\text{had}} \) events in the same sample. Assuming that systematic differences between data and MC in \( B_{\text{had}} \) reconstruction efficiency are the same for generic and signal events, the following relation holds:

\[ \frac{\varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}}}{\varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}, \text{MC}}} = \frac{\varepsilon_{B_{\text{had}}}^{BB}}{\varepsilon_{B_{\text{had}}}^{BB, \text{MC}}}, \]

Using Eqs. 8.6 and 8.7, the BF is finally obtained from:

\[ B(B \to K^* \nu \bar{\nu}) = \frac{N_{\text{sig}}}{\varepsilon_{B_{\text{sig}}} \cdot N_{B_{\text{had}}}^{BB}} \cdot \frac{\varepsilon_{B_{\text{had}}}^{BB, \text{MC}}}{\varepsilon_{B_{\text{had}}}^{K^* \nu \bar{\nu}, \text{MC}}}. \]
8.2 Statistical method for the upper limit extraction

A Bayesian approach [50] is used to extract the upper limit on the $B \rightarrow K^* \nu \tau$ branching fraction. Statistical and systematic uncertainties affecting the signal and background yield, the signal efficiency, and the normalization are taken into account.

A prior probability density function $P(B)$ for the branching fraction is defined as:

$$
P(B) = \begin{cases} 
P(B) \text{ uniform} & B > 0 \\ 
P(B) = 0 & B < 0 
\end{cases}$$  \hspace{1cm} (8.9)

imposing the physical constraint of a non-negative BF. Given a single $K^*$ channel, the Bayes theorem is applied to derive the posterior PDF for the branching fraction:

$$
P(B|data) \propto P(data|B)P(B) = \int d\varepsilonB_{sig}dN_{B\text{had}} \left[ \mathcal{L}(N_{sig} = B \cdot \varepsilonB_{sig} \cdot N_{B\text{had}})P(\varepsilonB_{sig})P(N_{B\text{had}})P(B) \right],$$  \hspace{1cm} (8.10)

where $\mathcal{L}(N_{sig})$ is likelihood function of the $NN_{out}$ distribution (in the Neural Network fit) or the Poissonian distribution for $N_{obs} = N_{sig} + N_{bkg}$ (in the cut and count analysis), both marginalized with respect to $N_{bkg}$, while $P(\varepsilonB_{sig})$ and $P(N_{B\text{had}})$ are suitable distributions describing the signal efficiency and $B_{\text{had}}$ yield. For simplicity, they are defined as Gaussian functions centered at the signal efficiency and $B_{\text{had}}$ yield estimations, respectively, and whose width is equal to the corresponding uncertainties.

The integral in Eq. (8.10) is evaluated numerically, with the following procedure:

1. generate a random value for the BF according to Eq. 8.9;
2. generate $\varepsilonB_{sig}$ and $N_{B\text{had}}$ according to the previously defined Gaussian distributions;
3. assign to each generated BF value a weight given by the $N_{sig}$ likelihood;
4. iterate several times;

The BF posterior distribution is give by the weighted distribution of the branching fraction. The value of the $B_{90\%}$ for which:

$$
\frac{\int_{0}^{B_{90\%}} P(B|data)d(B)}{\int_{0}^{\infty} P(B|data)d(B)} = 90\% 
$$  \hspace{1cm} (8.11)

sets the Upper Limit at 90% confidence level.
This approach allows also to combine different channels, i.e. \( K^{*+} \rightarrow K^0_s \pi^+ \) and \( K^{*+} \rightarrow K^+ \pi^0 \), by the following generalization of Eq. (8.10):

\[
P(B|data) \propto \int dN_{B_{\text{had}}} \prod_{\alpha} \int d\varepsilon_{\text{sig},\alpha} \mathcal{L}(N_{\alpha,\alpha} = B \cdot \varepsilon_{\text{sig},\alpha} \cdot N_{B_{\text{had}}}) P(\varepsilon_{\text{sig},\alpha}) P(N_{B_{\text{had}}}) P(B) \tag{8.12}
\]

where \( \alpha \in \{ K^{*+} \rightarrow K^+ \pi^0, K^{*+} \rightarrow K^0_s \pi \} \). In this case, different signal efficiency values are generated for the different channels, and the weight to the BF is given by the product of the different \( N_{\text{sig}} \) likelihoods. Correlation between various channels due to systematic uncertainties are taken into account. In principle, also cross-feed efficiencies, that are found to be zero for all the \( K^* \) decay modes, can be treated. With the same philosophy, the method allows to combine the semileptonic and hadronic recoil analyses as independent results.

### 8.3 Branching fraction measurement

Before applying the analysis to the real data, its sensitivity is estimated. The results obtained on the simulated sample are used to compute the expected upper limits evaluated as follow: for the cut and count analysis the number of observed events in data is assumed to correspond to the expected number of background; for the Neural Network fit, the signal yield and its statistical error are taken from toy MC studies. In Tab. 8.1 the quantities with which the expected result is estimated are listed. The expected ULs at 90% of confidence level are:

\[
\begin{align*}
\text{cut and count: } & \mathcal{B}(B^\pm \rightarrow K^{*\pm} \nu \overline{\nu}) < 2.8 \times 10^{-4} \tag{8.13} \\
\mathcal{B}(B^0 \rightarrow K^{*0} \nu \overline{\nu}) & < 1.8 \times 10^{-4} \\
\text{Neural Network fit: } & \mathcal{B}(B^\pm \rightarrow K^{*\pm} \nu \overline{\nu}) < 1.4 \times 10^{-4} \\
\mathcal{B}(B^0 \rightarrow K^{*0} \nu \overline{\nu}) & < 1.5 \times 10^{-4}
\end{align*}
\]

The Neural Network method results in better performances: by applying a loose selection on the discriminating variables, a larger statistics with respect to the cut and count is maintained and the larger background contamination is compensated by the better discriminating power of the fit with respect to the event counting.
The analysis results on data (Tab. 8.1) allow to determine the measured ULs:

\[
\begin{align*}
\text{cut and count: } & \mathcal{B}(B^\pm \to K^{*\pm} \nu \overline{\nu}) < 3.3 \times 10^{-4} \\
& \mathcal{B}(B^0 \to K^{*0} \nu \overline{\nu}) < 2.4 \times 10^{-4} \\
\text{Neural Network fit: } & \mathcal{B}(B^\pm \to K^{*\pm} \nu \overline{\nu}) < 2.1 \times 10^{-4} \\
& \mathcal{B}(B^0 \to K^{*0} \nu \overline{\nu}) < 1.1 \times 10^{-4}
\end{align*}
\]  

The posterior PDFs for the branching fraction (as computed from the Neural Network fit results) are plotted in Fig. 8.1.

