UNIVERSITÀ DEGLI STUDI DI CAGLIARI

Scuola di Dottorato in Scienze e Tecnologie Fisiche
(Ciclo XX)

A study of rare B-meson decays with muons in the final state with the LHCb detector

Supervisors:
Prof. Biagio Saitta
Dott. Walter Bonivento

Thesis by:
Nicola Serra

Dall'altra parte Antìloco pietoso
lagrimando dirotto, e di cordoglio
spezzato il petto rattenea d'Achille
le terribili mani, onde col ferro
non si squarciasse per furor la gola.
Udì del figlio l'ululato orrendo
la veneranda Teti che del mare
sedea ne' gorghi al vecchio padre accanto.
Mise un gemito, e tutte a lei dintorno
si raccolser le Dee, quante ne serra
il mar profondo, di Nerèo figliuole
Glauce, Talìa, Cimòdoce, Nesea
e Spio vezzosa e Toe ed Alie bella
per bovine pupille, e la gentile
Cimòtoe ed Attea: quindi Melìte
e Limnòria e Anfitòe, Jera ed Agave,
Doto, Proto, Ferusa e Dinamena
e Desamena ed Amfinòma e seco
Callïanìra e Dori e Panopea,
e sovra tutte Galatea famosa;
v'era Apseude e Nemerte e con Janira
Callïanassa ed Ïanassa; alfine
l'alma Climene, e Mera ed Oritìa
ed Amatea dall'auree trecce, ed altre
Nerëidi dell'onda abitatrici.

Contents

List of Figures

List of Tables

Introduction

The Standard Model gives a successful description of known phenomena, however there are various indications of the existence of New Physics. The main purpose of the LHC experiments will be the search of the Higgs boson and the search of physics beyond the Standard Model at the TeV scale.
The LHC is a proton-proton collider, which will operate at a center of mass energy of $14TeV$. It is under construction at the CERN in Geneva. It is foreseen to start in the middle of 2008. At the LHC there are three experiments mainly devoted to the search of new physics: ATLAS, CMS and LHCb.
ATLAS and CMS are the biggest experiments and they will look for direct evidence of new particles, or new degrees of freedom. Their very first targets will be the search of the Higgs boson and super-symmetric particles. LHCb is dedicated to the physics of b-hadrons, it will look for indirect evidences of new physics measuring branching ratios, decay amplitudes and CP asymmetry effects in b-hadrons, which can be sensitive to new physics effects.
The direct and indirect approach to the search of NP is quite complementary. In particular, the indirect approach is useful not only to discover but also to understand and discriminate between the different theories of NP.
This work is a study of rare B-meson decays with muons in the final state, in context of the LHCb experiment.
Three different analyses are here presented:
-The sensitivity to the $B_s^0 \to e^{\pm}\mu^{\mp}$ decay;
-the LHCb sensitivity to the $B_s^0 \to \mu^+\mu^-\gamma$ decay;
-the correction of angular biases in the $B_d^0 \to K^{*0}\mu^+\mu^-$ decay [1].
This thesis is organized into six chapters. In the first chapter the theoret-

---

[1] The study of the correction of angular biases in the $B_d^0 \to K^{*0}\mu^+\mu^-$ channel was carried out in collaboration with the high energy group of the Imperial College of London.

ical framework is briefly reviewed. In chapter 2 the LHCb experiment and the different parts of the detector are described. In chapter 3 the statistical methods used for this work are summarized. In chapter 4 and 5 the studies of the LHCb sensitivity to the $B^0_{s,d} \to e^\pm\mu^\mp$ [5] and $B^0_{s,d} \to \gamma\mu^+\mu^-$ [6] channels are presented. In chapter 6 a study of the correction of angular biases in the $B^0_d \to K^{*0}(892)\mu^+\mu^-$ is discussed [7].
All the three analyses here presented are originals, the two sensitivity study ware never attempted before in LHCb.

Chapter 1

Physics beyond the Standard Model involved in rare B-meson decays

1.1 Problems of the Standard Model

The Standard Model is a quantum field theory based on the local gauge symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. There are three families of quarks and leptons. Each family is composed of two $SU(2)_L$ doublets, and three $SU(2)_L$ singlets. For instance the family $[Q, u^c, d^c, L, e^c]$ with $Q = (u, d)$ and $L = (e, \nu)$ $SU(2)_L$ doublets and u^c, d^c and e^c being charge conjugate $SU(2)_L$ singlets transforms as $\left[(3, 2, \frac{1}{3}), (\overline{3}, 1, -\frac{4}{3}), (\overline{3}, 1, \frac{2}{3}), (1, 2, -1), (1, 1, 2)\right]$ under the gauge symmetry group [1]. The SM has 19 free parameters: 9 charged fermion mass, 4 CKM mixing angles, 3 gauge coupling constants, the Higgs vacuum expectation value v, the Higgs quartic coupling λ and the QCD θ parameter [2]. Recently the neutrino oscillation experiments proved that the neutrino has a mass, if we consider 3 light Majorana neutrinos there are at least another 9 parameters: 3 masses , 3 mixing angles and 3 phases. This description of the elementary particle interactions has been very successful

---

[1] Here we use the convention that the fields are left-handed and the electric charge is $Q_{EM} = T_{3L} + \frac{Y}{2}$.

[2] The most general QCD lagrangian is $L = -\frac{1}{4}trF_{\mu\nu}F^{\mu\nu} - \frac{n_f g^2 \tilde{\theta}}{32\pi^2}trF_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\gamma^\mu D_\mu - me^{i\theta'\gamma^5})\psi$, the two parameters $\tilde{\theta}$ and θ' contribute to the physical measurable parameter θ which can be responsible for strong CP-violation. The fact that $\theta \simeq 0$ to fit the experimental bound on the electric dipole moment of the neutron is considered a fine-tuning problem of the SM.

until now. However there are reasons to look for theories beyond the SM. The SM is not considered a good candidate to be the ultimate theory of nature because it is affected by **naturalness**, **hierarchy** and **fine tuning problems**. In particular any quantum field theory (QFT) is considered an effective theory valid up to a determinate cut-off Λ, beyond which a new high energy theory is valid.

The problem of fine-tuning is related to the behavior of the low energy theory parameters with respect to the Ultra Violet (UV) cut-off Λ.

The parameters of a low energy theory can be either sensitive or insensitive to the UV cut-off. Being m a UV sensitive parameter of the low energy theory the following classification can be made. The theory is called **natural** if $m \sim \Lambda$ (i.e. $m \simeq \frac{g\Lambda}{4\pi}$ for the Higgs mass). It is called **symmetry-natural** if there is a symmetry limit in which $m = 0$ but because of symmetry breaking we can have $m \ll \Lambda$. Another possibility is that there is some tuning at the UV scale Λ due to some feature of the high energy QFT, i.e. $m_1 = m_2$, and radiative corrections to this relation leads to logarithmic divergences in the low energy QFT parameter m. In this case the theory is called **super-natural**. Finally the last possibility is that the theory is **unnatural** in which case there is a positive power law dependency on the UV scale and a fine tuning mechanism allows $m \ll \Lambda$. The latter case is unwanted as we don't know of any physical mechanism which produces fine-tuned theories.

In the SM there is only one parameter which has a power dependence on the UV cut-off, namely the Higgs boson mass. The Higgs boson of the SM is the missing particle needed to confirm the exactness of the model, however it is also a source of problems [1]. The Higgs mass is a free parameter in the SM, it is related to its vacuum expectation value (VEV) by the relations:

$$m_{Higgs} = \sqrt{2\lambda} v \tag{1.1}$$

$$v = (\sqrt{2} G_F)^{-\frac{1}{2}} = 247 GeV, \tag{1.2}$$

where v is the VEV, G_F is the Fermi coupling and λ is the quartic Higgs coupling, which is a free parameter. However precise measurements on the SM electroweak sector and measurements of the masses of the top quark and of the $W^{\pm}$ bosons are sensitive to $log(m_{Higgs})$ through radiative corrections.

This allows us to put an upper bound on the Higgs mass. The present best limit is $m_{Higgs} < 186 GeV$ at 95% confidence level. This is in agreement with theoretical constrains obtained requiring that the SM remains a perturbative theory up to the grand unified theory (GUT) scale[3].

Whatever the Higgs boson mass is, it receives divergent contributions from each particle that couples with it. Each fermion that couples with the Higgs contributes to the radiative correction of the Higgs boson mass through the diagram of Figure 1.1 with terms like:

$$\Delta m_{Higgs} = -\frac{\lambda_f^2}{8\pi^2}\Lambda_{UV} + ... \quad , \tag{1.3}$$

where Λ_{UV} is the cut-off momentum of the virtual particle in the loop.

If the SM is valid up to the Planck scale M_{Planck}, it would receive from all quarks and leptons radiative contributions of the order of M_{Planck}. Therefore, to maintain $m_{Higgs} < 190 GeV$, it is necessary to fine-tune the Higgs bare and visible mass [4]. This fine-tuning is considered unnatural and difficult to explain because it invokes an accidental UV relation. Moreover, if the cut-off of the SM is the Planck scale, it is very unnatural that the electroweak scale is 16 orders of magnitude less than M_{Planck}, this is known

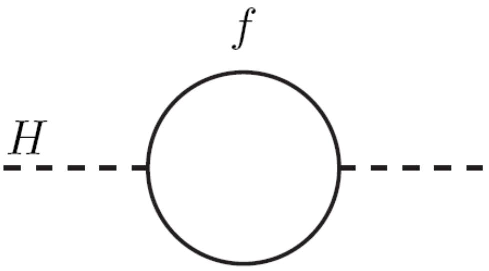


Figure 1.1: One loop quantum correction to the Higgs mass due to a fermion f.

---

[3] The condition that the SM remains a perturbative QFT up to $\Lambda_{GUT} = 10^{16} GeV$ constrains the Higgs boson mass to be in the range $130 GeV - 190 GeV$

[4] At one loop the radiative correction to the Higgs mass is: $\delta m_H^2 = \frac{\Lambda_{SM}}{16\pi^2 v^2}(2m_W^2 + m_Z^2 + m_H^2 - 4m_t^2)$, where Λ_{SM} is the cut-off of the SM. To avoid fine-tuning between these parameters, i.e. $\frac{\delta m_H}{m_H} < 10$, it must be that $\Lambda \sim fewTeV$ [2]

as the hierarchy problem.

The three main problems of the SM are therefore very closely related, and for all these problems there are common solutions.

The most likely solution is that the SM is a natural or symmetry-natural theory, in this case the cut-off of the SM must be of the order of the electroweak scale: few TeV[5]. Obviously the new theory which replaces the SM at the TeV scale must explain why the Planck scale is 16 orders of magnitude higher. However there are various possible solutions to this question. For instance the theory can contain only logarithmic divergences and in this case it is a super-natural theory. Another possibility is that there are several effective QFT with increasing cut-offs approaching the Planck scale. Finally it is possible that the quantum gravity cut-off is not the Planck scale but significantly lower.

We do not have any experimental evidence of quantum gravity effects, moreover the most precise experimental data of classical gravitation is at the centimeter scale.

Up to now only the *string theory* is considered a possible candidate to be a quantum theory of gravity. Its basic idea is to replace the point-like particles of the QFT with $1-$dimensional elementary objects called strings. Different kinds of string theories have been proposed. In all these theories, to avoid anomalies, the number of space-time dimensions must be greater then four, typically ten or eleven. Moreover other multidimensional elementary objects called $D-$branes have been studied in the context of string theory. In these models the extra-dimensions can be either compactified, i.e. warped with a very small radius, or extended with the requirement that the gauge fields are constrained to live in the usual $(3+1)-$dimensional space time. In particular various models which predicts extra-dimensions have shown that the quantum gravity cut-off can be at the TeV scale (see for instance [32]). Therefore by assuming that the quantum gravity cut-off is given by the Cavendish experiment ($(G_{Newton})^{-\frac{1}{2}} \simeq 10^{19} GeV$) we are extrapolating the experimental results up to 31 orders of magnitude, from classical gravity

---

[5] The most natural explanation is that $m_{Higgs} \sim \frac{g\Lambda}{4\pi}$, in this case $\Lambda \sim 1TeV$. However many new physics couplings at the scale of $\Lambda \sim 1TeV$ have already been ruled out [4]. If the SM is a symmetry-natural theory the cut-off can be increased up to $10TeV$.

to quantum gravity. This approach could be very naive.
Moreover, even if we choose to allow for the fine-tuning of the SM there are other reasons, related to the perturbative behavior of the SM [1], to choose a cut-off much lower than the Planck scale.
However theoretical physicists don't consider this model a good candidate to be the ultimate theory of nature. as it has too many free parameters. A great deal of effort has been made trying to solve the problems of the SM within a theory that unifies all the forces of nature.

1.2 Super-symmetric extensions of the Standard Model

The most promising candidate to replace the SM at the TeV scale is the supersymmetric extension of the SM (SUSY). Many SUSY models are supernatural theories up to the Planck scale, therefore they would be able to solve the hierarchy problem of the SM. Moreover, the unification of gauge couplings doesn't occur in the SM and it is achievable in several SUSY theories. Supersymmetry is an additional symmetry of Nature which involves bosonic and fermionic degrees of freedom. This symmetry transforms boson into fermion and vice versa. The simplest supersymmetric action of non-interacting particles is given by the Weiss-Zumino model, Equation 1.4:

$$S = -\int d^4x \partial_\mu \phi \partial^\mu \phi^* + i\psi^\dagger \overline{\sigma}^\mu \partial_\mu \psi. \tag{1.4}$$

Where ψ is a fermionic field and ϕ is a bosonic field and they transform under supersymmetric transformation as in Equation 1.5.

$$\begin{aligned} &\delta\phi = \epsilon\psi, \delta\phi^* = \epsilon^\dagger\psi^\dagger \\ &\delta\psi_\alpha = -i(\sigma^\mu\epsilon^\dagger)_\alpha \partial_\mu\phi,\ \delta\psi^\dagger_{\dot\alpha} = i(\epsilon\sigma^\mu)_{\dot\alpha}\partial_\mu\phi^*, \end{aligned} \tag{1.5}$$

where ϵ_α is an infinitesimal anticommuting Weyl fermion object.
The Minimal extension of the SM (MSSM) is the minimal SUSY model which can be obtained from the SM without external assumption. This model has a large number of free parameters, there are 19 parameters which correspond to the SM free parameters (see section 1.3) and 105 genuinely new parameters

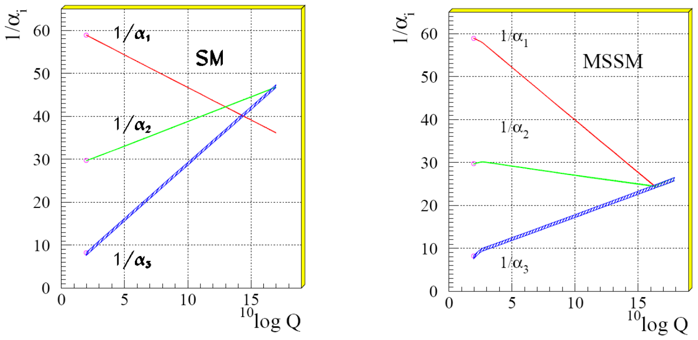
of the MSSM. For each fermion of the SM there are two super-symmetric partners associated to the left-handed and right-handed fermions. In general they are not mass eigenstates but they can mix. These particles are called *squarks* and *sleptons*. The Higgs sector of the MSSM is composed of two hypercharge ($Y = \pm 1$) Higgs doublets which generate masses for up-type and down-type quarks and charged leptons. There are five physical Higgs particles: a charged pair $H^{\pm}$, a CP-odd neutral boson A^0 and two neutral CP-even bosons h^0 and H^0. The tree level Higgs sector parameters depend on two quantities: the ratio between the two Higgs VEVs ($\tan\beta = \frac{v_u}{v_d}$) and the mass of the A^0 boson. Additional parameters enter when one loop radiative corrections are included.

Finally there are super-symmetric partners of the gauge and Higgs bosons. The name of these particles is obtained by adding the suffix '-ino' to the name of the correspondent partner of the SM, i.e. the super-partner of the gluon is the gluino, which is a spin 3/2 fermion. Even if the general MSSM has 124 free parameters, most of them can be constrained by asking for the consistency of the MSSM with the present experimental bounds. For instance most of the parameter regions of the MSSM don't conserve the leptonic numbers for the three families, exhibit new sources of CP-violation inconsistent with the actual experiments and the flavor changing neutral currents (FCNC) are unsuppressed. Requiring that the MSSM respects these phenomenological bounds we obtain SUSY models with significantly less parameters, in this case people speak of CMSSM (constrained MSSM). One interesting model is the minimal Super-gravity model (mSUGRA), in this case the MSSM and its interactions are determined by only five parameters in addition to the usual 19 SM parameters: m_0, A_0, $m_{1/2}$, $\tan\beta$ and the sign of the μ parameter. The MSSM possesses a multiplicative conserved quantum number, the R-parity $R = (-1)^{3(B-L)+2S}$. All the SM particles are even R-parity eigenstates and the super-partners are odd R-parity eigenstates. R-parity conservation implies that the lightest super-symmetric particle (LSP) is stable and it is a good candidate for the cold Dark Matter.

1.3 Grand Unified Theory (GUT)

The theory that unifies the electromagnetic force, the strong and the weak force in a unique force is called GUT (Grand Unified Theory). The difference with the so called TOE (Theory Of Everything) is that GUT doesn't include gravity. There are at least three good reasons to deal with gravitation separately with respect to the other forces:
1)The natural scale of Quantum Gravity is $M_P \sim 10^{19} GeV$ whereas the unification of the other forces should happen at $M_G \sim 10^{16} GeV$;
2)Up to now there has not been a universally accepted quantum theory of gravity to look at, as it happens for the other forces;
3)Gravitation is a particular force, that is responsible for the space-time curvature, it is not completely clear that the difference between the Planck mass and the GUT scale poses a fine tuning problem.
Since quarks and leptons are both fundamental and without structure it is desirable to have a symmetry that unify these two particles. The easiest way to unify quarks and leptons is through the well known Pati-Salam gauge group $SU(4)_C \times SU(2)_L \times SU(2)_R$ [18]. One family is now expressed by $[(Q, L), (Q^c, L^c)]$ transforming as $\left[(4, 2, 1), (\overline{4}, 1, \overline{2})\right]$. In this model the $Q = (u^c, d^c)$ and $L = (e^c, \nu^c)$ are doublets under $SU(2)_R$ and the left/right-handed symmetry is restored at high energy. The electric charge is given by the relation $Q_{EM} = T_{3L} + T_{3R} + \frac{1}{2}(B - L)$, B being baryon number and L being lepton number. The right-handed neutrino is desirable when considering neutrino masses and it is needed to complete the $SU(2)_R$ lepton doublet. It has not SM quantum number and it is therefore "sterile". This kind of unification doesn't unify the gauge couplings, there are still 2 independent couplings if we impose $L \leftrightarrow R$ parity symmetry and 3 if we do not. To unify the gauge couplings it is necessary to insert the SM in a simple unified gauge group. The simplest way to do this is the Georgi-Glashow theory based on the symmetry group $SU(5)$, where this group breaks in one step to the SM. The requirement for the couplings to be equals at the GUT scale (M_G) put a prediction on the nucleon life time.