In the SL analysis \[37\], ULs of \(0.8 \times 10^{-4}\) for the charged channel and \(1.8 \times 10^{-4}\) for the neutral one have been set; by combining the Neural Network fit results and the semileptonic analysis the following upper limits are obtained:

\[
\begin{align*}
\mathcal{B}(B^\pm \to K^{*\pm} \nu \overline{\nu}) < 0.8 \times 10^{-4} \\
\mathcal{B}(B^0 \to K^{*0} \nu \overline{\nu}) < 1.2 \times 10^{-4} \\
\mathcal{B}(B \to K^* \nu \overline{\nu}) < 0.8 \times 10^{-4}
\end{align*}
\]

Since no constraints have been applied on the \(K^*\) or neutrino pair kinematics, the results can be interpreted as a limit on the BF of \(B \to K^* +\text{invisible}\), in models where invisible particles, other than neutrinos, can produce missing energy \[27\], \[30\].

For comparison, the Belle results (already discussed in Sec. 1.5) are here recalled: 
\[
\mathcal{B}(B \to K^{*\pm} \nu \overline{\nu}) < 1.4 \times 10^{-4}, \mathcal{B}(B \to K^{*0} \nu \overline{\nu}) < 3.4 \times 10^{-4}.
\]

It can be notice that with less statistic (535 million \(B\overline{B}\) pairs for the Belle measurement to be compared to the 454 million used in this analysis) the present work sets the most stringent limits reported to date. In addition, the measured upper limits are still consistent with the SM expectation \[23\].

### 8.4 Phenomenological constraints from \(B \to K^{(*)}\nu \overline{\nu}\)

#### Analysis method and Standard Model expectation

Several New Physics models predict enhancements in the \(B \to K^{(*)}\nu \overline{\nu}\) branching fractions with respect to the SM expectation. The magnitude of NP contributions is related
Table 8.1: Values used to compute the expected and the measured upper limits on $\mathcal{B}(B^\pm \to K^{\pm}\nu\bar{\nu})$ and $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$.

<table>
<thead>
<tr>
<th></th>
<th>$K^{*+} \to K^+\pi^0$</th>
<th>$K^{*+} \to K_S^0\pi^+$</th>
<th>$K^{*0} \to K^+\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{B_{\text{had}}}$</td>
<td>1,012,788 ± 34,580</td>
<td>1,012,788 ± 34,580</td>
<td>717,490 ± 22,256</td>
</tr>
<tr>
<td></td>
<td>Cut and Count</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{B_{\text{sig}}}$</td>
<td>$(4.0 \pm 0.4) \times 10^{-2}$</td>
<td>$(3.0 \pm 0.3) \times 10^{-2}$</td>
<td>$(10.3 \pm 1.0) \times 10^{-2}$</td>
</tr>
<tr>
<td>$N_{\text{exp}}^{\text{bkg}}$</td>
<td>18.8 ± 12.6</td>
<td>8.5 ± 6.1</td>
<td>14.0 ± 7.4</td>
</tr>
<tr>
<td>$N_{\text{obs}}$</td>
<td>20</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Neural Network Fit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{B_{\text{sig}}}$</td>
<td>$(5.8 \pm 0.5) \times 10^{-2}$</td>
<td>$(5.2 \pm 0.6) \times 10^{-2}$</td>
<td>$(16.6 \pm 1.4) \times 10^{-2}$</td>
</tr>
<tr>
<td>$N_{\text{sig}}^{\text{exp}}$</td>
<td>1 ± 6</td>
<td>1 ± 6</td>
<td>2 ± 10</td>
</tr>
<tr>
<td>$N_{\text{sig}}$</td>
<td>5 ± 7</td>
<td>3 ± 8</td>
<td>$-10 \pm 11$</td>
</tr>
</tbody>
</table>

Figure 8.1: Posterior PDF for $\mathcal{B}(B^\pm \to K^{\pm}\nu\bar{\nu})$ on the left and $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$ on the right: the green dark (light) area shows the 90% (95%) CL region. The results are for the hadronic Neural Network analysis.

to parameters that enter the NP under exams. It has been shown how the Effective Hamiltonian formalism allows to decouple SM and NP effects. In particular, writing the decay amplitude for a $B$ exclusive decay as in Eq. 1.26:

$$\mathcal{A}(B \to F) = \langle F | \mathcal{H}_{\text{eff}} | B \rangle = \sum_i B_i V_{\text{CKM}}^{ij} \eta^{ij}_{\text{QCD}}(\mu, \mu_W) F_i(x_i),$$

the non-SM contributions can show up as new complex phases in $V_{\text{CKM}}^{ij}$, new terms in the sum not proportional to the CKM elements involved in the SM amplitude, new operators $B_i$ or contributions to the Inami-Lim functions $F_i$ [21]. The current experimental knowledges can be exploited to constraint parameters defining one of the previous contributions.

Two example are discussed in the following: constraint on the Inami-Lim function in the
8.4 Phenomenological constraints from $B \to K^{(*)} \nu \bar{\nu}$

The starting point is the SM Effective Hamiltonian for $b \to s \nu \bar{\nu}$, defined in Sec 1.3 and here recalled:

$$H_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{ts}^* V_{tb} C_L Q_L^
u,$$

where the SM Wilson coefficient $C_L$ and operator $Q_L$ are respectively given by:

$$C_L^\nu|_{\text{SM}} = \frac{4B(x_t) - C(x_t)}{\sin^2 \theta_W},$$

$$Q_L^\nu = \frac{e^2}{4\pi^2} \bar{s}_L \gamma_\mu b_L \gamma^\mu (1 - \gamma^5) \nu.$$  \hfill (8.15)

The two New Physics models under discussion can give:

- an additive term to $C(x_t)$, $\Delta C$, due to non-standard $Z^0$ vertices that would modify the Wilson coefficient as:

$$C_L^\nu|_{\text{SM}} = \frac{4B(x_t) - (C(x_t) + \Delta C)}{\sin^2 \theta_W}.$$  \hfill (8.16)
Since the box is not affected from new couplings, the $B(x_t)$ is expected to be unchanged. In addition, choosing a Minimal flavor violation framework [51], in which all the NP effects can enter only through corrections to the SM Wilson coefficients, no right-handed operators $Q^R_L$ are allowed to enter the effective Hamiltonian: as a consequence $C^v_R$ is considered to be zero;

- a new operator that accounts for the presence of the light dark matter in the final state [27] and which act as an additive term to the decay amplitude.

The procedure adopted to set the constraints consists of several steps:

- compute the SM values of the quantities entering the experimental observable (in the case of the $B \to K^{(*)}\nu\bar{\nu}$ branching fractions, the SM Inami-Lim function should be evaluated);

- add the NP contribution to the SM parameters ;

- evaluate the theoretical expectation for the experimental observables using SM and NP terms;

- compare the theoretical expectation to the experimental results to constraint the NP parameters.

In Tab. 8.2 a list of the physical parameters used in the phenomenological analysis is presented, while Tab. 8.3 shows the experimental measurements of the BF that have been exploited.