The simplest $SU(5)$ model was already ruled out by Superkamiokande



(a) (b)

Figure 1.2: Evolution of the inverse of the three gauge coupling constants within the SM (a) and the MSSM (b).

results of the proton lifetime and LEP measurement of $sin^2\theta_W$. Moreover, non-SUSY GUTs with only one breaking scale were already ruled out because it is not possible to obtain the unification of the three couplings. In Figures 1.2(a)-1.2(b) the predicted behavior of the gauge couplings defined in 1.6, assuming the SM and the MSSM, is shown [8]. The unification of the couplings is achievable only in the latter case.

$$\begin{aligned}\alpha_1 &= \frac{5}{3}\frac{g'^2}{4\pi}\\ \alpha_2 &= \frac{g^2}{4\pi}\\ \alpha_3 &= \frac{g_s^2}{4\pi}.\end{aligned}\tag{1.6}$$

However non-Susy GUTs with a more complicated breaking mechanism can still fit experimental data. For instance a great deal of attention has been put on non-SUSY models $SO(10) \rightarrow SU(4)_C \times SU(2)_L \times SU(2)_R \rightarrow SM$. It was shown that these models with the second breaking scale determined by light neutrino masses using the see-saw mechanism fits well with the low gauge coupling behavior [9], it is in agreement with the present nucleon decay bounds [10] and it provides a natural accommodation for the light neutrino masses. In [11] (see below) it was also stressed that the $SU(4)_C \times SU(2)_L \times SU(2)_R$ symmetry breaking can give an alternative explanation to the hierarchy problem with respect to the usual SUSY or large extra-dimensions theories. Recently a great deal of progress has been made constructing intersecting D-brane models within the string theory that can naturally incorporate the Pati-Salam symmetry group in 4 dimensions (see for instance [32]).

1.4 Two Higgs doublet model (2HDM).

The most straightforward extension of the SM is the two Higgs doublet model (2HDM). This model contains two complex $SU(2)$ higgs doublet scalar fields:

$$\phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix}, \qquad \phi_2 \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix}. \tag{1.7}$$

These fields acquire VEVs ($<\phi_i^0>= v_i$), breaking in this way the electroweak symmetry. In this model the W boson mass is given by $M_W^2 = \frac{1}{2}g^2(v_1^2 + v_2^2) = \frac{1}{2}g^2 v_{SM}^2$, where v_{SM}^2 is the SM Higgs VEV. Usually this model is parametrized with the v_{SM} value and with $\tan\beta = \frac{v_2}{v_1}$.

The 2HDM can be thought as an effective model. It has the same particle content as the Higgs sector of the MSSM.

When $\tan\beta$ is large, the Yukawa coupling to the b quark is of the order 1 and large effects in B decays are therefore expected.

1.5 New Physics in the $B_s^0 \to e^\pm\mu^\mp$ channel.

1.5.1 The Pati-Salam model

In the usual Pati-Salam model a particle called Leptoquark (LQ) arises when the gauge group $SU(4)_C$, which unifies quarks and leptons, is spontaneously broken [18]. This particle acquires a mass dependent on the scale at which the symmetry is broken. This model explains why quarks experience the strong force while leptons do not, why the quark charge is fractional and unifies in an elegant and natural way quarks and leptons. The main difference between quarks and leptons is that quarks experience strong interaction. This difference is removed if we interpret the lepton number as the fourth color. The easiest way to do this is to arrange quarks and leptons into families like:

$$\begin{pmatrix} u_R & u_G & u_B & \nu_l \\ d_R & g_G & d_B & l \end{pmatrix}_L \tag{1.8}$$

$$\begin{pmatrix} u_R & u_G & u_B & \nu_l \end{pmatrix}_R \tag{1.9}$$

$$\begin{pmatrix} d_R & d_G & d_B & l \end{pmatrix}_R, \tag{1.10}$$

where R,G,B are the usual three colors Red, Green and Blue and where L(R) index denotes left-handed (right-handed) chirality.

In this model the break-down of the $SU(4)_C \rightarrow SU(3)_C \times U(1)_{L-B}$ allows lepton number violation for each family but preserves the $B - L$ quantum number and the total lepton number as well. This mechanism allows elementary interactions represented by the vertex in Figure 1.3, in which the dashed line is the LQ. This particle carries both baryon and lepton numbers. Because the total lepton number and the B-L number must be conserved, the LQ cannot mediate pure fermionic interactions like $\mu \rightarrow e\gamma$ or the nucleon decay.

However effective four fermion interactions should appear in the neutral mesons lagrangian. Because we have three families of quark and leptons there are 6 possible different couplings. Choosing for instance the coupling (d_C, e) and (s_C, μ) an effective four fermion interaction appears in the neutral K-meson decay through the effective lagrangian:

$$L_{eff} = \frac{g_4}{2M_C^2}\bar{d}\gamma^{\mu} e\bar{\mu}\gamma_{mu} s + h.c., \tag{1.11}$$

where M_C is the LQ mass, the g_4 coupling is the Pati-Salam coupling and the sum over colors is implicit. Because of the fact that $SU(4)_C$ breaks to $SU(3)_C$ the g_4 coupling is just the strong coupling at the M_C scale. The lagrangian of Eq. 1.11 corresponds to the $K \rightarrow e^{\pm}\mu^{\mp}$ decay, through the tree level Feynman diagram of Figure 1.4. The branching ratio for the $K_L^0 \rightarrow e^{\pm}\mu^{\mp}$ decay can be easily computed [19] evolving the quark masses and the coupling constant to the M_C scale. The branching ratio depends on the inverse of the LQ mass at the fourth power. Therefore a limit on the $K_L^0 \rightarrow e^{\pm}\mu^{\mp}$ decay is directly translated into a limit on the LQ mass.

However there is not a deep reason to couple the d-quark family with the e-lepton family and the s-quark family with the μ-lepton family, but (as was

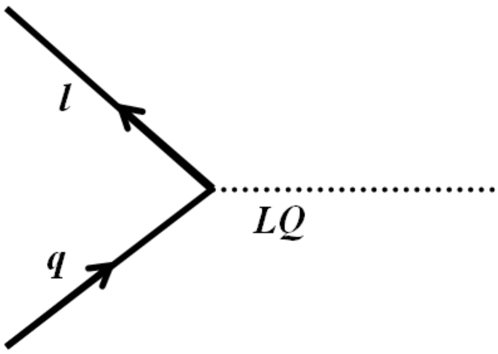


Figure 1.3: Quark-Leproquark-Lepton vertex which arises in the Pati-Salam model.

Table 1.1: Lower bound on Pati-Salam LQ mass (TeV) from rare K, π and B meson decays. The coupling scheme is indicated in the first column.

	$K_L^0 \to \mu^\pm e^\mp$	$\frac{\pi^+\to e^+\nu}{\pi^+\to\mu^+\nu}$	$\frac{K^+\to e^+\nu}{K^+\to\mu^+\nu}$	$B_d^0 \to e^\pm\mu^\mp$	$B_s^0 \to e^\pm\mu^\mp$
$e\mu\tau$	1950	250	4.9		
$\mu e\tau$	1950	76	130		
$e\tau\mu$		250		50	
$\mu\tau e$		76		50	
$\tau\mu e$			4.9		20.7
$\tau e\mu$			130		20.7

suggested in [19]) we have to keep an open mind on all the possible couplings. In Table 1.1 all the best limits currently available on quark-lepton couplings [6] are shown. These limits are obtained in the approximation $V_{CKM} \simeq I$, a more general coupling scheme without this assumption is discussed in Section 1.5.2.

In the case of the coupling scheme with the τ-lepton family associated to the third or the second quark family the most sensitive probes of the existence of the LQ couplings are lepton flavor violating (LFV) decays of neutral B mesons in electron and muon ($B_{s,d}^0 \to e^\pm\mu^\mp$). The branching ratio for these LFV decays can be computed with the Feynman diagrams in Figure

---

[6] This table is taken from [19] and updated by us.

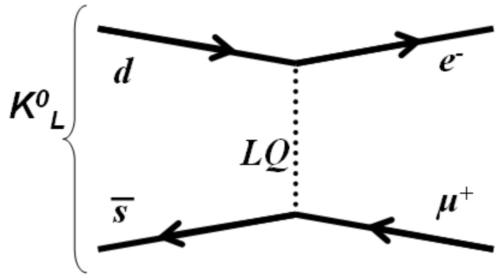


Figure 1.4: Leptoquark mediated LFV neutral K^0_L meson decay into electron and muon ($K^0_L \to e^{\pm}\mu^{\mp}$).

1.6 and is given by the following Equation 1.12:

$$\Gamma(\overline{B}^0_{d,s} \to e^-\mu^+) = \pi\alpha_s(M_C)\frac{F^2_{B(d,s)}m^3_{B(d,s)}R^2}{M_C^4}$$
$$R = \frac{m_{B(d,s)}}{m_b}\left(\frac{\alpha_s(M_C)}{\alpha_s(m_t)}\right)^{-\frac{4}{7}}\left(\frac{\alpha_s(m_t)}{\alpha_s(m_b)}\right)^{-\frac{12}{23}} \tag{1.12}$$

where the coefficients $F_{B(s,d)}$ are $B^0_{(s,d)}$ decay constants, $m_{B(s,d)}$ is the $B^0_{s,d}$ mass and m_t and m_b are the b and t quark masses. These decays were already studied by the CDF and Belle experiments. The world's best limits at present are: $Br(B^0_s \to e^{\pm}\mu^{\mp}) < 6.1 \cdot 10^{-6}$ @ 90% ([26]) and $Br(B^0_d \to e^{\pm}\mu^{\mp}) < 1.7 \cdot 10^{-7}$ @ 90% ([27]).

1.5.2 The Pati-Salam model with Leptoquark-W boson mixing

In the previous section the possible lepton-quark couplings were examined, however chirality eigenstates are not mass eigenstates. Quarks and leptons acquire mass through the Higgs mechanism. Since it is not possible to diagonalize the quark mass matrix and the quark flavor matrix together, there are transitions between different quark families. Choosing the basis in which (u, c, t) are mass eigenstates, we have:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} \mid d\rangle \\ \mid s\rangle \\ \mid b\rangle \end{pmatrix} = \begin{pmatrix} \mid d'\rangle \\ \mid s'\rangle \\ \mid b'\rangle \end{pmatrix}. \tag{1.13}$$

It can be shown that a mixing in the interaction of LQ currents with the W boson necessarily occurs [22]. Moreover to avoid ultraviolet divergent contributions to $l \to l'$ transition via quark-leptoquark pairs it is necessary to include this mixing at the tree-level. Let us indicate with (u_l, d_l) the quark doublet associated with the lepton l. When we choose the basis in which the charged leptons are diagonal we have:

$$\begin{pmatrix} u_e & d_e \\ \nu_e & e \end{pmatrix}, \begin{pmatrix} u_\mu & d_\mu \\ \nu_\mu & \mu \end{pmatrix}, \begin{pmatrix} u_\tau & d_\tau \\ \nu_\tau & \tau \end{pmatrix}, \tag{1.14}$$

The states in 1.14 are not mass eigenstates but for the charged leptons. We can write:

$$\nu_l = K_{in}\nu_n \tag{1.15}$$

$$u_l = U_{in}u_n \tag{1.16}$$

$$d_l = D_{in}d_n \tag{1.17}$$

where the states ν_n, u_n, d_n are the mass eigenstates and the matrix K, D, U are unitary matrices. If we insert the 1.15-1.17 in the lagrangian for the charged weak interaction we obtain:

$$L_W = \frac{g}{2\sqrt{2}}\left[K^*_{li}\overline{\nu}_i\gamma^\mu(1-\gamma_5)l + U^*_{ln}D^*_{lm}\overline{u}_n\gamma^\mu(1-\gamma_5)d_m\right]W^*_\mu + h.c. \tag{1.18}$$

In Equation 1.18 can be seen that the usual CKM matrix is related to the U, D matrices through the relation $V = U^\dagger D$. In the breakdown of the $SU(4)_C$ three charged and colored LQ arise, their interaction with fermions is described by the lagrangian:

$$L_{LQ} = \frac{g_S(M_C)}{\sqrt{2}}\left[D_{ln}\bar{l}\gamma_\mu d^c_n + (K^\dagger U)_m k(\overline{\nu}_m\gamma_\mu u^c_k)\right]X^\mu_c + h.c. \tag{1.19}$$

At low energy, when the transferred momentum is negligible with respect to the LQ mass, the effective lagrangian leads four-fermion quark-lepton interactions. Once again this interactions could be responsible for LFV neutral meson decays. However, now it is not possible to translate directly the limit in the branching ratios of these decays into a limit in the LQ masses, as was done in Table 1.1. Therefore the LQ mass lower bound could be lower than

Table 1.2: Lower bound on the LQ mass (TeV) from rare meson decays when the mixing with the $W^{\pm}$ boson is taken into account.

Experimental limit	Bound on LQ	Reference
$Br(K^+ \to \pi^+\mu^-e^+) < 5.2 \cdot 10^{-10}$	$\frac{M_{LQ}}{\lvert D_{es}D^*_{\mu d}\rvert^{1/2}} > 96TeV$	[23]
$Br(K^+ \to \pi^+\mu^+e^-) < 1.3 \cdot 10^{-11}$	$\frac{M_{LQ}}{\lvert D_{es}D^*_{\mu s}\rvert^{1/2}} > 240TeV$	[24]
$Br(K_L \to \mu e) < 4.7 \cdot 10^{-12}$	$\frac{M_{LQ}}{\lvert D_{ed}D^*_{\mu s}+D_{es}D^*_{\mu d}\rvert^{1/2}} > 1950TeV$	[25]
$Br(B^0_s \to \mu e) < 6.1 \cdot 10^{-6}$	$\frac{M_{LQ}}{\lvert D_{es}D^*_{\mu b}+D_{eb}D^*_{\mu s}\rvert^{1/2}} > 20.5TeV$	[26]
$Br(B^0_d \to \mu e) < 1.7 \cdot 10^{-7}$	$\frac{M_{LQ}}{\lvert D_{ed}D^*_{\mu b}+D_{eb}D^*_{\mu d}\rvert^{1/2}} > 50TeV$	[27]

expected when the mixing is neglected.

In Table 1.3 various processes sensitive to LQ interaction with the matrix elements involved are listed. The only mixing independent LQ mass bound is given by the cosmological measurement $Br(\pi^0 \to \nu\overline{\nu}) < 2 \cdot 10^{-13}$ [28] which corresponds at $M_c > 18TeV$.

It is clear that a better knowledge of the LFV neutral meson decays is necessary in order to confirm or exclude this model. The partial width for the processes $B^0_{d,s} \to e^{\pm}\mu^{\mp}$ is related to the Leptoquark mass M_{LQ} through [22]:

$$\begin{aligned}
BR(B^0_{d,s} \to e^{\pm}\mu^{\mp}) &= \Gamma(B^0_{d,s} \to e^{\pm}\mu^{\mp}) \cdot \frac{2\pi \cdot \tau_{B(d,s)}}{h} \\
\Gamma(B^0_{d,s} \to e^{\pm}\mu^{\mp}) &= \pi\alpha_s^2(M_{LQ})\frac{F^2_{B(d,s)}m^3_{B(d,s)}R^2F^{d,s}_{mix}}{M^4_{LQ}} \\
R &= \frac{m_{B(d,s)}}{m_b}\left(\frac{\alpha_s(M_{LQ})}{\alpha_s(m_t)}\right)^{-4/7}\left(\frac{\alpha_s(m_t)}{\alpha_s(m_b)}\right)^{-12/23} \qquad (1.20)
\end{aligned}$$

where $F_{B(d,s)}$ are the $B^0_{d,s}$ decay constants, $m_{B(d,s)}$ the $B^0_{d,s}$ mass, m_b and m_t the running b and t quark masses in the modified minimal subtraction ($\overline{MS}$) scheme, $\tau_{B(d,s)}$ the $B^0_{d,s}$ lifetimes and $\alpha_s(\mu)$ the strong coupling constant evaluated at the scale μ. The factor $F^{d,s}_{mix}$ takes into account generation mixing within the Pati-Salam Model and is, e.g. for the decay $B^0_s \to e^+\mu^-$:

$$F^s_{mix} = \left(\mathcal{D}_{es}\mathcal{D}^*_{\mu b} + \mathcal{D}_{eb}\mathcal{D}^*_{\mu s}\right) \qquad (1.21)$$

where $\mathcal{D}_{ij}$ are the elements of the unitary mixing matrix of the down-type quarks in the representation 1.14.

1.5.3 The Pati-Salam $SU(4)_C \times SU(2)_L \times SU(2)_R$ alternative model

As discussed above, in order to solve the hierarchy and fine tuning problems, there should be a new QFT at the TeV scale which replaces the SM. In [32] it was shown that the hierarchy problem can be solved constructing a gravitational theory in (4+2)-dimensions, with two compactified warped dimensions. In this model the Planck energy ($10^{19}GeV$) is replaced with the EW energy of the order of the TeV and the SM is embedded in a $SU(4)_C \times S(2)_L \times SU(2)_R$ Pati-Salam gauge group in the bulk. In the usual Pati-Salam model the $SU(4)_C$ breaking must be at least at 18TeV, due to the experimental limit on $Br(\pi^0 \to \nu\overline{\nu}) < 2 \cdot 10^{-13}$.

Recently, in [11], a model in which the $SU(4)_C \times S(2)_L \times SU(2)_R$ breaking allows the lepton-quark unification at the TeV scale was proposed. It was also stressed that this unification can give a possible solution to the hierarchy problem.