To validate the analysis procedure, the Standard Model predictions for the $B \to K^{(*)}\nu\bar{\nu}$ branching fractions have been computed. They are calculated starting from the differential decay rate of Eqs. 1.34 and 1.37, here recalled:

$$
\frac{d\Gamma(B \to K\nu\bar{\nu})}{ds} = \frac{G_F^2 \alpha^2 m_B^5}{256 \pi^5} |V_{ts} V_{tb}|^2 \lambda^{3/2}_K(s) f_+(s) |C'_L + C'_R|^2 ,
$$

$$
\frac{d\Gamma(B \to K^{*}\nu\bar{\nu})}{ds} = \frac{G_F^2 \alpha_s^2 m_B^5}{1024 \pi^5} |V_{ts} V_{tb}|^2 \lambda^{1/2}_K(s) \left\{ \frac{8 s \lambda_{K^{*}}(s) V^2(s)}{(1 + \sqrt{r_{K^{*}}})^2} |C'^v_L + C'^v_R|^2 \
+ \frac{1}{r_{K^{*}}} \left[ (1 + \sqrt{r_{K^{*}}})^2 (\lambda_{K^{*}}(s) + 12 r_{K^{*}} s) A_1(s)^2 + \frac{\lambda_{K^{*}}(s) A_3(s)}{(1 + \sqrt{r_{K^{*}}})^2} \right] \right\} \left| C^v_L - C^v_R \right|^2 \right\} .
$$
Table 8.2: Parameters and their values used in the phenomenological analysis [4].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{W^\pm}$ (GeV/c^2)</td>
<td>80.425</td>
</tr>
<tr>
<td>$m_{Z^0}$ (GeV/c^2)</td>
<td>91.1867</td>
</tr>
<tr>
<td>$m_{B^+}$ (GeV/c^2)</td>
<td>5.279</td>
</tr>
<tr>
<td>$m_{B^0}$ (GeV/c^2)</td>
<td>5.2794</td>
</tr>
<tr>
<td>$m_{K^*}$ (GeV/c^2)</td>
<td>0.892</td>
</tr>
<tr>
<td>$m_K$ (GeV/c^2)</td>
<td>0.494</td>
</tr>
<tr>
<td>$m_t$ (GeV/c^2)</td>
<td>161.2 ± 1.7</td>
</tr>
<tr>
<td>$m_b$ (GeV/c^2)</td>
<td>4.21 ± 0.08</td>
</tr>
<tr>
<td>$m_c$ (GeV/c^2)</td>
<td>1.3 ± 0.1</td>
</tr>
<tr>
<td>$m_s$ (GeV/c^2)</td>
<td>0.95 ± 0.25</td>
</tr>
<tr>
<td>$\mu_0$ (GeV)</td>
<td>80.425</td>
</tr>
<tr>
<td>$\mu$ (GeV)</td>
<td>4.6</td>
</tr>
<tr>
<td>$\tau_{B^+}$ (ps)</td>
<td>1.638</td>
</tr>
<tr>
<td>$\tau_{B^0}$ (ps)</td>
<td>1.530</td>
</tr>
<tr>
<td>$G_F$ (GeV^{-2})</td>
<td>$1.16637 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1/137</td>
</tr>
<tr>
<td>$\alpha_s$</td>
<td>0.199 ± 0.003</td>
</tr>
<tr>
<td>$\bar{\rho}$</td>
<td>0.167 ± 0.051</td>
</tr>
<tr>
<td>$\bar{\eta}$</td>
<td>0.386 ± 0.035</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.2258 ± 0.0014</td>
</tr>
<tr>
<td>$\sin \theta_W$</td>
<td>0.23154</td>
</tr>
</tbody>
</table>

Table 8.3: Measurement of $B \to K^{(*)}\nu\bar{\nu}$ exploited in the phenomenological analysis.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Lumi. (fb^-1)</th>
<th>Normalization</th>
<th>Signal Efficiency</th>
<th>$N_{exp}^{bkg}$</th>
<th>$N_{obs}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \to K^+\nu\bar{\nu}$ [31]</td>
<td>492</td>
<td>(535 ± 7) $\times 10^6$</td>
<td>$26.7 \pm 2.9 \times 10^{-5}$</td>
<td>20.0 ± 4.0</td>
<td>10</td>
</tr>
<tr>
<td>$B^+ \to K^+\nu\bar{\nu}$ [33]</td>
<td>319</td>
<td>(351 ± 4) $\times 10^6$</td>
<td>$1.64 \pm 0.22 \times 10^{-3}$</td>
<td>30.7 ± 0.7</td>
<td>38</td>
</tr>
<tr>
<td>$B^0 \to K^0\nu\bar{\nu}$ [31]</td>
<td>492</td>
<td>(535 ± 7) $\times 10^6$</td>
<td>$5.0 \pm 0.3 \times 10^{-5}$</td>
<td>2.0 ± 0.9</td>
<td>2</td>
</tr>
<tr>
<td>$B^+ \to K^{*+}\nu\bar{\nu}$ [31]</td>
<td>492</td>
<td>(535 ± 7) $\times 10^6$</td>
<td>$5.8 \pm 0.7 \times 10^{-5}$</td>
<td>5.6 ± 1.8</td>
<td>4</td>
</tr>
<tr>
<td>$B^0 \to K^{*0}\nu\bar{\nu}$ [31]</td>
<td>492</td>
<td>(535 ± 7) $\times 10^6$</td>
<td>$5.1 \pm 0.3 \times 10^{-5}$</td>
<td>4.2 ± 1.4</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>Lumi. (fb^-1)</th>
<th>Normalization</th>
<th>Signal Efficiency</th>
<th>$N_{sig}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \to K^{*+}\nu\bar{\nu}$</td>
<td></td>
<td>(454 ± 5) $\times 10^6$</td>
<td>$5.6 \pm 0.7 \times 10^{-4}$</td>
<td>-22 ± 21</td>
</tr>
<tr>
<td>$K^{*+} \to K^+\pi^0$ (SL) [37]</td>
<td>413</td>
<td>(454 ± 5) $\times 10^6$</td>
<td>$6.7 \pm 0.6 \times 10^{-2}$</td>
<td>5 ± 7</td>
</tr>
<tr>
<td>$K^{*+} \to K^0\pi^+$ (HAD) [37]</td>
<td>413</td>
<td>(454 ± 5) $\times 10^6$</td>
<td>$4.3 \pm 0.6 \times 10^{-4}$</td>
<td>3 ± 23</td>
</tr>
<tr>
<td>$K^{*+} \to K^0\pi^+$ (SL) [37]</td>
<td>413</td>
<td>(454 ± 5) $\times 10^6$</td>
<td>$6.1 \pm 0.7 \times 10^{-2}$</td>
<td>3 ± 8</td>
</tr>
<tr>
<td>$B^0 \to K^{*0}\nu\bar{\nu}$</td>
<td></td>
<td>(8.9 ± 0.4) $\times 10^6$</td>
<td>$9.2 \pm 0.8 \times 10^{-4}$</td>
<td>35 ± 16</td>
</tr>
<tr>
<td>$K^{*0} \to K^+\pi^-$ (HAD) [37]</td>
<td>413</td>
<td>(6.3 ± 0.3) $\times 10^6$</td>
<td>$19.2 \pm 1.6 \times 10^{-2}$</td>
<td>-10 ± 11</td>
</tr>
</tbody>
</table>
with the coefficient $C_R$ assumed to be zero in the SM. The adopted parametrization for the form factors ($f_+(s), V(s), A_1(s),$ and $A_2(s)$ in the previous equations) is:

$$f(s) = f(0)e^{c_1s+c_2s^2+c_3s^3}$$  \hspace{1cm} (8.17)

where the parameter values are computed in Ref. [52]. A Bayesian approach consisting of three steps, is used to evaluated the expected SM branching fractions:

- the inputs defining the $B \to K\nu\pi$ and $B \to K^*\nu\pi$ dilepton spectrums are treated as follow: the experimental parameters (Tab. 8.2) are generated according to a Gaussian distribution with mean and width corresponding to the measured central value and error, respectively; the parameters entering the form factor are generated according to a uniform distribution in the ranges extracted in Ref. [52];

- for each parameter generation, the value of $d\Gamma/ds$ is computed and an integration over the $s$ physical range is performed;

- the resulting BF is computed.

The previous steps are iterated for a given number of experiments and the SM predictions for the isospin averaged $B \to K\nu\pi$ and $B \to K^*\nu\pi$ branching fractions are obtained. The distributions of the branching fraction posterior are shown in Fig. 8.3 and the allowed ranges are summarized in Tab. 8.4. The isospin average consists on averaging the charged and neutral channels which, in absence of a dynamical isospin asymmetry in the electroweak interaction between the $b$ and a lighter quark, differ only for the $B$ meson lifetime. The resulting central values at 68% Confidence Level are:

$$BR(B \to K\nu\pi) = (0.32 \pm 0.08) \times 10^{-5},$$

$$BR(B \to K^*\nu\pi) = (1.1 \pm 0.3) \times 10^{-5}.$$  \hspace{1cm} (8.18)

They are in agreement with the published SM expectations [23]: $\mathcal{B}(B \to K\nu\pi) = (3.8^{+1.2}_{-0.6}) \times 10^{-6}$ and $\mathcal{B}(B \to K^*\nu\pi) = (1.3^{+0.4}_{-0.3}) \times 10^{-5}$. Eq. 8.19 differs from the previous values for to the usage of most recent measurements of input parameters and a different treatment of the uncertainties on the form factors.
Figure 8.3: Standard Model prediction for $B \rightarrow K^* \nu \bar{\nu}$ (left) and $B \rightarrow K \nu \bar{\nu}$ (right): the yellow (red) area shows the 68% (95%) CL region.

Table 8.4: SM expectations for $B(B \rightarrow K \nu \bar{\nu})$ and $B(B \rightarrow K^* \nu \bar{\nu})$.

<table>
<thead>
<tr>
<th></th>
<th>$B(B \rightarrow K \nu \bar{\nu})(\times 10^{-6})$</th>
<th>$B(B \rightarrow K^* \nu \bar{\nu})(\times 10^{-6})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>68% CL range</td>
<td>[2.4, 4.0]</td>
<td>[7.7, 14.2]</td>
</tr>
<tr>
<td>95% CL range</td>
<td>[1.8, 5.1]</td>
<td>[5.6, 18.3]</td>
</tr>
</tbody>
</table>
Constraints on New Physics parameters

As anticipated before, in presence of non-standard $Z^0$ couplings, the Inami-Lim function describing the penguin diagram is modified. Other than the SM expectation, an additive term containing the extra-factor $\Delta C$ should be accounted for. The measured quantities exploited to constraint $\Delta C$ are: the number of observed event in a counting approach or the fitted number of signal events in case of a maximum likelihood fit (Tab. 8.3). In the event counting analysis the likelihood of $N_{\text{obs}}$ takes the form of a Poisson distribution, defined as:

$$L(N_{\text{th}}, N_{\text{obs}}) = \frac{(N_{\text{th}})^{N_{\text{obs}}} e^{N_{\text{th}}}}{N_{\text{obs}}!}$$  \hspace{1cm} (8.20)

being $N_{\text{th}} = B_{\text{th}} \cdot \epsilon_{\text{sig}} \cdot N_B + N_{\text{bkg}}$ where $B_{\text{th}}$ is the theoretical BF accounting for SM and NP terms, while $\epsilon_{\text{sig}} \cdot N_B$, and $N_{\text{bkg}}$ are the measured signal efficiency and normalization factor and the expected background yield respectively. In the fit analysis, the $N_{\text{sig}}$ likelihood is a Gaussian centered in the fitted value and whose width corresponds to the sum in quadrature of the statistical and systematic errors (the gaussianity of the likelihood is ensured from the toy MC studies discussed in Sec. 6.2.3).

The steps to derive the $\Delta C$ constraints are the following:

- the value $\Delta C$ is generated according to a flat distribution in the range $[-10, 10]$;
- the SM expectation for $d\Gamma/ds$ is computed as specified before and similarly for the NP terms;
- the total dilepton spectrum (standard plus non-standard contribution) is calculated and integrating over the whole $s$-range the BF is obtained;
- the expected value for the experimental observable ($N_{\text{sig}}$ or $N_{\text{obs}}$) is computed by using the total BF as defined in the previous step;
- the likelihood associated to the experimental observable is evaluated in correspondence of its expectation value;
- such likelihood value is assigned as weight to the generated $\Delta C$;

The procedure is iterated and the weighted distribution of the generated $\Delta C$ gives the posterior PDF (left plot in Fig. 8.4), from which the 68% and 95% CL region for the
parameter range are set:

\begin{align}
68\% \text{ CL range} & : \ [-3.6, -1.4] U [-0.6, 1, 2] \quad (8.21) \\
95\% \text{ CL range} & : \ [-4.7, 1.9] \quad (8.22)
\end{align}

The result is obtained by combining \( B \to K \nu \bar{\nu} \) and \( B \to K^* \nu \bar{\nu} \) measurements: comparing the plot on the left (which include the \( B \to K^* \nu \bar{\nu} \) BABAR measurement) and the plot on the right (which do not) the impact of the measurement presented in this thesis can be appreciated.

The second NP model investigated predicts the existence of light dark matter which can act as invisible energy in \( B \) final state with a \( K^* \). Note that the \( B \to K \nu \bar{\nu} \) measurements and the Belle \( B \to K^* \nu \bar{\nu} \) result are not exploited since the kinematic properties of the kaon enter the selection and the analysis can not be considered as model independent.

The SM branching fraction obtained by integrating Eq. 1.37, should be summed to the BF obtained by integrating the following decay rate (Eq. 1.42):

\[
\frac{d\Gamma(B \to K^* \nu \bar{\nu})}{d\hat{s}} = \frac{x_t^2 C_{DM}^2 A_{0}(\hat{s})^2 I_{K^*}(\hat{s}m_S) h(\hat{s})}{512\pi^3 m_B},
\]

over the physical range of \( \hat{s} = (p_B - p_{K^*})^2 \). Recalling that \( C_{DM} \) is defined as:

\[
C_{DM} = \frac{\lambda}{m_h^2} \frac{3g_W^2 V_{ts} V_{tb}}{32\pi^2} x_t,
\]
there are three parameters that define the dark-matter branching fraction: the coupling between the Higgs and the dark matter field $\lambda$, the Higgs mass $m_h$, and the scalar WIMP mass $m_S$. The first two can be combined to define $k \propto \lambda/m_h^2$; as discussed in Ref. [27] such quantity is related to the computation of dark matter abundance in the universe and, for $m_S \sim O$(some GeV/$c^2$), it should be $k \sim O(1)$. In the light of these considerations, the ranges for $k$ is chosen to be $[0, 2]$, while $m_S$ is constrained, from kinematical arguments, in the range $[0, (m_B - m_K^-)^2]$. The constraints on the $m_S - k$ plane are set with the same procedure adopted for $\Delta C$ but in this case the parameters to be generated with flat distribution are $m_S$ and $k$. The results of the analysis are shown in Fig. 8.5. For large values of $k$, small values of the $S$ mass are excluded: in fact, considering that the decay rate is proportional to $k^2/m_S$, large $k$ and small $m_S$ would lead to high branching fraction; for small $k$ is difficult to exclude any $m_S$ range since the smallness of the BF could be due both to a high $S$ mass or to the low value of the coupling which define $k$.