The fermion multiplets are:

$$Q_L = \begin{pmatrix} d_R & u_R \\ d_G & u_G \\ d_B & u_B \\ E^- & E^0 \end{pmatrix}_L , Q_R = \begin{pmatrix} u_R & d_R \\ u_G & d_G \\ u_B & d_B \\ \nu & e \end{pmatrix}_R , f_L = \begin{pmatrix} (E_R^-)^c & \nu_L \\ (E_R^0)^c & e_L \end{pmatrix} \tag{1.22}$$

where E^-, E^0 is an $SU(2)$ exotic fermion doublet.

They transform under this gauge group as follow:

$$Q_L \sim (4,2,1) \tag{1.23}$$

$$Q_R \sim (4,1,2) \tag{1.24}$$

$$f_L \sim (1,2,2). \tag{1.25}$$

The minimal choice for scalar multiplets that break the gauge symmetry giving mass to the charged fermions is:

$$\chi_L \sim (4,2,1) \tag{1.26}$$

$$\chi_R \sim (4,12) \tag{1.27}$$

$$\phi \sim (1,2,2) \tag{1.28}$$

notice that the scalar multiplets transform with the same quantum numbers as the fermion multiplets of each generation. The lagrangian which takes into account the scalar-fermion interaction is:

$$L = \lambda_1 \overline{Q}_L f_L \tau_2 \chi_R + \lambda_2 \overline{Q}_R f_L \tau_2 \chi_L + \lambda_3 \overline{Q}_L \phi \tau_2 Q_R + \lambda_4 \overline{Q}_L \phi^c \tau_2 Q_R + h.c. \quad (1.29)$$

where $\phi^c = \tau_2 \phi^* \tau_2$. The symmetry is broken when the component $(T = -1, I_{3R} = 1/2)$ of χ_R, the component $(T = -1, I_{3L} = 1/2)$ of χ_L and the two components $I_{3R} = -I_{3L} = \pm 1$ of the ϕ field gain non zero vacuum expectation value (VEV) as shown in 1.30-1.33:

$$< \chi_R(T = -1, I_{3R} = 1/2) >= \omega_R \quad (1.30)$$

$$< \chi_L(T = -1, I_{3L} = 1/2) >= \omega_L \quad (1.31)$$

$$< \phi(I_{3R} = -I_{3R} = -1/2) >= u_1 \quad (1.32)$$

$$< \phi(I_{3R} = -I_{3R} = 1/2) >= u_2, . \quad (1.33)$$

If the VEVs satisfy the relation $\omega_R > u_{1,2}, \omega_L$ [7] the symmetry is broken as in the 1.34.

$$\begin{array}{c} SU(4) \otimes SU(2)_L \otimes SU(2)_R \\ \downarrow \langle \chi_R \rangle \\ SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \\ \downarrow \langle \phi \rangle, \langle \chi_L \rangle \\ SU(3)_c \otimes U(1)_Q \end{array} \quad (1.34)$$

In this model there are four electrically neutral leptons: $\nu_{L,R}$ and $E^0_{L,R}$. The exotic leptons $E^0_{L,R}$ gain masses by the VEV ω_R, their masses must be heavier than $m_Z/2 \sim 45 GeV$ as they don't contribute to the Z decay amplitude. The right-handed neutrino is sterile with respect to the SM gauge group and it gains mass mixing with the E lepton. Its mass is given by the relation:

$$m_{\nu_R} \sim \frac{2 m_u m_e}{m_E}. \quad (1.35)$$

---

[7] If $\omega_R \gg u_{1,2}, \omega_L$ the model reduces to the SM.

where $\phi^c = \tau_2 \phi^* \tau_2$.

The left handed neutrino ν_L is massless at the tree level, but it gains mass radiatively through the interaction of the Feynman diagram shown in Figure 1.5. It is not possible to give a precise estimation of the ν_L mass because there are too many free parameters, such as the E-lepton masses, however one would expect that neutrino mass arising from this radiative process would be rather light. In the hypothesis that Higgs scalar contributions to one-loop correction of the radiative neutrino mass are negligible, it is possible to put a naive theoretical upper limit of 50eV on the neutrino mass [11]. It can be shown that this model respects the present bounds on nucleon decay. As for the other Pati-Salam type models, also this model allows LFV neutral meson decays. This can occur with interactions of the form:

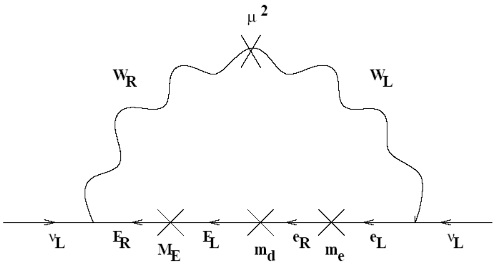


Figure 1.5: Feynman diagram for the radiative neutrino mass.

$$L = \frac{g_s}{\sqrt{2}} \overline{D}^i_R W'_\mu \gamma^\mu K'^{ij} l^j_R + h.c. \tag{1.36}$$

where W' is a colored leptoquark of charge 2/3, $D^{1(2,3)=d(s,b)}$, $l^{1(2,3)} = e(\mu, \tau)$ and the matrix K'^{ij} is a CKM type matrix. The difference with the previous model is that this charged leptoquarks couple only with right-handed fermions and are therefore called chiral leptoquarks. In [11] several LFV decays were examined. The results were interpreted for the possible matrices K'. A particularly interesting configuration is when this matrix is given by 1.37 or 1.38:

$$K' \sim \begin{pmatrix} 0 & 0 & 1 \\ cos\alpha & sin\alpha & 0 \\ -sin\alpha & cos\alpha & 0 \end{pmatrix} \tag{1.37}$$

$$K' \sim \begin{pmatrix} 0 & cos\alpha & sin\alpha \\ 0 & -sin\alpha & cos\alpha \\ 1 & 0 & 0 \end{pmatrix} \tag{1.38}$$

In this case the most sensitive LFV decay to the lepton-quark unification are the decays of neutral B-mesons via the diagram in Figure 1.6. The branching ratio for this process is approximately given by the formula:

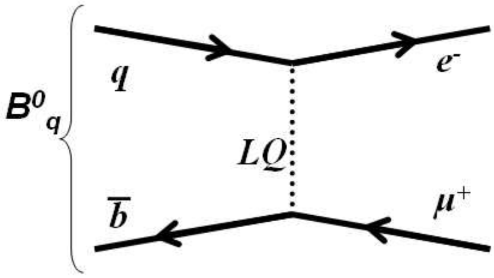


Figure 1.6: Leptoquark mediated LFV B^0_q decay into electron and muon ($B^0_{s,d} \to e^{\pm}\mu^{\mp}$).

$$Br(B^0 \to e^+\mu^-) \simeq 3 \times 10^{-6} \left(\frac{TeV}{M_{W'}}\right)^4 \tag{1.39}$$

where $M_{W'}$ is the mass of the W' boson. In order to solve the hierarchy problem we can put the bound

$$M_{W'} \lesssim few \cdot TeV. \tag{1.40}$$

Notice that this couple scheme allows lepton-quark unification at the TeV scale without invoking super-symmetry and provides a natural explanation for the smallness of neutrino masses. Using the relation 1.40 it is possible to predict LFV neutral B-meson decays with a branching ratio larger than 10^{-9} (this prediction can be tested at the LHCb experiment).
Therefore this model extends the SM at the TeV scale allowing LFV neutral meson decays and adding additional $W'^{\pm}$ bosons, a new Z' boson and a doublet of exotic fermions $E^{0,-}$.

Table 1.3: Possible LQ couplings consistent with the SM gauge group. All the couplings are dimensionless, B and L conserving. In the first four columns the LQ quantum numbers are listed.

Spin	3B+L	SU(3)$_c$	SU(2)$_W$	U(1)$_Y$	coupling
0	-2	$\overline{3}$	1	1/3	$\overline{q}^c_L l_L$ or $u^c_R e_R$
0	-2	$\overline{3}$	1	4/3	$\overline{d}^c_R e_R$
0	-2	$\overline{3}$	3	1/3	$\overline{q}^c_L l_L$
1	-2	$\overline{3}$	2	5/6	$\overline{q}^c_L \gamma^\mu e_R$ or $\overline{d}^c_R \gamma^\mu l_L$
1	-2	$\overline{3}$	2	$-1/6$	$\overline{u}^c_R \gamma^\mu l_L$
0	0	3	2	7/6	$\overline{q}_L e_R$ or $u_R l_L$
0	0	3	2	1/6	$\overline{d}_R l_L$
1	0	3	1	2/3	$\overline{q}_L \gamma^\mu l_L$ or $d_R \gamma^\mu e_R$
1	0	3	1	5/3	$\overline{u}_R \gamma^\mu e_R$
1	0	3	3	2/3	$\overline{q}_L \gamma^\mu l_L$

1.5.4 Leptoquarks in model independent analysis

In the previous sections some Pati-Salam models were discussed. All these models predict LFV B-meson decays. However there are many other models which predict the leptoquark existence [29] and consequently LFV neutral B-meson decays. LQs can be scalars or vectors, their couplings may or may not depend on the fermion chirality (as for instance in the $SU(4)_C \times SU(2)_L \times SU(2)_R$ alternative model). In addition to GUT models in which leptons and quarks are expected to form multiplets, there are other models in which the LQ arises, such as the extended Technicolor and compositness models. All the possible LQ couplings that are dimensionless, conserving B and L number and consistent with the SM gauge symmetry group are listed in Table 1.3 for both scalar and vector LQ. In [29] a model independent analysis of the possible LQ couplings was discussed, however in such an analysis the LQ is less constrained because there are two free parameters instead of one: the LQ mass M_{LQ} and the LQ coupling λ_{LQ}.

1.6 New Physics in the $B^0_{s,d} \to \mu^+\mu^-\gamma$ channel

1.6.1 New Physics in the $B^0_s \to \mu^+\mu^-$ channel

In the indirect search for new physics particularly important are flavor changing neutral currents (FCNC) . These decays are forbidden at tree level in the SM and they are highly suppressed by the GIM mechanism. Because of the breakdown of the GIM mechanism at low energy these branching ratios are not vanishing. In the SM they can proceed only via penguin and box diagrams. Therefore SM contributions via virtual particles might be at the same level as NP contributions.

The experimental measurements of $b \to s\gamma$ transitions ([39], [40] and [41]) allowed us to put very strong constrains on various NP models, such as the 2HDM and MSSM. Other NP sensitive processes are the $b \to sl^+l^-$ transitions.

Among these, the $B_{s,d} \to \mu^+\mu^-$ decay, the $B_{s,d} \to \gamma\mu^+\mu^-$ decay and the $B_d \to K^{*0}\mu^+\mu^-$ decay are particularly interesting.

The fully leptonic decays of the neutral B-meson ($B_{s,d} \to l^+l^-$) are predictable in the SM. They have a very small branching ratio. This is due not only to the FCNC suppression but also to helicity suppression, in fact in this decay the two leptons are emitted with the same helicity. Therefore these branching ratios are smaller for lighter leptons (Eq. 1.41).

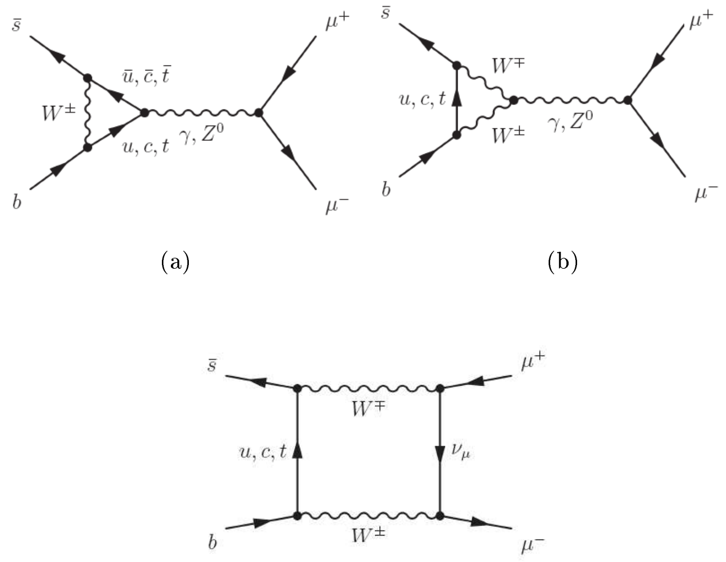
$$\begin{aligned} Br(B^0_s \to e^+e^-) &\sim 10^{-14} \\ Br(B^0_s \to \mu^+\mu^-) &\sim 10^{-9} \\ Br(B^0_s \to \tau^+\tau^-) &\sim 10^{-6} \end{aligned} \tag{1.41}$$

All these decays are very promising for what concerns the search of physics beyond the SM, in particular it was shown in [35] that they can be useful for the constrain of the CMSSM parameters. Even if the $B_s \to \tau^+\tau^-$ decay has the largest branching ratio [8] it is difficult to measure because of the low detection efficiency. The most interesting decay is the $B^0_s \to \mu^+\mu^-$ which can be mediated by the Feynman diagrams in Figure 1.7. Thanks to the latest

---

[8] The leptonic decays of the B^0_d meson have smaller branching ratio because of Cabibbo suppression.

measurement of the B^0_s oscillation frequency ($\Delta M_s = 17.8 \pm 0.1 ps^{-1}$) the theoretical prediction of this branching ratio within the SM is very precise ($Br(B^0_s \rightarrow \mu^+\mu^-) = (3.4 \pm 0.4) \cdot 10^{-9}$). Moreover the branching ratio of the $B^0_s \rightarrow \mu^+\mu^-$ channel within the MSSM is proportional to the Higgs vacuum expectation value ($\tan\beta$) at the sixth power. Therefore this decay is very sensitive to SUSY models with high $\tan\beta$ value. In the context of the CMSSM the branching ratio of $B^0_s \rightarrow \mu^+\mu^-$ and the anomalous magnetic moment of the muon $g_\mu - 2$ are function of few parameters, i.e. the gaugino mass and $\tan\beta$.



(a) (b)

(c)

Figure 1.7: Feynman diagrams for the $B^0_s \rightarrow \mu^+\mu^-$ decay within the SM.

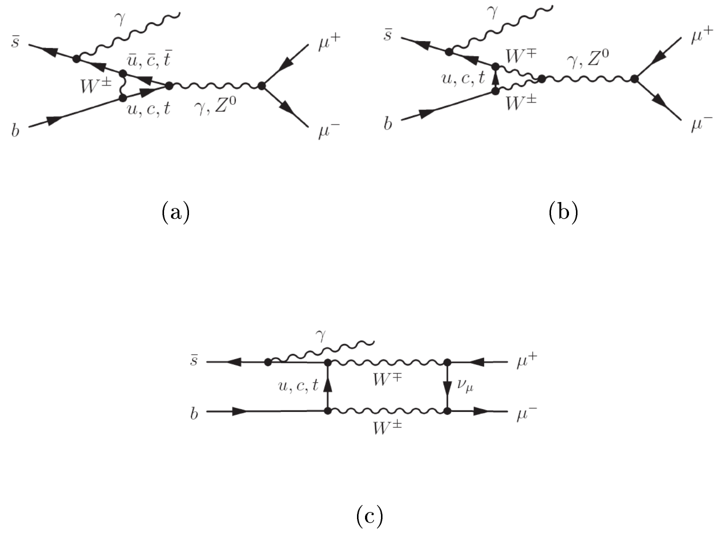
The best limit at present on the $B^0_s \rightarrow \mu^+\mu^-$ decay is put by the CDF experiment ($Br(B^0_s \rightarrow \mu^+\mu^-) < 5.8 \cdot 10^{-8}$ @95%CL).

However it is expected that this limit will be significantly improved by the LHCb experiment, which will be able to measure the SM branching ratio in just one year of running. This measurement is one of the most interesting early measurement for the LHCb experiment. The LHCb sensitivity for this decay was already studied in detail in [38].

This analysis is not subject of this thesis.

1.6.2 The $B^0_{s,d} \to \mu^+\mu^-\gamma$ decay

Another interesting decay is the $B^0_s \to \mu^+\mu^-\gamma$ channel. This channel proceeds via the Feynman diagrams of Figure 1.8. Even if the diagrams for the $B^0_s \to \gamma\mu^+\mu^-$ decay have one more vertex with respect to those of the $B^0_s \to \mu^+\mu^-$, the branching ratio for the two channels is predicted to be of the same order of magnitude. This is due to the fact that the presence of the photon removes the helicity suppression.



(a) (b)

(c)

Figure 1.8: Feynman diagrams for the $B^0_s \to \mu^+\mu^-\gamma$ decay within the SM.

However, in this case, because of the photon radiation, an additional form factor must be taken into account. This calculation can be carried out using various approaches, which invoke different approximations. All these calculations ([45], [46], [47] and [42]) give an estimation for the branching ratio of the $B^0_s \to \gamma\mu^+\mu^-$ decay in the range $(1 \div 5) \cdot 10^{-9}$, when long distance contribution are not taken into account.

In particular the most complete calculation of the branching ratio and of the dimuon invariant mass spectrum within the SM was carried out by N.

Nikitin and D. Melikhov in [42].

The spectrum for the square of the dimuon invariant mass (q^2) as computed [42] in is shown in Figure 1.9. Recently new physics effects in the

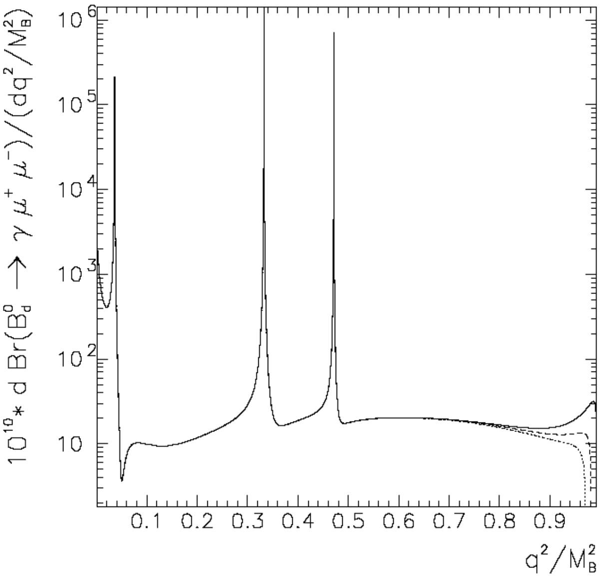


Figure 1.9: Dimuon q^2 spectrum for the $B^0_s \to \gamma\mu^+\mu^-$ decay.

$B^0_s \to \mu^+\mu^-\gamma$ channel were studied with a model independent analysis [43] and in the context of specific NP models as well [44]. This branching ratio can be computed through the *Operator Product Expansion* (OPE), by the effective hamiltonian:

$$H_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb}V^*_{ts} \sum C_i(\mu)O_i(\mu), \tag{1.42}$$

where the O_i are local operators, V_{ij} are CKM matrix elements, μ is the renormalization scale and the coefficients C_i are known as Wilson coefficients. New physics can contribute to the branching ratio of $B^0_s \to \mu^+\mu^-\gamma$ either changing the Wilson coefficients already present in the SM or through new operators. The effective Hamiltonian of the $b \to ql^+l^-$ can be written in terms of twelve model-independent four-Fermi operators [43] as in Equation 1.43.

$$H_{eff} = \quad \frac{G_\alpha}{\sqrt{2}\pi} V_{tq}V^*_{tb}\{C_{SL}\bar{q}i\sigma_{\mu\nu}\frac{q^\nu}{q^2}Lb\bar{l}\gamma^\mu l + C_{BR}\bar{q}i\sigma_{\mu\nu}\frac{q^\nu}{q^2}Rb\bar{l}\gamma^\mu l$$

$$
\begin{aligned}
&+C_{LL}^{tot}\bar{q}_L\gamma_\mu b_L \bar{l}_L\gamma^\mu l_L + C_{LR}^{tot}\bar{q}_L\gamma_\mu b_L \bar{l}_R\gamma^\mu l_R + C_{RR}^{tot}\bar{q}_R\gamma_\mu b_R \bar{l}_L\gamma^\mu l_L \\
&+C_{RR}\bar{q}_R\gamma_\mu b_R \bar{l}_R\gamma^\mu l_R + C_{LRLR}\bar{q}_L b_R \bar{l}_L l_R \\
&+C_{LRRL}\bar{q}_L b_R \bar{l}_R l_L + C_{RLRL}\bar{q}_R b_L \bar{l}_R l_L \qquad (1.43) \\
&+C_T\bar{q}\sigma_{\mu\nu} b \bar{l}\sigma^{\mu\nu} l + iC_{TE}\epsilon^{\mu\nu\alpha\beta}\bar{q}\sigma_{\mu\nu} b \bar{l}\sigma_{\alpha\beta}\}
\end{aligned}
$$

In Figure 1.10 the behavior of Wilson coefficients as a function of the branching ratio $B_s^0 \to \mu^+\mu^-\gamma$ is shown [43].

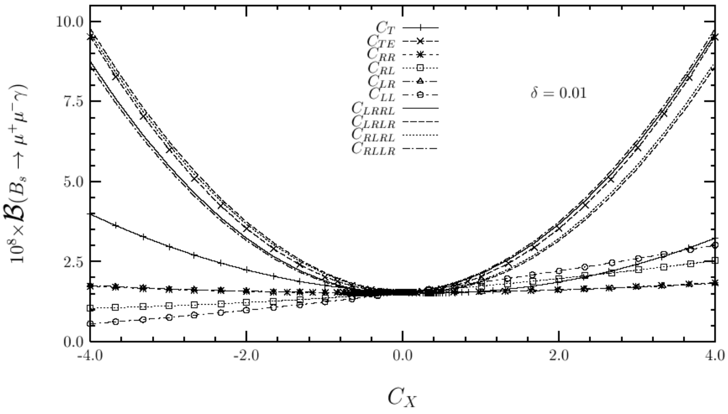


Figure 1.10: The dependence of branching ratio for the $B_s^0 \to \gamma\mu^+\mu^-$ decay on the new Wilson coefficients. A cut on $\delta = 0.01$ in the photon energy was imposed.

In [44] it was also shown that the contribution of theories beyond the SM involving enhanced Z-penguin can significantly enhance this branching ratio, the possible enhancements are listed in Table 1.4 for two different form factors.

Moreover, $Br(B_s^0 \to \gamma\mu^+\mu^-)$ can also be significantly enhanced in some SUSY models like the two Higgs doublets model (2HDM) and mSUGRA. Figure 1.11 shows $Br(B_s^0 \to \mu^+\mu^-\gamma)$ as a function of $\tan\beta$ for different values of charged Higgs doublets mass. At present there is no measurement available on the branching ratio of the $B_s^0 \to \mu^+\mu^-\gamma$ decay. The world's best limit on the $B_d^0 \to \mu^+\mu^-\gamma$ decay is set by the BaBar experiment [48] ($Br(B_d^0 \to \mu^+\mu^-\gamma) < 1.7 \cdot 10^{-7}$ @ 90% CL).

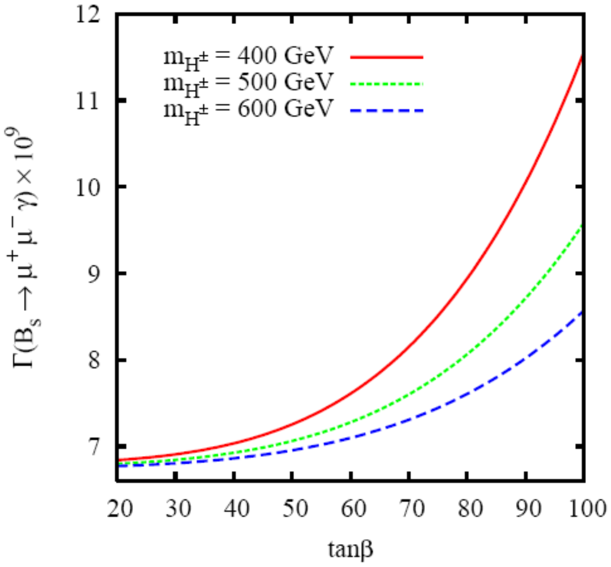


Figure 1.11: The branching ratio for the $B_s^0 \to \gamma\mu^+\mu^-$ decay in the 2HDM as a function of $\tan\beta$. The curves are obtained for different Higgs mass values [44].

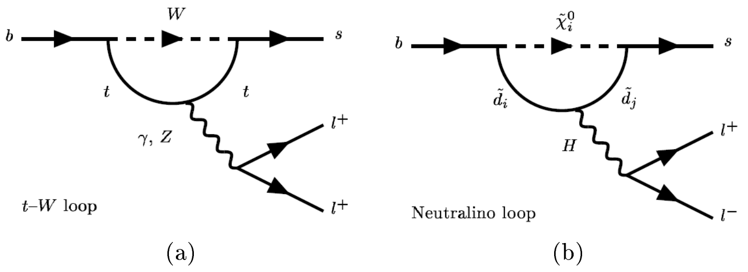
1.7 New Physics in the $B_d^0 \to K^{*0}\mu^+\mu^-$ channel

In Figure 1.12 the Feynman diagrams of the SM contributions (Electroweak penguins) and a possible SUSY contribution (Neutralino-loop) to this decay are shown. In [49] a model independent study via the OPE was performed, showing that the angular distribution in the kinematical limit of low di-muon invariant mass is sensitive to new CP sources and right-handed currents. Within the SM there are not right-handed currents and the only source of CP violation is due to a complex phase in the CKM matrix. However there are many extensions of the SM which predict new sources of CP violation due to the interference of new particles which may enter in the loops. Moreover right-handed currents arise in many SM extensions, as in theories involving the $SU(2)_L \times SU(2)_R$ gauge symmetry group.

The $B_d \to K^{*0}\mu^+\mu^-$ decay is completely described by the following formula

Table 1.4: The branching ratio for the $B_s^0 \to \gamma\mu^+\mu^-$ decay calculated within the SM and the bsZ model. The calculation was done in [44] using two different form factors: Kruger and Melikov's form factors (K and M) and Dincer and Sehgal's form factors (D and M).

Model	$Br(B_s^0 \to \gamma\mu^+\mu^-) \cdot 10^{-9}$
SM (K and M)	1.766
bsZ (K and M)	6.68
SM (D and S)	2.94
bsZ (D and S)	10.47



(a) (b)

Figure 1.12: Feynman diagrams for the $B_d^0 \to K^{*0}\mu^+\mu^-$ decay. Diagram (a) is one of the SM contributions (electroweak penguin). Diagram (b) is a possible SUSY contribution (neutralino loop).

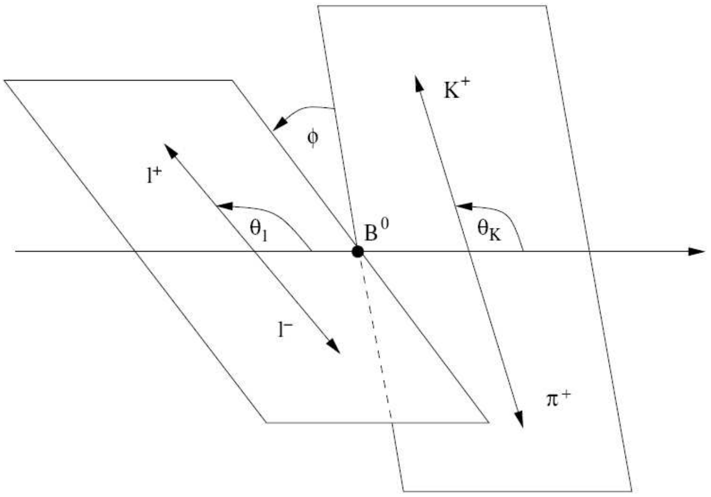


Figure 1.13: Kinematic variables for the $B_d^0 \to K^{*0}\mu^+\mu^-$ decay.

of the differential decay rate:

$$\frac{d^4\Gamma}{dsdcos\theta_K dcos\theta_l d\Phi} = \frac{9}{32\pi}\sum_{i=1}^{9} I_i(s,\theta_K)\cdot f_i(\theta_l,\Phi), \tag{1.44}$$

where θ_K, θ_l and Φ are three angles defined in the following which describe completely the angular distribution of the decay products and the variable s is the dimuon invariant mass.

Considering the $B_d \to K^{*0}\mu^+\mu^-$ decay (Figure 1.13), denoting with $\overrightarrow{p}$ the three-momentum measured in the B_d^0 rest frame, with $\overrightarrow{q}$ the three-momentum measured in the dimuon rest frame, with $\overrightarrow{r}$ the three-momentum measured in the K^{*0} rest frame, being the $z-axis$ the direction of the K^{*0} in the B_d^0 rest frame we define:

$$\begin{aligned}
&\overrightarrow{e}_z = \frac{\overrightarrow{p}_{K^+} + \overrightarrow{p}_{\pi^-}}{\mid \overrightarrow{p}_{K^+} + \overrightarrow{p}_{\pi^-} \mid} \quad , \quad e_l = \frac{\overrightarrow{p}_{\mu^+} \times \overrightarrow{p}_{\mu^-}}{\mid \overrightarrow{p}_{\mu^+} \times \overrightarrow{p}_{\mu^-} \mid} \quad , \quad e_K = \frac{\overrightarrow{p}_{K^+} \times \overrightarrow{p}_{\pi^-}}{\mid \overrightarrow{p}_{K^+} \times \overrightarrow{p}_{\mu^-} \mid} \\
&cos\theta_l = \frac{\overrightarrow{q}_{\mu^+} \cdot \overrightarrow{e}_z}{\mid \overrightarrow{q}_{\mu^+} \mid} \quad , \quad cos\theta_K = \frac{\overrightarrow{r}_{K^+} \cdot \overrightarrow{e}_z}{\mid \overrightarrow{q}_{K^+} \mid} \\
&sin\Phi = (\overrightarrow{e}_l \times \overrightarrow{e}_K)\cdot \overrightarrow{e}_z \quad , \quad cos\Phi = \overrightarrow{e}_K \cdot \overrightarrow{e}_l .
\end{aligned} \tag{1.45}$$

For the $\overline{B}_d^0 \to \overline{K}^{*0}\mu^+\mu^-$ decay the angles are measured with respect to the μ^- and the K^+. Therefore θ_K is the angle between the $K^-(K^+)$ and the

$\overline{B}^0_d(B^0_d)$ direction in the $\overline{K}^{*0}(K^{*0})$ rest frame, θ_l is the angle between the $\mu^+(\mu^-)$ and the $\overline{B}^0_d(B^0_d)$ in the dimuon rest frame and the Φ is the angle between the $\overline{K}^{*0}(K^{*0})$ decay plane and the dimuon plane in the $\overline{B}^0_d(B^0_d)$ rest frame (Figure 1.13). The angles are therefore defined in the intervals:

$$0 \leq \theta_l \leq \pi \quad , \quad 0 \leq \theta_K \leq \pi \quad , \quad -\pi \leq \Phi \leq +\pi. \tag{1.46}$$

Even if this channel suffers from the theoretical uncertainties related to the heavy-to-light form factors, these uncertainties can be greatly reduced in the limit in which the final meson has a large energy in the $B^0_d(\overline{B}^0_d)$ rest frame. The functions I_{1-9} can be written in terms of the amplitudes A_j (with $i = t,\ 0, \|,\ \perp$) which have both left-handed and right-handed components. The A_t amplitude correspondent to the scalar component of the off-shell K^{*0} is negligible in the kinematical limit in which the muon mass is small with respect to the dimuon invariant mass and the K^{*0} is almost on-shell. The A_j amplitudes can be expressed as a function of the helicity amplitude as in Equation 1.47.

$$A_{\perp,\|} = \frac{H_{+1} \mp H_{-1}}{\sqrt{2}} \quad , \quad A_0 = H_0 \quad , \quad A_t = H_t \tag{1.47}$$

The I_{1-9} function are explicitly written as a function of their left and right-handed components. The differential decay rate of Equation 1.44 has 10 free parameters: 6 complex quantities, in which the 2 phases between left-handed and right-handed components can be arbitrarily fixed as they don't interfere with each other.

$$\begin{aligned}
I_1 = \quad & \frac{3}{4}(\mid A_{\perp L} \mid^2 + \mid A_{\| L} \mid^2) \sin^2\theta_K + \\
& + \mid A_{0L} \mid^2 \cdot \cos^2\theta_K + (L \to R) && (1.48) \\
I_2 = \quad & \frac{1}{4}(\mid A_{\perp L} \mid^2 + \mid A_{\| L} \mid^2) \sin^2\theta_K - \\
& - \mid A_{0L} \mid^2 \cdot \cos^2\theta_K + (L \to R) && (1.49) \\
I_3 = \quad & \frac{1}{2}(\mid A_{\perp L} \mid^2 - \mid A_{\| L} \mid^2) \sin^2\theta_K + (L \to R) && (1.50) \\
I_4 = \quad & \frac{1}{\sqrt{2}}\Re(A_{0L}A^*_{\| L}) sin 2 + (L \to R) && (1.51) \\
I_5 = \quad & \sqrt{2}\Re(A_{0L}A^*_{\| L}) \sin 2 - (L \to R) && (1.52) \\
I_6 = \quad & 2\Re(A_{\| L}A^*_{\perp L}) \sin^2 - (L \to R) && (1.53)
\end{aligned}$$

$$I_7 = \sqrt{2}\Im(A_{0L}A^*_{\parallel L})\sin 2 - (L \to R) \tag{1.54}$$

$$I_8 = \frac{1}{\sqrt{2}}\Im(A_{0L}A^*_{\parallel L})\sin 2 + (L \to R) \tag{1.55}$$

$$I_9 = \Im(A^*_{\parallel L}A_{\perp L})\sin^2 +(L \to R) \tag{1.56}$$

In this decay it is possible to define some observables which are predictable within the SM and are sensitive to NP [49], [50], [51], [52], [53]. The most promising is the forward-backward asymmetry (FBA), which is the asymmetry between the $\mu^+(\mu^-)$ going forwards and backwards with respect to the $B_d^0(\overline{B}_d^0)$ direction in the dimuon rest frame. This asymmetry was computed in the SM as a function of the dimuon invariant mass square (s) (Figure 1.14). An observable particularly sensitive to NP is the zero crossing point of the FBA which is predicted with a very small error. It was shown that the FBA is enhanced by SUSY contributions for large $\tan\beta$ values and that it depends strongly on the sign of the μ parameter of the CMSSM (Figure 1.15). Other interesting asymmetries are $A_T^{(1)}$ and $A_T^{(2)}$ which are defined

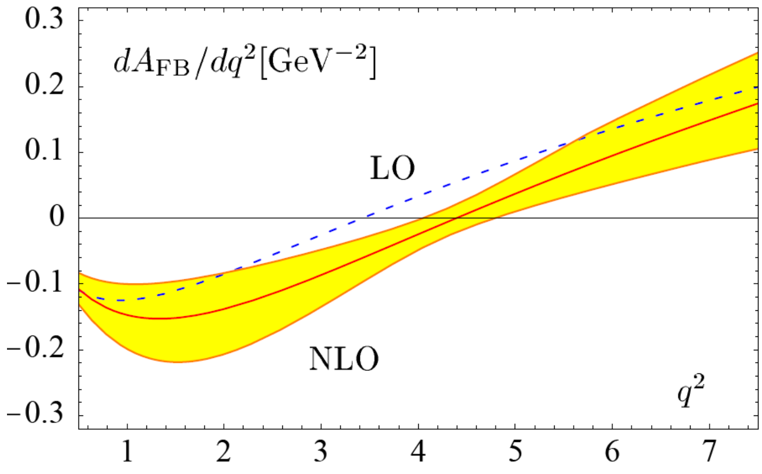


Figure 1.14: FBA versus the dimuon mass square s in the SM. The dotted line is computed at leading order. The continuous line at next to leading order. The yellow band represents the theoretical uncertainty [53].

as[9]:

$$A_T^{(1)} = \frac{-2\Re(A_\parallel A_\perp^*)}{\mid A_\parallel \mid^2 + \mid A_\perp \mid^2} \tag{1.57}$$

$$A_T^{(2)} = \frac{\mid A_\perp \mid^2 - \mid A_\parallel \mid^2}{\mid A_\parallel \mid^2 + \mid A_\perp \mid^2} \tag{1.58}$$

$$A_{FB} = \frac{3}{2} \frac{\Re(A_{L\parallel} A_{L\perp}^*) - \Re(A_{R\parallel} A_{R\perp}^*)}{\mid A_0 \mid^2 + \mid A_\parallel \mid^2 + \mid A_\perp \mid^2} \tag{1.59}$$

The quantities $A_T^{(1)}$ and $A_T^{(2)}$ are predicted with a small theoretical error.