Figure 8.5: Constraint on the $m_S$-$k$ plane from the BABAR $B \to K^*\nu\bar{\nu}$ result: the continuous (dashed) line represent the lower limit of the 68\% (95\%) probability region.
Conclusions

The first search for the $B \to K^* \nu \bar{\nu}$ decay at the BABAR experiment has been presented. The channel under investigation, a Flavor Changing Neutral Current process, is of particular interest: it is prohibited at tree level in the Standard Model and the final state can be reached by a $B$ meson only through higher order diagrams. As a consequence New Physics can show up in terms of new particles entering loop and box diagrams or non-standard invisible states that would be responsible for the missing energy: this can affect the experimental measurement by enhancing the branching fraction of some order of magnitude.

The analyzed data sample consists of 454 million $B\bar{B}$ pair collected at the $\Upsilon(4S)$ resonance with the BABAR detector at the SLAC PEP-II $B$ Factory. The experimental method used to reconstruct the signal decay is called “recoil technique”: one $B$ meson decay is fully reconstructed in hadronic final states as $B \to D^{(*)} X$, $X$ being a combination of pions and kaons. On the other hemisphere of the event a charged or neutral $K^*$ associated to missing energy is searched for. The following $K^*$ decay modes have been reconstructed: $K^{*+} \to K^+ \pi^0$, $K^{*+} \to K_S^0 \pi^+$, and $K^{*0} \to K^+ \pi^-$. Two different approaches have been exploited to discriminate signal from background events: a cut and count analysis and an extended maximum likelihood fit to the distribution of a Neural Network output. Whit both the methods no evidence for signal has been found. The second one proved to be better: the loose selection on the discriminant variables allows to maintain a larger signal statistics and the larger background contamination is compensated by the better discriminating power of the Neural Network fit with respect to the cut and count analysis. The following values for the signal efficiency has been found: $(5.8 \pm 0.5) \times 10^{-2}$ for $K^{*+} \to K^+ \pi^0$, $(5.2 \pm 0.6) \times 10^{-2}$ for $K^{*+} \to K_S^0 \pi^+$, and $(16.6 \pm 1.4) \times 10^{-2}$ for $K^{*0} \to K^+ \pi^-$. The main background contamination comes from $B\bar{B}$ events in which the kaon and the pion used to reconstruct the $K^*$ are randomly paired or come from a $K^*$ on the hadronic
B side. Moreover events in which a B decays semileptonically, with a lepton daughter lying outside the detector acceptance, are a source of irreducible background.

The number of observed events is in agreement with the background expectation. In the cut and count analysis the background event estimation is $18.8 \pm 12.6$ for $K^{*+} \to K^+\pi^0$, $8.5 \pm 6.1$ for $K^{*+} \to K^0_\pi \pi^+$, and $14.0 \pm 7.4$ for $K^{*0} \to K^+\pi^-$ to which corresponds a selected data sample of 20, 11, and 19 events, respectively. Also in the Neural Network fit, the signal yield estimation is consistent with 0 for all $K^*$ decay modes.

The Neural Network approach allows to set the following upper limits at 90% confidence level:

$$B(B^\pm \to K^{*\pm} \nu \overline{\nu}) < 2.1 \times 10^{-4},$$
$$B(B^0 \to K^{*0} \nu \overline{\nu}) < 1.1 \times 10^{-4}.$$  

The present work has been combined with an analysis in which to reconstruct the $B$ accompanying the signal candidate, in place of hadronic decays, semileptonic final states are considered. The resulting upper limits, consistent with the Standard Model expectations [23], represent the most precise measurement on the $B \to K^* \nu \overline{\nu}$ search reported to date:

$$B(B^\pm \to K^{*\pm} \nu \overline{\nu}) < 0.8 \times 10^{-4},$$
$$B(B^0 \to K^{*0} \nu \overline{\nu}) < 1.2 \times 10^{-4},$$
$$B(B \to K^* \nu \overline{\nu}) < 0.8 \times 10^{-4}.$$  

In addition they are completely model-independent, since no assumptions have been made on the kinematics of the signal decay, and can be interpreted as upper limits on $B \to K^* + invisible$.

The experimental measurements have been used to constrain parameters defining particular New Physics scenarios, which predict, for example, non-standard $Z^0$ couplings in a Minimal Flavor Violation framework or the presence of light dark matter which can replace the neutrino pair in the final state.

The analysis has been recently published by Physical Review D [37].

Regarding the future prospects, the $B \to K^* \nu \overline{\nu}$ decay will be one of the channels to be investigated in a future Super Flavor Factory [53], where the higher luminosities and possible improvements on the detector hermeticity and on the analysis strategy will make the experimental sensitivity approach the Standard Model prediction.
Appendix A

$B$ decay product reconstruction

In this appendix, the reconstruction of charged tracks and neutral clusters is described (Secs. A.1). Such objects are used to identify and select composite particles for the $B_{\text{had}}$ and $B_{\text{sig}}$ reconstruction. In Sec. A.2 the criteria applied to the $B_{\text{had}}$ decay products are listed. Different requirements are adopted for kaons and pions of the signal side, as presented in Secs. 5.1-5.2.

A.1 Track and Cluster reconstruction

In the two following sections the selection of charged tracks and neutral clusters are described. The particle identification algorithm used to identify the charged tracks as pion or kaons is also discussed. Control samples have been used to produce the plots of this section which do not correspond to the final sample used in the present analysis.

A.1.1 Charged Track selection

The charged track selection is based on information coming from both tracking devices: SVT and DCH. A track is defined by a set of five parameters ($d_0$, $z_0$, $\phi_0$, $\lambda$, $1/p_T$), and the associated error matrix, which is measured at the point of closest approach to the $z$-axis; $d_0$ and $z_0$ are the distance between the point of closest approach to the $z$-axis and the origin of the coordinate system in the $x$-$y$ plane and along the $z$-axis respectively; $\phi_0$ is the azimuth of the track, $\lambda$ the dip angle relative to the transverse plain and $p_T$ the measured transverse momentum. A reconstructed track has to satisfy the following selection requirements:

- $d_0 < 1.5$ cm and $z_0 < 10$ cm, to veto fake tracks and cosmic muons;
• $N$ (DCH hits) $> 0$ for high momentum tracks ($p_T > 0.2$ GeV/$c$), where $N$ (DCH hits) is the number of hits in the drift chamber; this request is not applied to low momentum tracks since they do not reach the DCH at all;

• $p_{lab} < 10$ GeV/$c$, $p_{lab}$ being the momentum in the laboratory frame, to remove tracks not compatible with the beam energy;

• $0.41 < \theta_{lab} < 2.54$, where $\theta_{lab}$ is the polar angle in the laboratory frame.