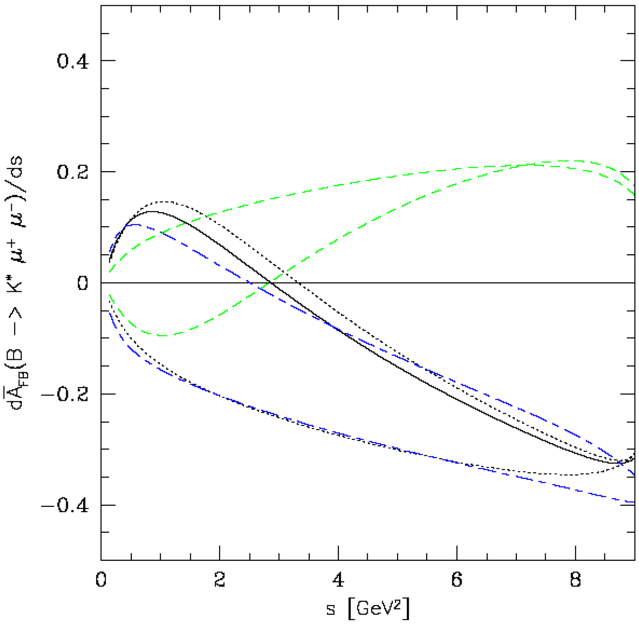


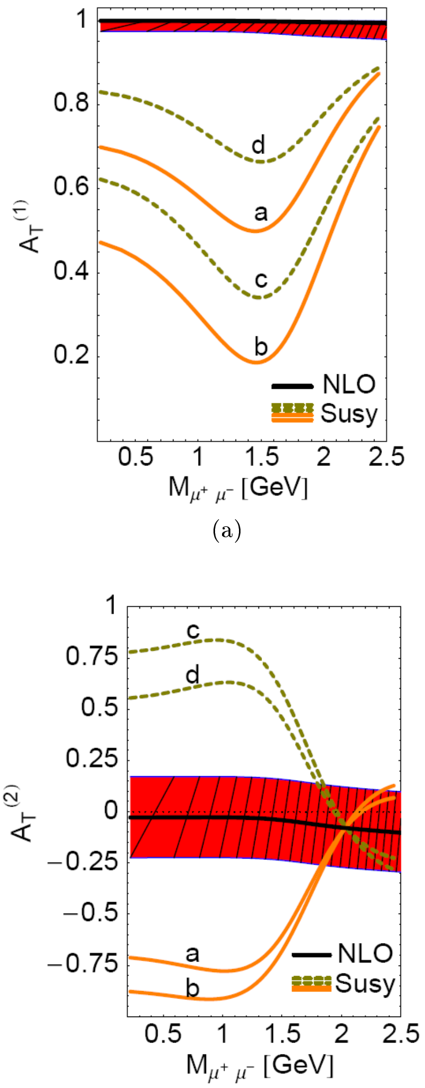
Figure 1.15: FBA versus s as predicted in some SUSY models and in the SM (solid line) for the $B_d^0 \to K^{*0}\mu^+\mu^-$ decay [50].

In [49] it was shown that SUSY contributions generally change the values of these asymmetries. In Figure 1.16 the asymmetries $A_T^{(1)}$ and $A_T^{(2)}$ as a function of the variable s are plotted. The black line is the SM prediction, the band represents the theoretical uncertainty. The green and yellow lines are predictions of SUSY models. It is important to stress that the asymmetry values are significantly changed also by SUSY models with low $tan(\beta)$ value, this makes this measurement complementary to

---

[9] For completeness FBA is also written as a function of the amplitudes A_j.

most of the important channels for the indirect search of NP, such as the $B_s^0 \to \mu^+\mu^-$ decay. The $B_d^0 \to K^*\mu^+\mu^-$ decay was already studied by the BaBar and Belle experiments ([56] and [57]). Its branching ratio is $Br(B_d^0 \to K^*\mu^+\mu^-) = (0.88^{+0.33}_{-0.29} \pm 0.10) \cdot 10^{-6}$.
The FBA was measured with a large uncertainty due to lack of statistics. The present sensitivity to the zero crossing point of the FBA is no sufficient to discriminate between the SM and other NP theories.



(a)

(b)

Figure 1.16: $A_T^{(1)}$ (Figure (a)) and $A_T^{(2)}$ (Figure (b)) as a function of the dimuon mass. The black line is the next to leading order prediction within the SM. The band represents the theoretical uncertainty. The other lines are predictions of SUSY models with low $\tan\beta$ value [49].

Chapter 2

The LHCb Experiment

2.1 Beauty production at LHCb

The p-p interactions are dominated by QCD effects. There are long-distance and short distance effects. The long distance interaction is a non-perturbative interaction in which the protons see each other as point like particles without structure. The scattering is known as “soft”, as the transferred momentum is low, the out-coming particles are therefore produced at small polar angles with respect to the beam axis. This process is not relevant for b-hadron production.

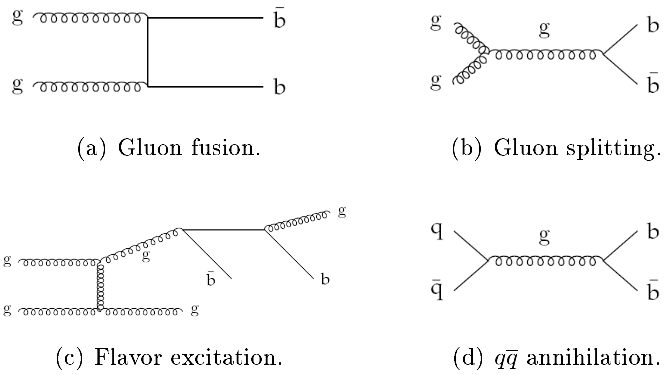
In short distance interactions the interacting particles are the partons of the incoming protons. In this case the interaction is described by the perturbative QCD followed by a non perturbative hadronization to colorless hadrons. The transferred momentum is large, therefore the out-coming particles are produced with large transverse momentum with respect to the beam axis. At LHCb the mean number of particles per events is about 300, the mean

Table 2.1: Cross section of various process in pp collisions at 14TeV at LHCb.

Total cross-section	$\sigma_{tot} = 100mb$
Inelastic cross-section	$\sigma_{in} = 5.5mb$
$c\bar{c}$ Production	$\sigma_{c\bar{c}} = 3.5mb$
$b\bar{b}$ Production	$\sigma_{b\bar{b}} = 500\mu b$
$Z^0W^{\pm}, Z^0$	$\sigma_{Z^0W^{\pm}} = 38pb$
$\gamma^*Z^0 + Jets$	$\sigma_{\gamma^*Z^0+Jets} = 73nb$
$W^{\pm} + Jets$	$\sigma_{Z^0+Jets} = 160nb$

Table 2.2: Production fraction for b-hadrons.

B^0, $B^\pm$	$39.8 \pm 1.0\%$
B_s^0	$10.4 \pm 1.4\%$
b baryons	$9.9 \pm 1.7\%$



(a) Gluon fusion.

(b) Gluon splitting.

(c) Flavor excitation.

(d) $q\bar{q}$ annihilation.

Figure 2.1: Feynman diagrams for the b-hadron production at LHCb.

charged stable multiplicity is about 100, of which about 20 are inside the LHCb pseudorapidity range.

In Table 2.1 the expected cross sections, for some processes as calculated with Pythia, are listed.

Among these processes there are the Z^0 and $W^\pm$ production, which have a small cross section if compared with the $b\bar{b}$ and $c\bar{c}$ cross sections and Jets emission.

Jets are originated by partons which scatter with very high transverse momentum, and then hadron showers are emitted. The study of jets reconstruction is important for the detection of light Higgs particles in LHCb.

For the LHCb experiment the most relevant processes consist of events which contain b and c hadrons. The trigger system (see [70]) is in fact designed to enhance these events, in particular $b\bar{b}$ events, with respect to minimum bias events.

The dominant Feynman diagrams for $b\bar{b}$ production at LHCb are: **gluon fusion**, **quark-antquark annihilation**, **gluon splitting** and **flavor excitation**. These processes are shown in the diagrams in Figure 4.14.

However, the production mechanism is affected by large theoretical uncertainties due to the phenomenological description of the next-to-leading order contributions.

At the LHC energy, it is unlikely that the two scattering partons will have similar momentum, but most likely, they will scatter with large differences in momentum.

As the mass of the $b\bar{b}$ quark pair is small with respect to the typical parton center of mass energy, the gluon is radiated with high momentum. Therefore b-hadrons are produced correlated and predominantly both forwards or backwards (Figure 2.2). The b and $\bar{b}$ quarks are then produced in the same forward (or backward) cone with a typical correlation of one unit in rapidity.

At the tree level, the cross section of the two processes $q\bar{q} \to Q\overline{Q}(1\%)$ and $gg \to Q\overline{Q}(99\%)$ at LHC is given by the formula [71]:

$$\frac{d^4\sigma}{dy_Q dy_{\overline{Q}} d^2 p_T} \propto \frac{1}{64\pi^2 m_T^4 (1 + \cosh(\Delta y))^2}, \tag{2.1}$$

where $m_T^2 = m^2 + p_T^2$ and where y indicates the rapidity of the out-coming quarks. The cross section is therefore suppressed for large p_T values and for large differences in rapidity between the two out-coming quarks, this justifies the choice of the forward geometry for the LHCb experiment.

In Table 2.2 the production fractions for the various b-hadrons, as evaluated with Pythia, are listed.

2.2 The LHCb Spectrometer

The LHCb (Large Hadron Collider Beauty Experiment) is a single arm forward spectrometer. It is the only experiment of the LHC completely devoted to the study of b-hadron physics, in particular to the measurement of CP violation and rare decays of b-hadrons.

The layout of the detector is shown in Figure 2.2, Its dimensions are (x = 6 m) x (y = 5 m) x(z = 20 m). It covers an angular range between $10mrad$ and $300mrad$ in the bending plane and between $10mrad$ and $250mrad$ in the non-bending plane, or equivalently a range of pseudorapidity of $1.9<\eta<4.9$.

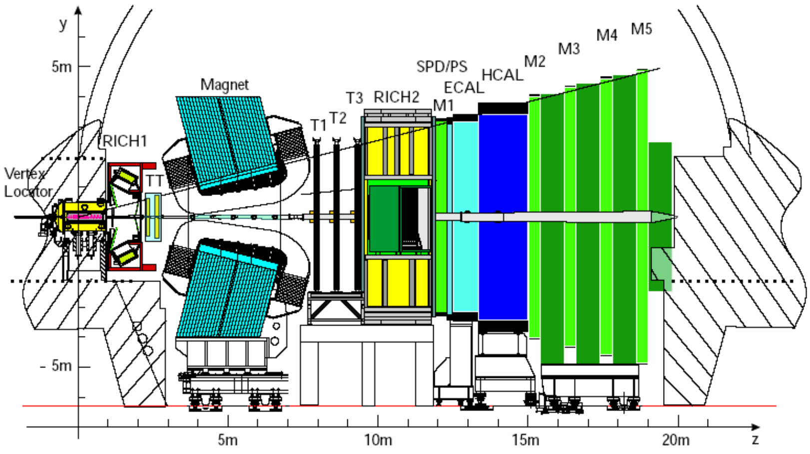


Figure 2.2: Side view of the LHCb detector (non-bending plane). The LHCb coordinate system is also shown.

The choice of the forward geometry is motivated by the fact that at high energy $b\bar{b}$ pairs are emitted correlated and predominantly forwards or backwards as shown in Figure 2.3 .
The LHCb coordinate system is right-handed and defined so that the z-axis points towards the muon station and the y axis points upwards. In this system the main magnetic field component is along the y axis.
The LHCb experiment will operate at a luminosity of $2 \times 10^{32} cm^{-2} s^{-1}$, a reduced luminosity with respect to the maximum design luminosity of the LHC. At this luminosity the events are dominated by a single $p - p$ interaction as shown in Figure 2.4. This implies an easier analysis and event reconstruction, a lower detector occupancy and a reduced radiation damage.
The LHCb apparatus consists of five sub-detectors:
-Vertex Locator Detector (VELO);
-Tracking System (TT, Tracking Stations T1, T2 and T3);
-Ring Imaging Cherenkov detectors (RICH1, RICH2);

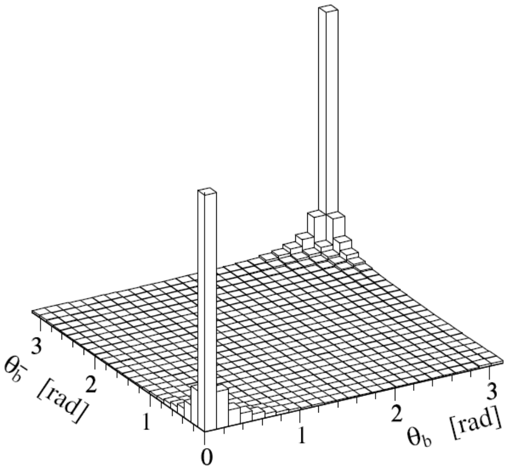


Figure 2.3: LEGO plot of the polar angle of b (θ_b) and $\bar{b}$ (θ_b) obtained with Pythia in the $p-p$ interaction at 14TeV. The $b\bar{b}$ pair is emitted predominantly forwards or backwards.

-Electromagnetic and Hadronic Calorimeters (ECAL, HCAL);

-Muon System (M1, M2, M3, M4 and M5).

The detector is traversed by the beam pipe. The beam pipe must be mechanically very stiff to withstand the pressure difference between the Ultra High Vacuum in the region of the LHCb detector and the ambient, but it must also be thin enough to minimize the number of particles from secondary interactions. For this purpose the chosen material was beryllium.

A dipole magnet is located close to the interaction point and is used to bend the trajectories of charged particles and allow the determination of the momentum from the deflection in the magnetic field. The magnet has a bending power of $\int Bdl \sim 4Tm$ and a non-uniformity below 5% . Figure 2.5 shows the behavior of B_y with respect to the z-axis. The choice of a warm magnet instead of a super-conducting magnet is due to the possibility of reversing the polarity in order to study systematics induced by possible left-right asymmetries in the detector.

The different parts of the detector will be briefly described in the remainder of this chapter.

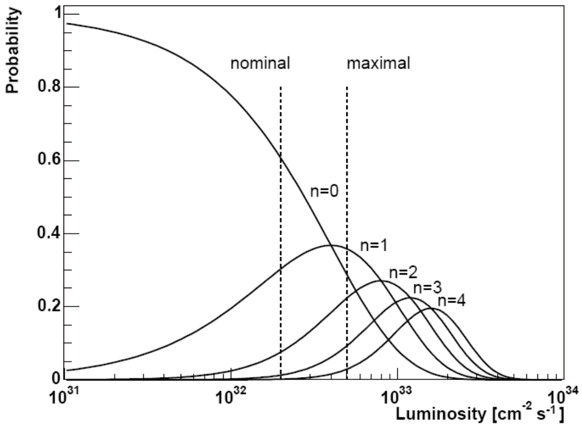


Figure 2.4: Probability of having n interactions as a function of the luminosity. The nominal and maximum luminosities are indicated by the dashed lines.

2.3 The Vertex Locator Detector (VELO)

The Vertex Locator (VELO) is one of the tracking subsystems of LHCb (Figure 2.6). It is a fundamental device to reconstruct both primary and secondary vertexes, as it provides precise measurement of the tracks close to the interaction point. The presence of secondary vertexes is an important signature which allows us to enrich the content of B-mesons in the data.

As the $c\tau$ of B-mesons is about $400\mu m$ and the average momentum is about $100GeV$, the distance between Primary Vertex (PV) and the B-meson decay secondary vertex is about 1cm, therefore a sub-mm resolution is needed to resolve secondary vertexes. Moreover to measure $B_s - \overline{B}_s$ oscillations, which are very fast ($\Delta M_s = (17.8 \pm 0.1)ps^{-1}$), a proper time resolution of 40 fs is required and therefore a very good secondary vertex resolution. The on-line secondary vertex resolution is expected to be $80\mu m$, this precision increases up to $40\mu m$ in the off-line analysis.

The LHCb vertex detector consists of 21 silicon stations along the beam direction. This detector has to operate in a very high radiation environment which is strongly non uniform in the radial direction.

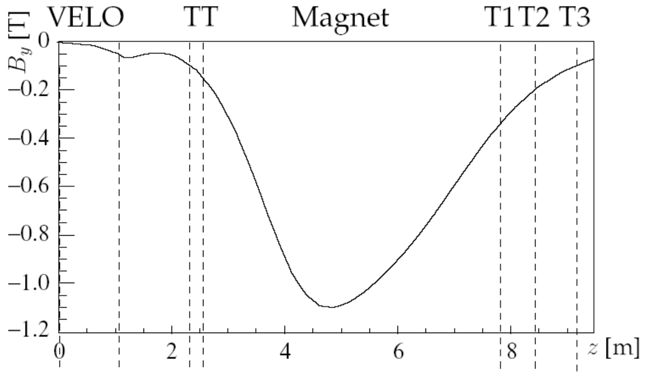


Figure 2.5: The B_y component of the magnetic field strength as a function of the z-coordinate.

The VELO detector is enclosed inside a large vacuum vessel. However, to protect the primary LHC vacuum allowing the sensor revoval, the detector stations are placed in an aluminium secondary vacuum container. The secondary vacuum cantainer must be radiation resistant and it must act as a wake field suppressor as well. In fact, the beam bunches, passing through the VELO, induce wake fields as a consequence of the geometrical changes and of the finite restivity of the wall materials. The wake field can affect both the VELO system (RF pick-up) and LHC beams.

In order to protect the sensors from RF currents, an aluminium alloy foil were designed. It is corrugated to dissipate the induced currents and minimize the material seen by the particles. However this shielding might be insufficient to protect the detector against high-frequency pick-up noise. The wake field suppression is achieved by enclosing the silicon stations in an aluminium box and ensuring that a continuous conductive surface guides the mirror charges from one end of the VELO vessel to the other.