A set of combined requirements on $\Delta \phi$, $\Delta p_T$, and $\Delta \theta$ is applied to each pair of reconstructed tracks to identify those which do not reach the EMC and therefore spiral inside the DCH (“loopers”). A dedicated selection has been developed to reject track fragments compatible with originating from loopers based on their distance from the beam spot.

Two tracks very closely aligned to each other are called “ghosts”. These cases arise when the pattern recognition algorithm splits the DCH hits belonging to a single track in two track fragments. In this case only the track with the largest number of DCH hits is retained.

A summary of the track selection criteria is shown in Tab. A.1.

Table A.1: Summary of track selection criteria.

<table>
<thead>
<tr>
<th>Track Selection</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance in $x$-$y$ plane</td>
<td>$</td>
</tr>
<tr>
<td>distance in $z$-axis</td>
<td>$</td>
</tr>
<tr>
<td>minimum number of DCH hits</td>
<td>$N_{DCH} &gt; 0$ if $p_T &gt; 0.2$ GeV</td>
</tr>
<tr>
<td>maximum momentum</td>
<td>$p_{lab} &lt; 10$ GeV/$c$</td>
</tr>
<tr>
<td>geometrical acceptance</td>
<td>$0.410 &lt; \theta_{lab} &lt; 2.54$ rad</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rejection of loopers and ghosts</th>
<th>Selection criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>loopers ($p_T &lt; 0.25$ GeV)</td>
<td>$\Delta p_t &lt; 120$ MeV from other tracks</td>
</tr>
<tr>
<td></td>
<td>Same sign: $</td>
</tr>
<tr>
<td></td>
<td>Opposite sign: $</td>
</tr>
<tr>
<td>ghosts ($p_T &lt; 0.35$ GeV)</td>
<td>$\Delta p_t &lt; 150$ MeV from other tracks</td>
</tr>
<tr>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

Charged hadron identification

Charged tracks are identified as kaons and pions by examining the loss of energy $dE/dx$ measured by the SVT and the DCH and the Čerenkov angle provided by the DIRC. For
A.1 Track and Cluster reconstruction

Each particle mass hypothesis \((e, \mu, \pi, K, \text{and } p)\), a likelihood function is defined, by using the quantities measured by the three different sub-detectors which can be considered as uncorrelated. The kaon and pion selectors evaluates the likelihood of a particle to be a kaon \(L_K\), a pion \(L_\pi\) and a proton \(L_p\); the likelihood ratios \(L_K/L_\pi\) and \(L_K/L_p\) verify the agreement between \(dE/dx\) and the Čerenkov angle measurements with the expected values for a given mass hypothesis. Fig. A.1 shows good discriminating power for momenta up to 700 MeV/c for the DCH and up to 600 MeV/c for the SVT; at higher momenta, the Čerenkov angle dominates the particle identification (Fig. A.2). Two particle selector modes are defined: a loose mode in which kaon efficiency is maximized and a tight one in which the pion mis-identification is minimized. Averaging over the track momentum, the kaon efficiency is 78% (58%) and the pion mis-identification is 4.5% (0.6%) for the loose (tight) selection.

Figure A.1: Energy loss for ionization as a function of the momentum in the SVT (left) and in the DCH (right) for different kinds of particles. The curves show the expected behaviors.

A.1.2 Photon selection

Neutral particles such as photon \((\pi^0)\) are reconstructed as (pairs of) clusters interacting with the EMC volume that do not match with tracks extrapolated from the tracker system to the EMC inner surface. The EMC clusters mostly originate from photons and momenta and angles are assigned to be consistent with a \(\gamma\) coming from the interaction region. A
lower bound of 50 MeV on the photon energy ($E_\gamma$) is imposed, in order to reduce beam-related background due to low energy photon. Requirements on the cluster shape are applied to suppress contamination from hadronic interactions (i.e. $K^0_L$ or neutrons). The variable that allows the discrimination between hadronic and electromagnetic showers is defined as:

$$LAT = \frac{\sum_{i=3}^{N} E_i r_i^2}{\sum_{i=3}^{N} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}$$ (A.1)

where $N$ is the number of crystals associated to the electromagnetic shower, $r_0$ is the mean distance between two crystals in the $\text{BABAR}$ EMC (about 5 cm), $E_i$ the energy associated to the $i$-th crystal, ordering them such that $E_1 > E_2 > E_3 > .. > E_N$, and $r_i$ is the polar radius in the plane perpendicular to the line pointing from the interaction point to the shower center. The summation in Eq. A.1 does not take into account the first two crystals accounting for the highest amount of energy. For electromagnetic showers most of the energy is deposited in two crystals while hadronic showers involves a higher number of crystals, so the formers have a small LAT value with respect to the latter. The following requirement is then applied: $LAT < 0.8$.

The second shape variable used in the selection is called $S9/S25$ and is defined as the ratio of the energy deposit in the 9 crystals closest to the cluster centroid and the energy deposit in the 25 closest clusters. $S9/S25$ is required to be greater than 0.9.

Clusters, which are considered as neutral candidates, although they are close to tracks,
need to be rejected. For this purpose the angle between the position of the cluster and the impact point of the nearest charged track at the EMC surface is exploited. The 3-D angular difference $\Delta \alpha$ is defined as:

$$\Delta \alpha = \cos^{-1} [\cos \theta_{cl} \cos \theta_{tr} + \sin \theta_{cl} \sin \theta_{tr} \cos(\phi_{cl} - \phi_{tr})]$$  \hspace{1cm} (A.2)

where $\theta_{cl, tr}$ and $\phi_{cl, tr}$ are the polar coordinates for cluster and track impact point on the EMC surface respectively. $\Delta \alpha$ is required to be $> 0.08$, considering the nearest track not identified as an electron.

Tab. A.2 shows a detailed summary of the selection criteria applied on the neutral candidates.

<table>
<thead>
<tr>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral energy</td>
</tr>
<tr>
<td>LAT</td>
</tr>
<tr>
<td>S9S25</td>
</tr>
<tr>
<td>unmatched clusters</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

**A.2 Composite particle reconstruction**

Clusters and charged kaon and pion candidates are used to reconstruct composite particles identified as $B_{\text{had}}$ daughters. Also the plots in this section are produced using control samples which do not correspond to the final sample used in the present analysis.