For safety reasons the closest approach allowed to the beam axis is about 5mm, however considering RF-shield and the guard-ring on the silicon the sensitive area starts at about $8mm$ from the beam axis. This distance is

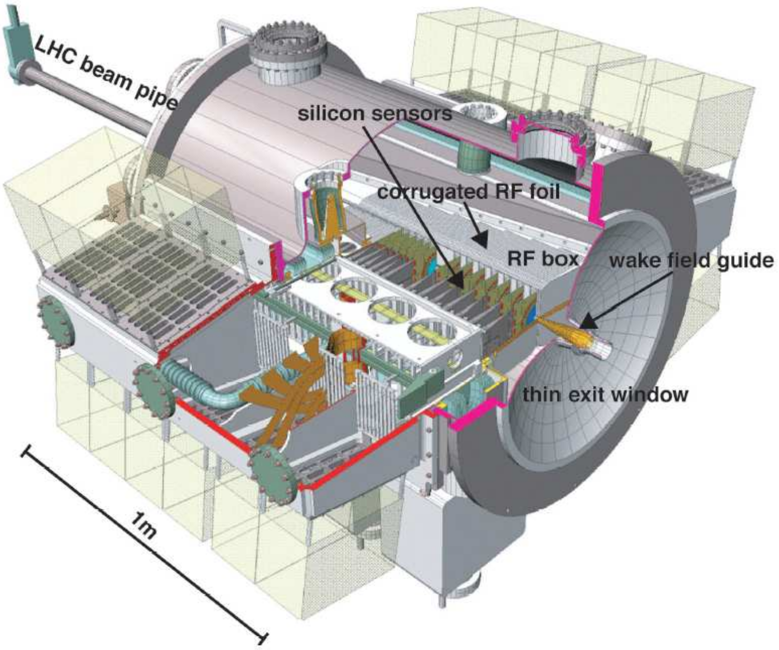
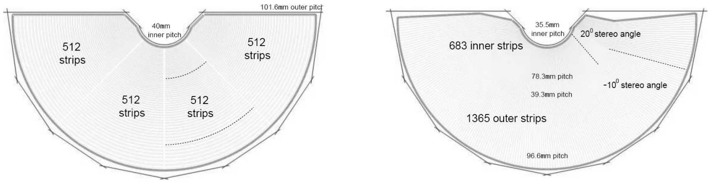


Figure 2.6: 3D view of the VELO vacuum vessel.

smaller than the aperture required by LHC during the injection to avoid radiation damage, hence the stations must be retractable. This is achieved by putting the whole system inside a vacuum vessel with a roman-pot to enable the removal of the two VELO halves.

The VELO active elements consist of 21 silicon disks spread along the beam line. Each station is constituted by a R-measuring sensor with circular strips centered around the beam axis (Figure 4.3(a)) and a Φ-measuring sensor with radial strips (Figure 4.3(b)).

The sensors are $220\mu m$ n-on-n single side silicon strips bounded back-to-back. The VELO is optimized to have the best impact parameter resolution. In order to select clean events in the $L0$ Trigger the VELO has also a **pile-up veto**. It consists of two R-sensors stations upstream the interaction point and it is used to detect the track multiplicity in the backward region and the number of PVs within the same bunch crossing (Figure 2.8). The VELO is refrigerated by a cooling block that uses CO_2 as refrigerant in orders to



(a) R-Sensor layout (b) Φ-Sensor layout

Figure 2.7: Schematic view of the R-sensor a) and of the Φ-sensor b).

avoid thermal and radiation damages.

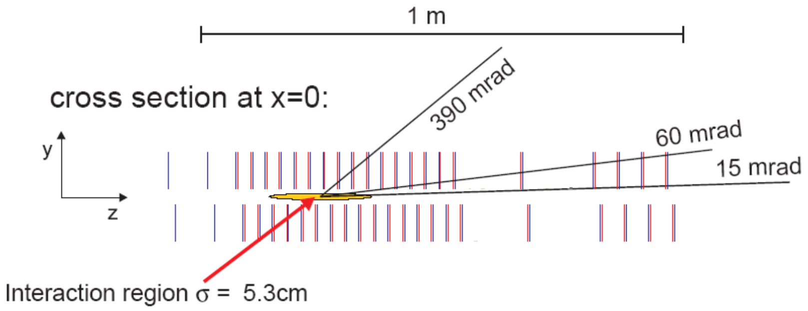


Figure 2.8: VELO stations in the $y-z$ plane. The pile-up veto is constituted of the two single planes at the smallest z position.

2.4 Tracking

The LHCb tracking system consists of three sub-detectors, whose aim is to reconstruct the tracks of charged particles:
-The VELO (already discussed above);
-The Trigger Tracker (TT);
-The tracking stations T1-T3.

Since the Tracking stations T1-T3 have to deal with different fluxes of particles as a function of the radius, two different technologies were adopted, named inner-tracker (IT) and outer-tracker (OT). A very good momentum resolution, of about $\frac{\delta p}{p} \sim 0.4\%$ is expected. This is necessary to achieve the required invariant mass resolution.

2.4.1 Trigger Tracking (TT)

The Trigger Tracker (TT) is located downstream of the RICH1 detector, before the magnet, at about $2.5m$ from the interaction point. It is fundamental in the reconstruction of **down-stream** and **upstream** particles (see Section 2.8) and it is also used in the online analysis to roughly measure the transverse momentum of particles with a large impact parameter.

The fringe magnetic field between the VELO and the TT (about $15Tm$) induces a few millimeters deflection from a straight line in a particle of momentum of a few GeVs. This is sufficient to measure the momentum with $10-40\%$ accuracy without the use of the other tracking stations.

Since the TT is in front of the magnet, it will deal with fluxes of particles as high as in the VELO. Therefore, a similar silicon based technology was chosen. The TT is also useful in off-line analysis to measure tracks of low momentum which are swept out of the acceptance by the magnet and hence cannot be seen by other tracking detectors, like for instance the low momentum pion from the $D^* \rightarrow D^0\pi$ decay.

Finally, the TT is necessary for the reconstruction of long lived particles, as for instance the K_s^0, which decays after the VELO. The active area of the detector, which covers the nominal acceptance of the spectrometer, is covered by silicon microstrip detectors with a strip pitch of $198\mu m$ and a length of $33cm$. The TT consists of two pairs of four layers: TTa centered at $z=232cm$ and TTb centered around $z=262cm$. The first and the last layer are x-layers with vertical readout strips, the second and third layer, called u/v-layers, are rotated by a stereo angle of ±5 degree respectively.

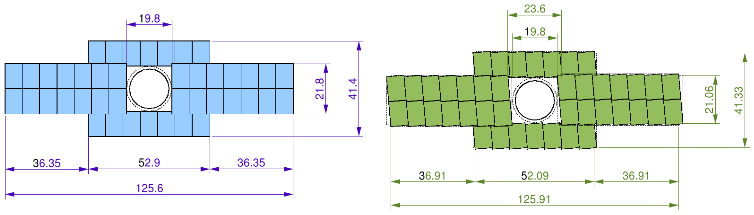
2.4.2 Tracking Stations (T1-T3)

The T1-T3 Tracking Stations are located after the magnet and upstream of the RICH2 detector. The purpose of the tracking stations is to give a precise measurement of particle momenta and to provide space points for long track reconstruction.

The T stations have to provide the direction of tracks for the reconstruction of the Cherenkov ring in the RICH detector and are used as seeds for reconstruction in the Calorimeter and in the Muon Detector for energy measurement and identification. Each station is divided into two parts: Inner Tracker and Outer Tracker.

2.4.3 Inner Tracker (IT)

The Inner Tracker (IT) has an overall sensitive area of about $4.2m^2$. Each station consists of four detector layers. The IT is a silicon microstrip detector which uses $320\mu m$ thick single side p^+-on-n sensors. The strip pitch is approximately $200\mu m$ and the strip length is up to $22cm$.



(a) x-layer (b) u-layer

Figure 2.9: Layout of the x-layer a) and of the u-layer b) in T2. The units are centimeters.

Each station of the IT consists of four detector boxes (Figure 2.9), each of which contains four detection layers. Each layer consists of seven staggered silicon ladders. The 28 ladder in the same detector box, are mounted onto a common cooling plate. The IT covers only about 1.3% of the tracking area, even if approximately 20% of the particles pass through it. In order to obtain high particle reconstruction efficiency and low occupancy, the read-out granularity of the detector has to be matched to the expected particle

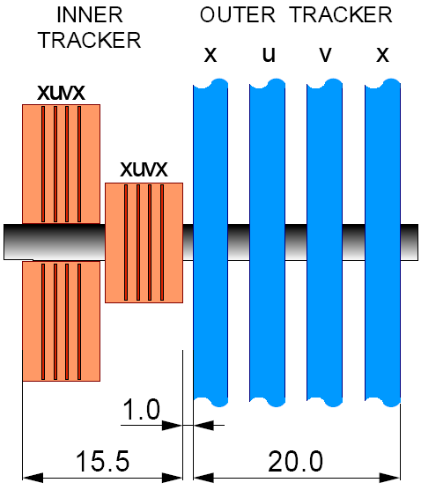


Figure 2.10: Layout of the tracking stations in the $y - z$ plane.

fluxes. Each tracking station has four detection layers with “xuvx”-topology described above for the TT. The IT is positioned upstream of the four detection layers of the OT (Figure 2.10) and the left/right boxes are placed upstream of the top/bottom boxes.

There is overlap between the sensitive areas of the different IT boxes and the adjacent OT. This allows us not only the full acceptance coverage but it is also necessary for alignment using shared tracks. The inner acceptance of the IT is a square around the beam pipe and the size is slightly different for the three stations to follow the shape of the beam pipe.

2.4.4 Outer Tracker (OT)

The rest of the tracking area is covered by the OT (Figure 2.11). The OT technology was chosen to have good spatial resolution and low cost. This is motivated by the fact that the OT covers a larger area than the IT with lower occupancy.

The OT consists of drift cells with a gas mixture of Ar(70%) and CO2(30%) for a drift time below 75 ns. The OT is built from 5mm straw tubes assembled in modules of rows of 64 or 96 straws depending on the region in which they are placed. Each module consists of two layers in which the straws are packed

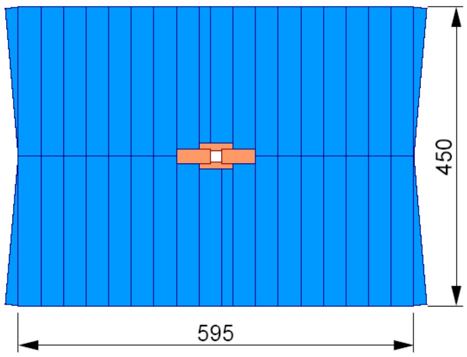


Figure 2.11: Front view of a x-layer of a T-station. The units are centimeters.

in such a way that the spatial resolution is less than the straw radius. For the OT the configuration follows the "xuvx"-topology, as for the IT, with two stereo u/v-layers embraced by two x-layers.

2.5 Ring Imaging Cherenkov (RICH) Detector

RICH is an acronym for Ring Imaging Cherenkov detector. These detectors are based on the known Cherenkov effect that occurs when a high energy charged particle traverses a dielectric medium with a velocity larger than the velocity of light in that medium. When this happens, the particle induces dipoles in the medium and a conical front of Cherenkov light is emitted.

The photons are collected in a spherical mirror with focal length f and focused onto a photon detector placed at the focal plane. The result is a circle of radius $r = f \times \theta_C$ independent on the emission point along the particle track. Measuring the θ_C angle it is possible to obtain the velocity $v = \beta \times c$ (β being the relativistic velocity fraction).

RICH detectors are used for particle identification. In fact, combining the independent measurements of the particle momentum from tracking detectors and the velocity measurements from the RICH detectors, it is possible to extract the mass of the particle and hence its identity. In the LHCb experiment, particle identification is of vital importance in order to make precise b-physics studies of channel which have the same topology but different final states, as for instance the $B \longrightarrow hh$ decays. It is also important for the B-mesons same side flavor tagging to find kaons accompanying the B-mesons. The LHCb experiment has chosen to have two RICH detectors, named RICH1 and RICH2.

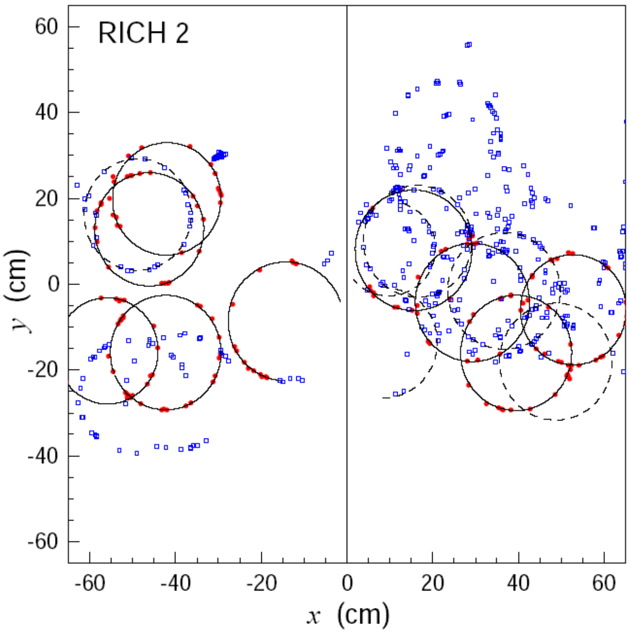


Figure 2.12: Typical RICH2 event with the reconstructed rings superimposed.

The RICH1 is placed between the VELO and the Trigger Tracker. It is far enough from the Magnet to see straight tracks and close enough to the VELO to see low momentum particles that may be swept out by the magnet. It covers a polar angle from $25mrad$ to $330mrad$. The RICH1 is designed to provide particle identification in the range $1GeV - 60GeV$ with two Cherenkov radiators: C_4F_{10} gas with $n = 1.0015$ for mid-momentum tracks and a $5cm$ thick panel made of silica aerogel with $n = 1.03$ suitable

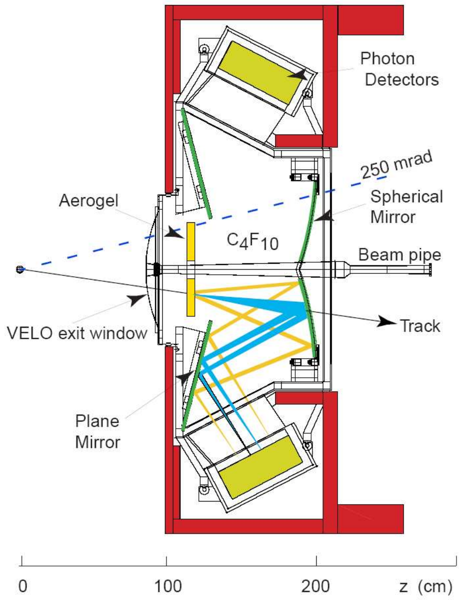


Figure 2.13: Side view of the RICH1 detector.

for low-momentum particles. The RICH2 uses CF4 gas as a single radiator with $n = 1.0005$ and provides particle identification up to $150 GeV$. For both detectors (RICH1 and RICH2) the light is focused onto the photon detector planes with tilted spherical mirrors and secondary plane mirrors. The RICH1 spherical mirrors have a radius of curvature of $2.400 m$. Each of the four mirrors can be individually adjusted in angle. The photon detection is done by pixel hybrid photon detectors (HPDs), they provide good spatial resolution and fast response time, and they are able to do single photon detection.

In Figure 2.12 a typical RICH2 event, with reconstructed rings superimposed, is shown. The RICH1 dimensions are $x = 3m$, $y = 2m$, $z = 1m$ and the optics are arranged vertically at the top and at the bottom, Figure 2.13. The RICH2 dimensions are $x = 7m$, $y = 7m$, $z = 2m$ and its optics are arranged horizontally on both sides of LHCb.

2.6 Calorimeters

The purpose of the calorimeters is to identify electrons, photons and hadrons and measure their energy. An incident particle interacting with the detector generates a shower, thus almost all the energy of the incident particle is absorbed in the medium. Moreover, calorimeters are the only devices that can record the energy of neutral hadrons and photons. Because of the fact that the energy resolution varies with the law $E^{-1/2}$, the precision is better as the energy increases.

In LHCb, the calorimeters are used not only for electron, photon and hadron identification and energy measurement in the off-line analysis, but also for various trigger algorithms. For this reason they must provide very fast energy signals.

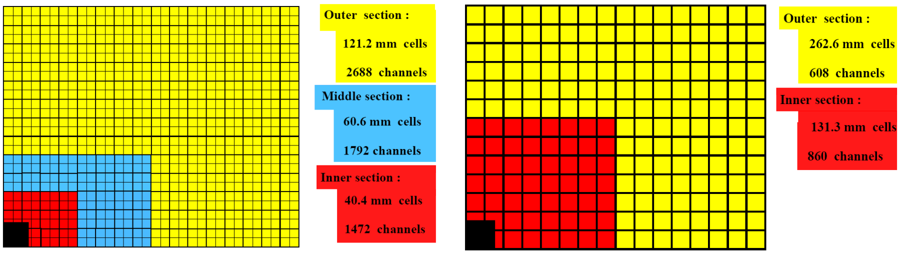
The LHCb calorimeters are placed in the middle of the LHCb detector, downstream of the first muon station (M1) and upstream of the second (M2).

There is an electromagnetic calorimeter (ECAL) followed by a hadron calorimeter (HCAL). The most demanding requirement for the calorimeter is the electron L0 trigger: it requires a rejection of 99% of inelastic pp interaction with an enrichment factor of at least 15 in B events. To reduce the background due to charged pions a preshower is placed upstream of the electromagnetic calorimeter, longitudinally segmenting the ECAL.

The background of the π^0 with large E_T is rejected introducing a scintillator pad detector (SPD) upstream of the PS. The main background at the L0 Trigger level is due to photon conversions in the detector material, which cannot be rejected at this stage. Because of the fact that the background increases rapidly at small angles, to avoid severe radiation damages and occupancy problems, there is a central hole of $30 mrad$ in both directions and in both the calorimeters. The outer limits are $300 mrad$ horizontally and $250 mrad$ vertically.

As the particle rate varies by two orders of magnitude over the calorimeter surface it was natural to choose a variable lateral segmentation. For the SPD, PS and ECAL a segmentation in three different cell-size zones was chosen, defining the so called inner, middle and outer regions (Figure 2.14(a)). For

HCAL two different cell-size zones with different granularity were defined (Figure 2.14(b)).



(a) ECAL, SPD, PS (b) HCAL

Figure 2.14: Schematic view of the transverse segmentation of the LHCb calorimeter system. In Figure a) the SPD, PS and ECAL segmentation is shown. In Figure b) the HCAL segmentation is shown.

2.6.1 Scintillator Pad Detector (SPD)

The SPD allows us to distinguish between charged and neutral particles. It is particularly useful for $e^{\pm}/\ \pi^0$ and $e^{\pm}/\gamma$ shower separation. The SPD consists of $15mm$ thick scintillator pads arranged in a matrix layer. A groove in the scintillator holds helicoidal wavelength shift fibers (WLS) to collect the scintillator light which is sent to photomultipliers located above and below the detector. About 25 photoelectrons are emitted in response to a MIP (minimum ionizing particle).