**A.2.1 Selection of $\pi^0$ mesons**

Neutral pions are identified by pairs of neutral clusters with minimum energy of 30 MeV. The energy of the resulting $\pi^0$ has to be above 200 MeV. The two photon invariant mass has to lie between 110-155 MeV/$c^2$, corresponding to a $(-4\sigma, +3\sigma)$ cut around the nominal value. In Fig.A.3 the invariant mass distributions for simulated events and data are shown.
A.2.2 Selection of $K_s^0$ mesons

The $K_s^0$ candidates are reconstructed in the decay $K_s^0 \rightarrow \pi^+\pi^-$. All possible charged tracks, geometrically constrained to come from a common vertex, are paired to form the kaon candidate. The vertex fit is then performed using the GEOKIN algorithm. It consists on a $\chi^2$ minimization and uses, as constraining point for the vertex finding, the three dimensional point of closest approach to the $z$-axis. A $\pm 3\sigma$ cut around the $K_s^0$ nominal mass is imposed: $0.486 < m_{\pi^+\pi^-} < 0.510$ GeV/$c^2$. The invariant mass distribution of the $\pi^+\pi^-$ obtained from data is shown in Fig.A.4. A comparison between data and Monte Carlo for the $K_s^0$ momentum and polar angle is shown in Fig.A.5.

A.2.3 Selection of charmed mesons

Selection of $D$ mesons

The $D$ meson candidates are reconstructed by combinations of neutral and charged $K$ and $\pi$. The selections of $D^0$ and $D^+$ candidates are summarized in Tab. A.3 and Tab. A.4 respectively. Requirements on the $D$ meson reconstructed mass, the fit vertex probability, the $D$ meson momentum in the CM frame, the momentum of tracks originating from the $D$ are applied. In the $D^0 \rightarrow K^-\pi^+\pi^0$ selection (Tab. A.3) also the angle between the $K^-$ and the $\pi^+$ in the CM frame is exploited. Moreover the $\pi^+\pi^0$ invariant mass is required.
Figure A.4: Mass distributions for $K_S^0 \rightarrow \pi^+\pi^-$ on data. A fit to a sum of a double Gaussian and a first order polynomial function is performed.

Figure A.5: $K_S^0$ momentum (left) and polar angle (right) distributions in data (solid markers) and Monte Carlo simulation (hatched histogram), normalized to the same area.
to lie in the region $m(\rho) \pm 150 \text{ MeV}/c^2$, where $m(\rho) = 775.4 \text{ MeV}/c^2$ [4]. For each $D$ mode, the reconstructed $D$ mass must lie within $\pm 2\sigma$ of the nominal $D$ mass [4], where $\sigma$ is calculated by fitting to a Gaussian the reconstructed mass in simulated signal events.

Table A.3: Selection requirements applied in the $D^0$ reconstruction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$ vertex fit probability</td>
<td>$p(D^0) (D^0$ from $D^*)$</td>
</tr>
<tr>
<td>$D^0 \rightarrow K^-\pi^+$</td>
<td>$m(D^0) \pm 15 \text{ MeV}/c^2$</td>
</tr>
<tr>
<td>$p(K^-), p(\pi^+)$</td>
<td>$&gt; 200 \text{ MeV}/c$</td>
</tr>
<tr>
<td>$m(K^-\pi^+\pi^0)$</td>
<td>$m(D^0) \pm 25 \text{ MeV}/c^2$</td>
</tr>
<tr>
<td>$p(K^-), p(\pi^+)$</td>
<td>$&gt; 150 \text{ MeV}/c$</td>
</tr>
<tr>
<td>$m(\pi^+\pi^0)$</td>
<td>$m(\rho) \pm 15 \text{ MeV}/c^2$</td>
</tr>
<tr>
<td>$</td>
<td>\cos(\theta^*_{K^-\pi^+})</td>
</tr>
<tr>
<td>$D^0 \rightarrow K^0\pi^+\pi^-$</td>
<td>$m(D^0) \pm 20 \text{ MeV}/c^2$</td>
</tr>
<tr>
<td>$p(\pi^-)$</td>
<td>$&gt; 150 \text{ MeV}/c$</td>
</tr>
<tr>
<td>$D^0 \rightarrow K^-\pi^+\pi^-\pi^+$</td>
<td>$m(D^0) \pm 15 \text{ MeV}/c^2$</td>
</tr>
<tr>
<td>$p(K^-), p(\pi^+)$</td>
<td>$&gt; 150 \text{ MeV}/c$</td>
</tr>
</tbody>
</table>

Selection of $D^*$ mesons

Neutral $D^*$ candidates are formed by combining a selected $D^0$ with either a $\pi^0$ or a photon. The selection requirements, imposed on the difference between the reconstructed $D^0$ mass and the $D^0-\pi^0$ invariant mass, and on the $\pi^0$ and $D^*0$ momenta in the CM frame are summarized in Tab. A.5. The distribution of $\Delta m = m(D^0\pi^0/\gamma) - m(D^0)$ is fitted on the corresponding MC sample to the sum of two Gaussian functions. The weighted average of the width of the two Gaussian’s, multiplied by a factor 3, is taken as width of the $\Delta m$ window. The central value of the window is computed by using the $D^{*0}$ and $D^0$ nominal masses.

Charged $D^*$ candidates are originated by a $D^0$ and a charged pion (the $D^{*+} \rightarrow D^+\pi^0$ events enter the $B^0 \rightarrow D^+X$ category). To improve the angular resolution of the pion, called “soft” due to its low momentum ($p < 450 \text{ MeV}/c$), a vertex fit to the $D^{*+}$ is performed. The convergence of the fit is required, but no selection on the $\chi^2$ vertex prob-
Table A.4: Selection requirements applied in the $D^+$ reconstruction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$ vertex fit probability</td>
<td>$&gt; 0.01$</td>
</tr>
<tr>
<td>$p^*(D^+)$</td>
<td>$&lt; 2.5$ GeV/$c$</td>
</tr>
<tr>
<td>$D^+ \to K^-\pi^+\pi^+$</td>
<td></td>
</tr>
<tr>
<td>$m(K^-\pi^+\pi^+)$</td>
<td>$m(D^+) \pm 20$ MeV/$c^2$</td>
</tr>
<tr>
<td>$p^*(D^+)$</td>
<td>$&gt; 1.0$ GeV/$c$</td>
</tr>
<tr>
<td>$p(K^-)$</td>
<td>$&gt; 200$ MeV/$c$</td>
</tr>
<tr>
<td>$p(\pi^+)$</td>
<td>$&gt; 150$ MeV/$c$</td>
</tr>
<tr>
<td>$D^+ \to K^-\pi^+\pi^+\pi^0$</td>
<td></td>
</tr>
<tr>
<td>$m(K^-\pi^+\pi^+\pi^0)$</td>
<td>$m(D^+) \pm 30$ MeV/$c^2$</td>
</tr>
<tr>
<td>$p^*(D^+)$</td>
<td>$&gt; 1.6$ GeV/$c$</td>
</tr>
<tr>
<td>$p(K^-)$</td>
<td>$&gt; 200$ MeV/$c$</td>
</tr>
<tr>
<td>$p(\pi^+)$</td>
<td>$&gt; 150$ MeV/$c$</td>
</tr>
<tr>
<td>$D^+ \to K_S^{0}\pi^+$</td>
<td></td>
</tr>
<tr>
<td>$m(K_S^{0}\pi^+)$</td>
<td>$m(D^+) \pm 20$ MeV/$c^2$</td>
</tr>
<tr>
<td>$p^*(D^+)$</td>
<td>$&gt; 1.0$ GeV/$c$</td>
</tr>
<tr>
<td>$p(\pi^+)$</td>
<td>$&gt; 200$ MeV/$c$</td>
</tr>
<tr>
<td>$D^+ \to K_S^{0}\pi^+\pi^0$</td>
<td></td>
</tr>
<tr>
<td>$m(K_S^{0}\pi^+\pi^0)$</td>
<td>$m(D^+) \pm 30$ MeV/$c^2$</td>
</tr>
<tr>
<td>$p^*(D^+)$</td>
<td>$&gt; 1.0$ GeV/$c$</td>
</tr>
<tr>
<td>$p(\pi^+)$</td>
<td>$&gt; 200$ MeV/$c$</td>
</tr>
<tr>
<td>$D^+ \to K_S^{0}\pi^+\pi^-\pi^+$</td>
<td></td>
</tr>
<tr>
<td>$m(K_S^{0}\pi^+\pi^-\pi^+)$</td>
<td>$m(D^+) \pm 30$ MeV/$c^2$</td>
</tr>
<tr>
<td>$p^*(D^+)$</td>
<td>$&gt; 1.6$ GeV/$c$</td>
</tr>
<tr>
<td>$p(\pi)$</td>
<td>$&gt; 200$ MeV/$c$</td>
</tr>
</tbody>
</table>
ability is applied. The determination of the $\Delta m$ window is done adopting the procedure used for the $D^{*0}$ channels. The selection requirements of the $D^{**}$ selection are listed in Tab. A.6.

Table A.5: Selection requirements applied in the $D^{*0}$ reconstruction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^{*0} \rightarrow D^{0}\pi^{0}$</td>
<td></td>
</tr>
<tr>
<td>$m(D^{0}\pi^{0}) - m(D^{0})$</td>
<td>$\pm 4$ MeV/c$^2$</td>
</tr>
<tr>
<td>$p^*(D^{*0})$</td>
<td>[1.3, 2.5] GeV/c</td>
</tr>
<tr>
<td>$p^*(\pi^{0})$</td>
<td>[70, 450] MeV/c</td>
</tr>
<tr>
<td>$D^{*0} \rightarrow D^{0}\gamma$</td>
<td></td>
</tr>
<tr>
<td>$m(D^{0}\gamma) - m(D^{0})$</td>
<td>[127, 157] MeV/c$^2$</td>
</tr>
<tr>
<td>$p^*(D^{*0})$</td>
<td>[1.3, 2.5] GeV/c</td>
</tr>
<tr>
<td>$E^*(\gamma)$</td>
<td>[100, 450] MeV</td>
</tr>
</tbody>
</table>

Table A.6: Selection requirements applied in the $D^{*+}$ reconstruction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Selection Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^{**} \rightarrow D^{0}\pi^{+}$</td>
<td></td>
</tr>
<tr>
<td>$m(D^{0}\pi^{+}) - m(D^{0})$</td>
<td>$\pm 3\sigma$</td>
</tr>
<tr>
<td>$p^*(\pi^{0})$</td>
<td>[70, 450] MeV/c</td>
</tr>
</tbody>
</table>
Bibliography


[33] H. Kim on behalf of the BABAR collaboration, [arXiv:hep-ex/08052365].


[38] The BABAR Simulation Production Group, Generating Monte Carlo Events with GenFwkInt,

http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/generators/GenFwkInt/GenFwkInt.html


[42] Crystal Ball Collaboration, DESY F31-86-02.


[48] Root Class Index, TMultiLayerPerceptron

[49] MLPfit: a tool for Multi-Layer Perceptrons


Acknowledgments

I have had a very fortunate and pleasant experience during my PHD course, having had the chance to learn and to enjoy myself. First of all I would like to thank Prof. Maurizio Biasini for being a great supervisor and for always advising me on working on the right topics and with the right people. Essential for me it has been the support of Dr. Roberto Covarelli, whose precious collaboration has had a major impact on this work. I would like to thank Dr. Pasquale Lubrano for reviewing this thesis.

The present work has been feasible also thanks to the stimulating cooperation with Dr. Francesco Renga, Dr. Riccardo Faccini, and Dr. Paul Jackson, which I really want to thank.

Ora la parte dei ringraziamenti meno informale ma certo non meno importante.

A Roberto un grazie per la pazienza nel rispondere a domande piú o meno intelligenti, nell’ascoltare i miei lamenti (e questo é stato reciproco!!), per le discussioni sempre edificanti e per le pause caffé.

Venendo al mondo “BABAR” ringrazio....

Alessia e la Vir (in ordine alfabetico): per le belle chiacchierate al ROB, perché hanno portato un tocco di tacco e unghie dipinte. Ringrazio in particolare la prima per le nottate estive a parlare di $m_{ES}$ (..che se avessi avuto una bambolina voodoo....) e per l’impegno come PR! Ringrazio invece la seconda per i frappuccini..che sono durati finché non ci hanno chiesto soldi per la guerra..e per il detto “ognuno balla con sua nonna”!

Silvano: per i momenti trash nell’ufficio di Genova-Roma-resto del mondo; grazie per essere stato il primo maestro di IFR, mio e di quell’altra sprovveduta barese e per tanti altri momenti che ricordo con piacere!

la sora Cenci: impagabile per l’ospitalitá e le cene, che una volta capito con quale spugna va pulito il lavandino poi é tutta in discesa, per quella ventata di “Valle del Tevere” nell’altro emisfero..e come promesso..per il RooWorkspace.....
Francesco (R): perché è così demanding che per stargli dietro mi sono dovuta dar da fare parecchio..e perché la bevuta della brocca di vodka alla mia festa di laurea è indelebile nella memoria dei miei parenti.

Betta e Viola: per la-bira-prima e la-bira-dopo, per il coraggio di Viola di indossare l’abito argentato alla festa di Giampy..e per tante altre serate..

Tina e Antonio: che è come avere la mamma e papá a Palo Alto ...e poi “vuoi tu Concita” chi se lo dimentica????

Mavvio: cioè lui non so se lo ringrazio tanto..l’estate del commissioner l’ho stramaledetto..la sua acidità la odio...ma ha rischiato con me la vita nei taxi-matrix a Calcutta e l’infezione nel bar dell’inteligenzia..quindi sí alla fine un pochino lo ringrazio!

Cervelli: per la sua impareggiabile classe.

Sir Sacco: per le lezioni di ingelse, perché riusciva a bere il caffé del ROB nella tazzina con tanto di piattino conferendogli un minimo di dignitá...ma soprattutto per avermi regalato un bonus della sua triple-A quando ho lasciato le chiavi nella macchia accesa..e chiusa..sempre colpa di Mavvio.

La Lory: per il bel mese a casa-IFR e le chiacchierate notturne.

Fra (P): soprattutto per l’aperitivo prima di fuki sushi quella sera che la tetesca ci fece spellare vivi.

..e poi Maurizio per i mondiali 2006, Sarvati per il sushi a casa, Finocchiaro e Anulli per le onion rings, Neus e Feltresi per le lezioni di tango..

Tornando a Perugia, ringrazio i compagni di ufficio e corridoio in particolare la Bene, il Bonilla e Ale.

Ringrazio infine la mia famiglia e naturalmente Marco, che è il più “stremato” di tutti da questi tre anni di dottorato e dal continuo vai e vieni...grazzzzzzzie!