This allows us a very clean separation between electrons and photons. The discrimination is done by putting a threshold of 0.7 MIP energy within an SPD cell.

2.6.2 Preshower (PS) Detector

The Preshower detector is specifically useful to distinguish charged pions from electrons. It consists of a $12mm$ thick lead wall used as a converter followed by a layer of scintillator pad tiles as in the SPD. Because of the difference in the mean free path for electrons and pions in lead, the electrons produce a shower which starts in the lead whereas the pions do not. A threshold within an SPD cell allows us the discrimination between showers and MIPs.

2.6.3 Electomagnetic Calorimeter (ECAL)

The ECAL has to provide electron and photon identification and their energy measurement as well as π^0 identification. To satisfy LHCb requirements and to maintain a reasonable cost a “Shashlik” technology was chosen. With this technology, a resolution of $\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\%$ can be achieved. The ECAL is made of a sampling structure of $2mm$ thick lead sheets interspersed with $4mm$ thick scintillator plates. The scintillator light is collected as in the SPD and PS by the WLS fibers. To have sufficient energy resolution for high energy photons ECAL is $25X_0$ in depth.

2.6.4 Hadronic Calorimeter (HCAL)

The HCAL is $1.2m$ long, which corresponds at 5.6 interaction lengths. The structure consists of iron/scintillating tiles. The readout is done with WLS fibers. The scintillator and iron plates are parallel to the beam. The sampling structure is made of a $4mm$ scintillator thick and $16mm$ thick iron. The energy resolution can be $\frac{\sigma(E)}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\%$ as confirmed by a Test-beam with pions.

2.7 Muon Detector

The aim of the muon detector is to provide muon identification and fast P_t measurement for the Level 0 trigger. Muon identification is of key importance for LHCb. They are used, for instance, in same side tagging to identify the B-meson flavor. Moreover they are present in many rare decays and in the J/ψ decay.

The muon detector has five stations, called M1-M5 and each station is divided into four regions R1-R4 (Figure 2.15).

The M1 station is placed upstream of the SPD. The M2 station is positioned just after the HCAL. After each station except M1 a 80cm thick iron filter is placed, in order to absorb other particles. Between M1 and M2 the calorimeter also acts as an additional filter. The Muon detector consists of

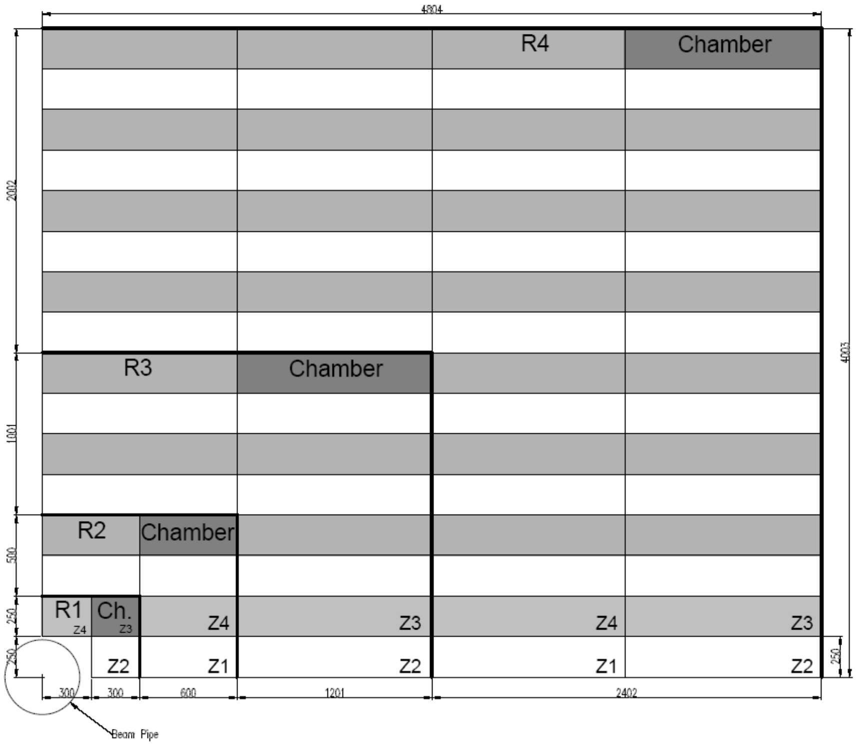


Figure 2.15: Schematic view of a quadrant of a muon station. The regions R1-R4 are indicated. The units are centimeters.

1392 chambers and covers an area of $435m^2$, the acceptance is in the range $20mrad - 306mrad$ in the bending plane and $16mrad - 258mrad$ in the non-bending plane. For all the chambers except for the ones placed in the M1R1 region the technology chosen is the Multi Wire Proportional Chamber (MWPC).

For the M1R1 region a triple-GEM technology, more radiation tolerant, was chosen as it has to cope with higher fluxes of particles. Since the L0 trigger asks for a coincidence between the 5 stations, each station has to have a trigger efficiency of at least 95% within a time window smaller than 25ns. This is essential to unambiguously identify the bunch-crossing. In order to have good time resolution, a fast gas mixture of Ar, CO_2 and CF_4, was chosen and the chambers are composed of two or four OR-ed gas gaps.

The chambers M1-M3 are mainly used for track-finding and momentum computation, while M4-M5 are used for particle identification.
At the $L0$ Trigger level the muon detector provides 20% momentum resolution and a muon indentification efficiency of 55%. This efficiency increases up to 99% in the off-line analysis (with about 4% of pion misidentification probability).

2.8 Track Reconstruction

One of the most important features of the LHCb experiment is good tracking performance, with high efficiency and very good momentum resolution. This is necessary to obtain the required invariant mass resolution. For instance, an invariant mass resolution of 18MeV in the $B_s \to \mu^+\mu^-$ channel allows us to separate up to 3σ the signal from the misidentified background $B_{s,d} \to hh$. The tracking reconstruction algorithm combines the hits of the four tracking detectors : VELO, TT, IT and OT. A reconstructed track consists of a state

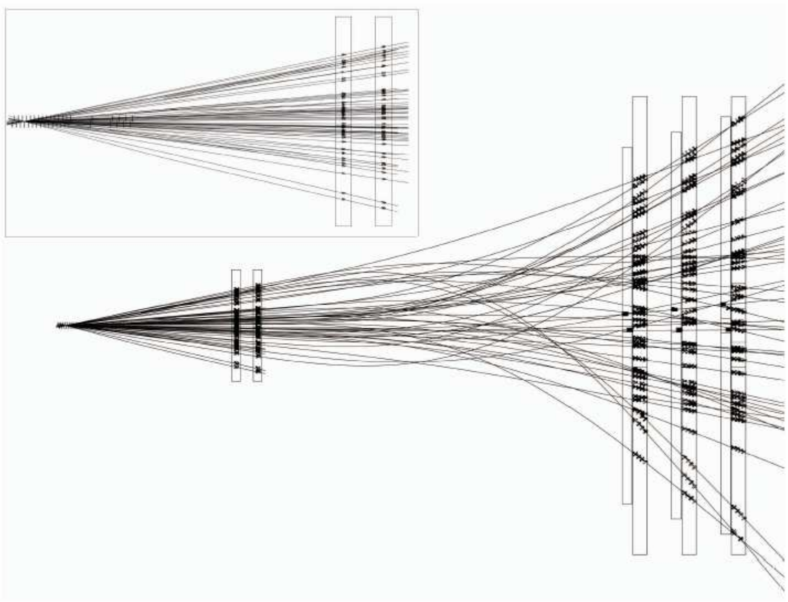


Figure 2.16: A typical LHCb event. The reconstructed tracks are also shown.

vector specified for different z-positions (x, y, $\frac{dx}{dz}$, $\frac{dy}{dz}$, $\frac{Q}{p}$). The track reconstruction algorithm starts with the search for track “seeds” in the VELO or in the T-stations. As the magnetic field in the VELO is negligible (Figure 2.5), the VELO seed search algorithm looks for straight lines. A T-seed is instead parametrized as a parabola, due to the presence of a fringe magnetic field in the T-stations.

T-seeds provide also an initial momentum estimation either from the curvature in the magnetic field or assuming that the particle comes from the interaction region, in the latter case a precision of up to 1% can be achieved.

Starting from these seeds, four kinds of tracks are classified as shown in Figure 2.17.

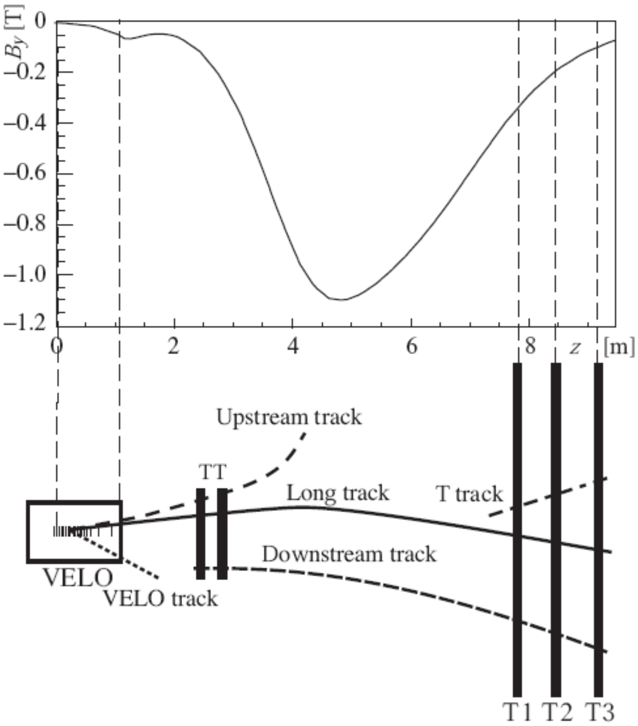


Figure 2.17: Schematic illustration of the different types of tracks in the LHCb detector. The illustration is in the bending plane and is not to scale. The magnitude of the magnetic field along the z-axis is also shown.

Long Tracks are tracks that traverse all the tracking detectors from the

VELO to the T-stations. These kinds of tracks are the most important concerning B-physics studies. The long track algorithm tries to combine VELO seeds with hits in the T-stations. When at least one hit in the T-stations is combined with the VELO hits, a momentum measurement is provided and hence a trajectory can be extrapolated. Hits in the other stations and in the TT are looked for, around this trajectory. The efficiency of this procedure is about 90%, an additional 5% can be obtain by reconstructing the tracks starting from T-seeds and looking for hits in the VELO which match the trajectory extrapolation.

Upstream Tracks traverse the VELO and the TT. Since the magnetic field between the VELO and the TT is low, they have a poor momentum resolution. However these tracks traverse the RICH1 and therefore have identification information. These tracks are usually particles of low momentum and are used both for understanding the background and for B-decay reconstruction or tagging.

The reconstruction algorithm looks for hits in TTa which match with the VELO seed. This trajectory defines the momentum of the particle. Moreover hits in TTb are required for confirmation, to minimize the ghost rate.

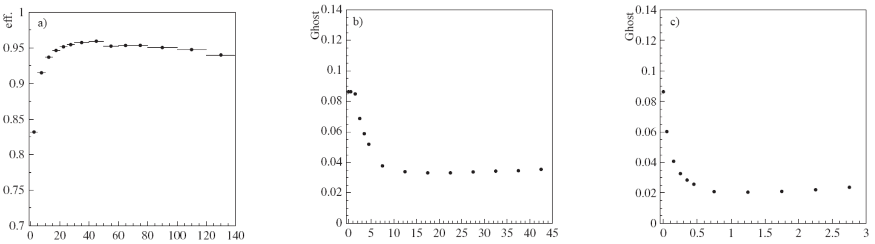
Downstream Tracks come mainly from particles with long life time that decay outside the VELO, such as K^0_S and Λ . They traverse only the TT and the T-stations. The reconstruction algorithm looks for hits in the TTa and TTb stations which match with the trajectory extrapolated with the T-station momentum measurement.

What remains are **T-tracks** and **VELO-tracks**, which have not been associated to other kind of tracks. VELO tracks are mainly backward tracks and typically at large angle: they are used for Primary Vertex reconstruction.

After reconstruction all tracks are refitted with a Kalman filter algorithm and the vector states are updated.

A global χ^2 of the track is stored as well as the “pull”-distributions of the track parameters. A typical LHCb event is shown in Figure 2.16.

In Figure 2.18(a) the reconstruction efficiency as a function of the particle momentum is shown. For long tracks of $P > 10GeV/c$, the reconstruction efficiency is about 94%. The ghost rate for long tracks is shown in Figures



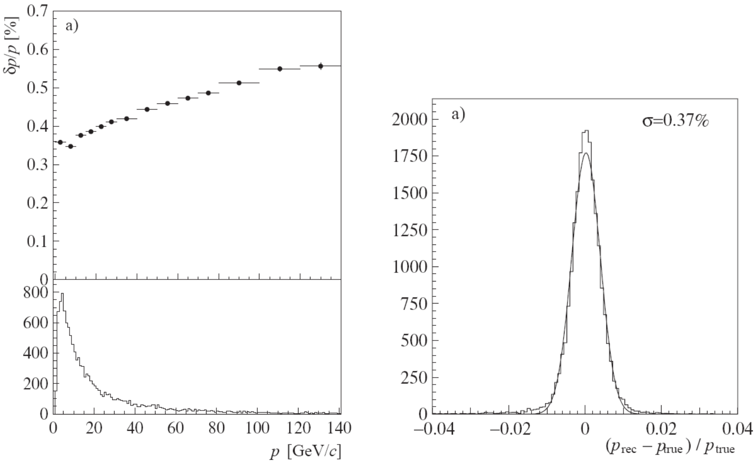
(a) Reconstruction efficiency as a function of momentum $P(GeV/c)$.

(b) Ghost rate as a function of the momentum $P(GeV/c)$.

(c) Ghost rate as a function of the transverse momentum $P_t(GeV/c)$.

Figure 2.18: Performance of the long track finding.

2.18(b) and 2.18(c). The two plots show that the ghost rate is higher at low P and P_t. For particle with $p_T > 0.5GeV$ the ghost rate is about 3%. The momentum resolution for long tracks is about 0.37% and is gaussian distributed (Figure 2.19(a) , 2.19(b)). For downstream tracks, as these particles traverse most of the magnetic field, the resolution is still good, about 0.47%.

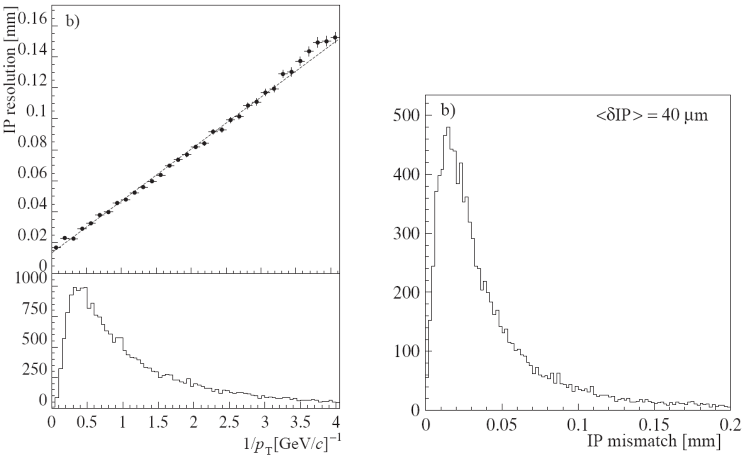


(a) Momentum resolution as a function of the particle momentum. The average particle momentum distribution is also shown on the bottom.

(b) Momentum resolution with a single Gaussian fit.

Figure 2.19: Momentum resolution.

The resolution is worse ($\frac{\delta p}{p} \sim 15\%$) for upstream tracks which see a very small magnetic field.
The impact parameter resolution for long tracks can be parametrized as : $\sigma_{IP} = 14\mu m + 35\mu m/p_T$ (Figure 2.20(a)), with an average resolution of $< \delta IP >= 40\mu m$.



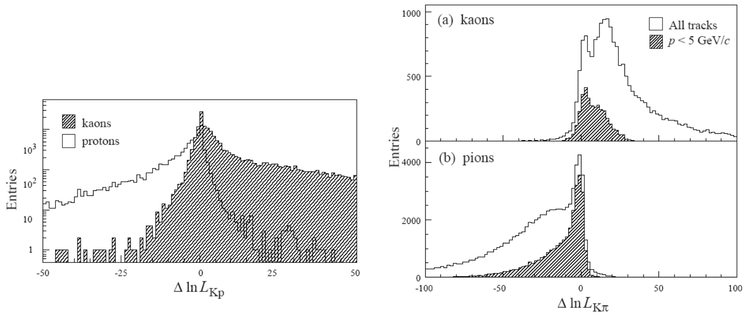
(a) Impact Parameter resolution as a function of $\frac{1}{P_t(GeV/c)}$. The average $\frac{1}{P_t(GeV/c)}$.

(b) Impact parameter precision.

Figure 2.20: Impact parameter resolution.

2.9 Particle Identification

Particle identification is a very important task for the LHCb experiment. $e^{\pm}/h^{\pm}$ identification is mainly done with the electromagnetic calorimeter, $\mu^{\pm}/h^{\pm}$ identification with the muon detector and $p^{\pm}/\pi^{\pm}/K^{\pm}$ identification with the RICH system. The neutral particles as π^0 and γ are mainly identified with the calorimeter system.
In this Section the different particle identification criteria will be discussed. The purpose of the RICH system is to distinguish $\pi^{\pm}$ from $K^{\pm}$ and $p^{\pm}$.
A likelihood is built in as follows: the pattern of hits observed in the RICH is compared with the pattern we would expect under a given set of hypotheses, the likelihood is determined from this comparison. The difference between



(a) DLL between the proton and kaon hypotheses.

(b) DLL between the pion and kaon hypotheses.

Figure 2.21:

the logarithm of the likelihood (DLL) for the different hypotheses is used to distinguish the particles (2.2-2.4).

$$\Delta lnL_{\pi K} = ln(\frac{L(\pi)}{L(K)}) \tag{2.2}$$

$$\Delta lnL_{\pi p} = ln(\frac{L(\pi)}{L(p)}) \tag{2.3}$$

$$\Delta lnL_{Kp} = ln(\frac{L(K)}{L(p)}). \tag{2.4}$$

In Figure 2.21, the ΔlnL for various particle hypotheses, using only RICH information is shown, while in Figure 2.22 the significance $N_\sigma = \sqrt{2\times \mid \Delta lnL \mid}$ for $\pi - K$ separation is shown.

Particle identification performance depends on the likelihood cut chosen in the analysis. With default cuts, it is possible to achieve a kaon efficiency of 88% with a pion misidentification of 3% in the range $2GeV - 100GeV$. Figure 2.23 shows the efficiency and the misidentification performance for standard cuts in the DLL.

Muon identification is provided by the muon detector, by extrapolating tracks for which $P > 3GeV/c$ in the muon stations and looking for muon detector hits within the **Field of Interest** (FOI), around the trajectory. A muon candidate is a track with a minimum number of muon hits within the corresponding FOI. The FOIs are parametrized looking at the plots of the variable $\frac{(x(y)_{MC} - x(y)_{hit})}{d}$, being d the pad size. Since the penetrating power

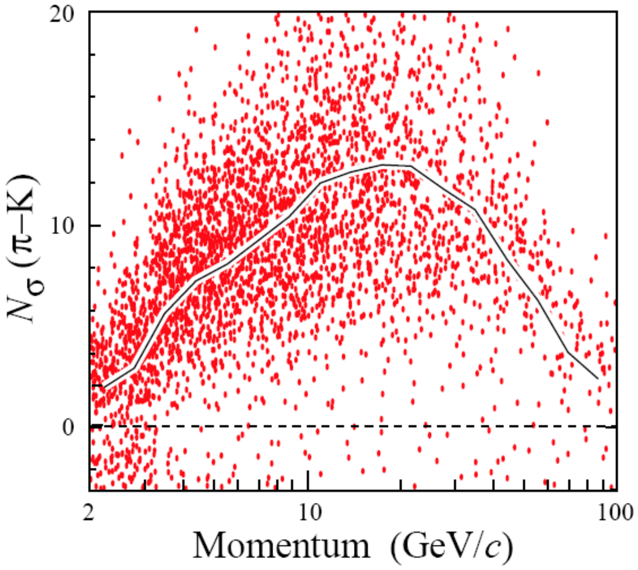


Figure 2.22: $\pi - K$ separation as a function of the momentum for true pions. The black line represents the average $\pi - K$ separation.

Table 2.3: Stations that must have hits within the relative FOI for the selection of a muon candidate.

Momentum (GeV)	Muon Stations
3<p<6	M2+M3
6<p<10	M2+M3+(M4 or M5)
p>10	M2+M3+M4+M5

of the muons increases with the energy, the minimum number of stations required to have hits in their correspondent FOIs depends on the track momentum. This is summarized in Table 2.3 for different muon momentum ranges.

The muon identification likelihood is based on the comparison of slopes in the muon system and in the tracker (Figure 2.24(a)) and the average track-hits distance of the hits within the FOIs (2.24(b)). The main background for muon identification consists of high energy pions. The *DLL* for muon and pion is shown in Figure 2.25. For standard loose cuts an average efficiency of about 95% and a misidentification of pions of 3% is achieved within a

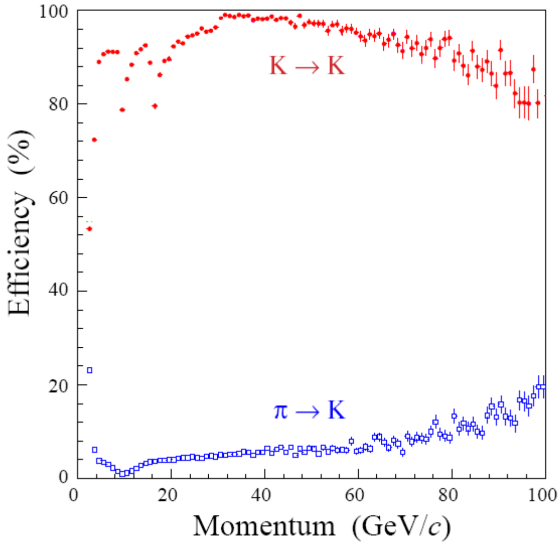


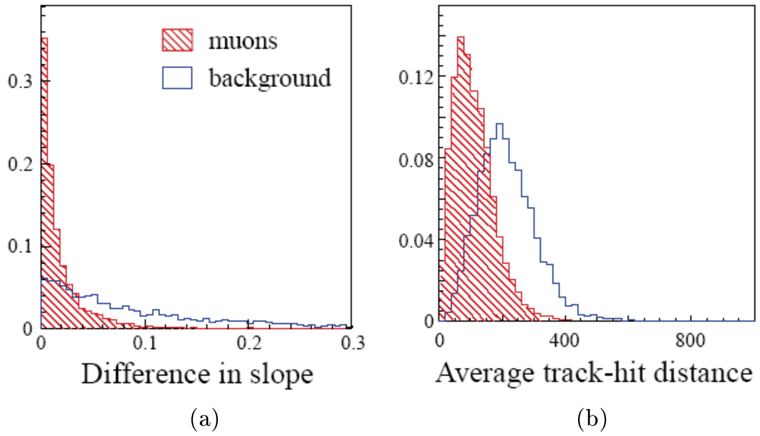
Figure 2.23: Kaon identification efficiency (red points) and pion misidentification rate (blue points).

momentum range $3GeV - 100GeV$ (Figure 2.26). However, as can be seen in the following, using a tighter cut in the $DLL(\mu - \pi)$ it is possible to have a misidentification of about 1% with an average efficiency higher than 90% for muons of $P > 3GeV/c$.

The electron identification is primarily done with the electromagnetic calorimeter. The main background for electron identification consists of charged hadrons.

A global χ^2 (χ^2_e) is constructed including the balance of track momentum with the energy of the track cluster in ECAL and the matching of the shower barycenter position with the extrapolated track trajectory (Figure 2.28(a)). A second estimator consists of searching neutral clusters expected in well known position due to the electron bremsstrahlung before the magnet. The estimator χ^2_{brem} is shown in Figure 2.28(b) for electrons and hadrons.

This estimator is also used by the bremsstrahlung recovering algorithm. Photons radiated by electrons are searched according to the χ^2_{brem}. The electron momentum before the bremsstrahlung is then recovered taking into account the photon radiation (Figure 2.27(b)).



(a) (b)

Figure 2.24: Variables used to build the muon identification likelihood.

Other estimators are: the track energy deposition in the Preshower and the deposition along the extrapolated trajectory in HCAL (Figure 2.28(c) and 2.28(d)). A likelihood distribution is constructed combining these estimators. The DLL between the electron and pion hypotheses is shown in Figure 2.29. With default cuts an average efficiency of about 95% and a pion misidentification of about 0.7% is reached (Figure 2.30).

A global likelihood for charged particles is obtained combining the various detector information as follows:

$$L(e) = L^{RICH}(e)L^{CALO}(e)L^{MUON}(non\ \mu) \tag{2.5}$$

$$L(\mu) = L^{RICH}(\mu)L^{CALO}(non\ e)L^{MUON}(\mu) \tag{2.6}$$

$$L(h) = L^{RICH}(h)L^{CALO}(non\ e)L^{MUON}(non\ \mu). \tag{2.7}$$

The photons are identified as neutral clusters in the electromagnetic calorimeter. ECAL clusters are reconstructed and matched with all the reconstructed tracks. A χ^2_γ is built extrapolating the trajectory to ECAL and matching the track impact point with the cluster barycenter. The minimal value of the χ^2_γ is used as an estimator, as shown in Figure 2.31.

Photons converted after the magnet can be identified requiring that there are no hits in the SPD. Using these identification criteria, in the $B^0 \to K^*\gamma$ analysis, a purity of 90% for non converted photons and of 79% for converted

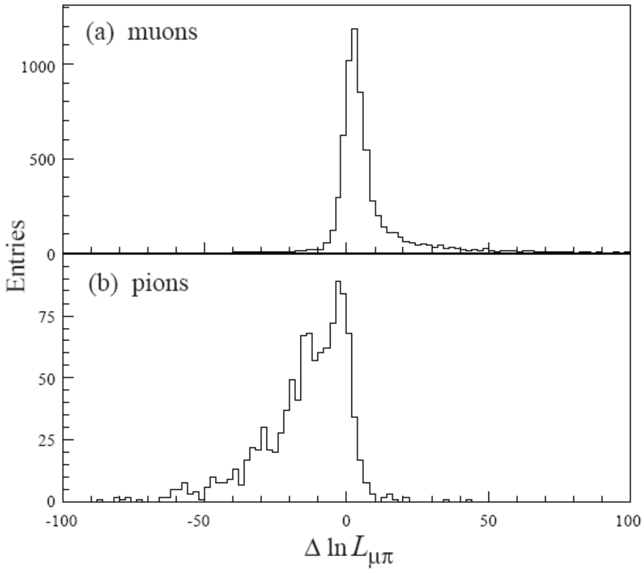


Figure 2.25: $DLL(\mu - \pi)$ for true muons a) and true pions b).

photons is achieved. The sample purity can be increased by putting a cut in the energy deposition in the PS.

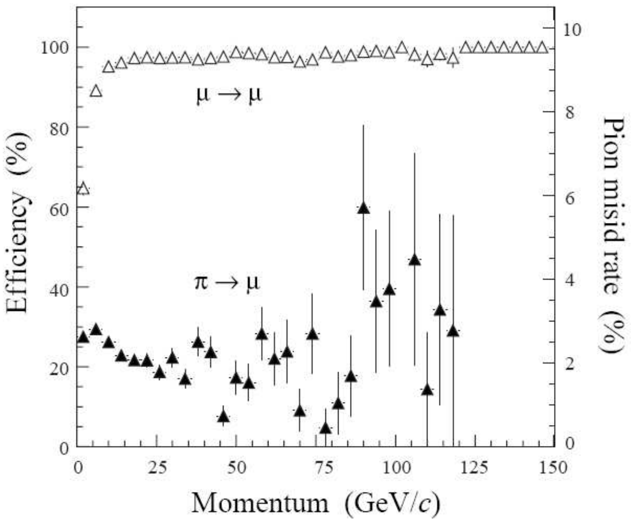
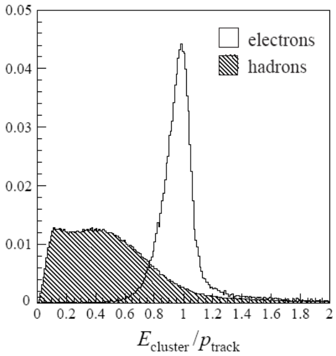
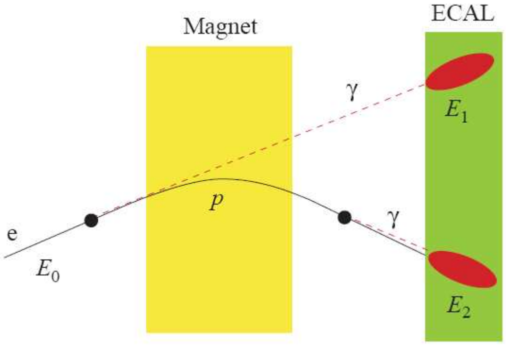


Figure 2.26: Muon identification efficiency (open points) and pion misidentification rate (black points) as a function of momentum.

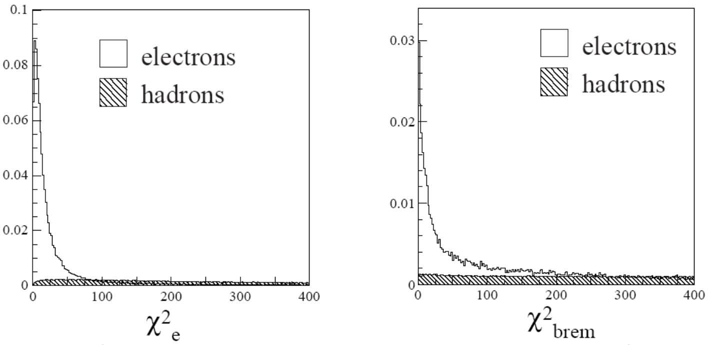


(a) Ratio of the uncorrected energy of the charged clusters in ECAL to the momentum reconstructed track for hadrons (shaded histogram) and for electrons (open histogram).



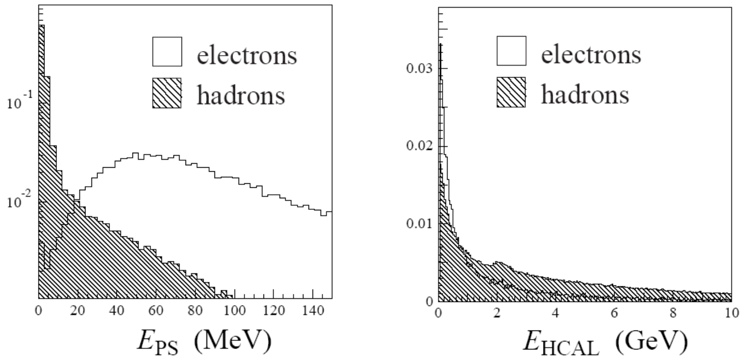
(b) Schematic illustration of the Bremsstrahlung correction. The electrons passing through the material of the detector can radiate photons. If they radiates before the magnet a defined amount of energy E_1 such that $E_0 = E_1 + E_2$, is seen in ECAL. For the electron identification we have $E_2 = P$.

Figure 2.27:



(a) The χ^2_e estimator of the matching of reconstructed tracks and charged cluster in ECAL.

(b) The value of the χ^2_{brem} estimator.



(c) The energy deposition in the preshower.

(d) The energy deposition along the extrapolated particle trajectory in HCAL.

Figure 2.28: Electron identification estimators.

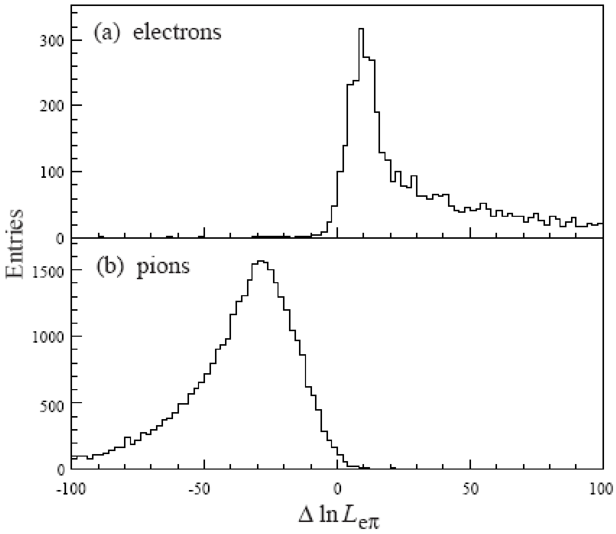


Figure 2.29: *DLL* between the electron and pion hypotheses for true electrons a) and true pions b).

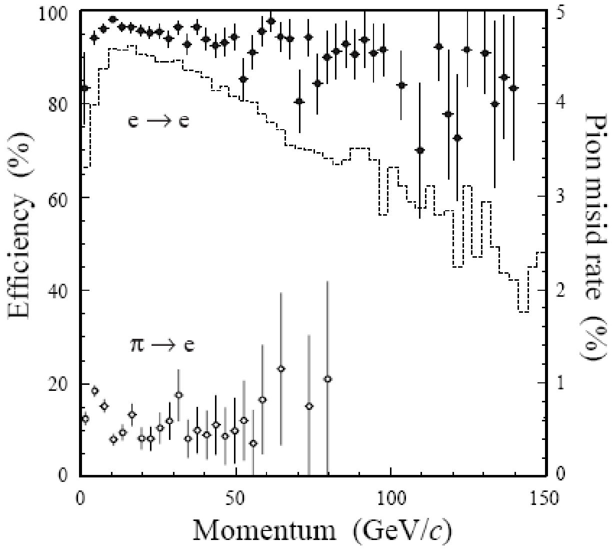


Figure 2.30: Electron identification efficiency and pion misidentification rate as a function of momentum. The plots are given for standard loose cuts.

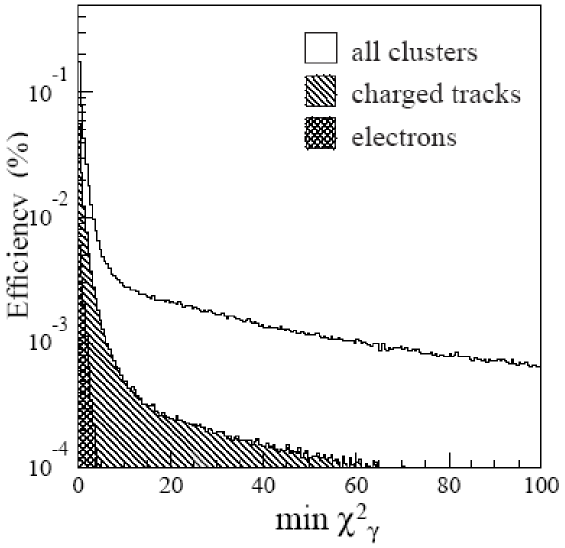


Figure 2.31: Value of the minimum of the χ^2_γ estimator.

2.10 Trigger

The task of the trigger system is to select interesting events in the pp interactions reducing the event rate to a sensible acquisition frequency. It is a very important component of the LHCb experiment. The LHC bunch crossing frequency is 40MHz, however due to the LHC bunch structure, at the LHCb interaction point, the visible crossing frequency is reduced to 10MHz. This frequency is still to high to allow us to record such an amount of data on tape. Moreover most of the events are not interesting for B-physics studies. At the luminosity of $L = 2 \times 10^{32} cm^{-2} s^{-1}$ the rate of events with $b\bar{b}$ pairs is about 100kHz and only a small fraction of these events have the decay products in the LHCb acceptance, reducing the visible rate to a few kHz. The trigger must reject the background and enhance the signal with very fast algorithms. The LHCb trigger system is divided into two levels: L0 with an output rate of 1MHz and the HLT with an output rate of 2kHz (Figure 2.32)

2.10.1 L0 Trigger

The L0 trigger is a “hardware” trigger, implemented with custom electronics with 1MHz output rate. It uses the fact that B-mesons decay into particles with high transverse momentum. The latency of the L0 trigger is $4.2\mu s$. It involves four subsystems: the pile-up, the calorimeter, the muon system and the global decision unit (L0DU) which provides the final decision.

The purpose of the pile-up system is to veto events with multiple interactions. However, the dimuon trigger overrules the pile-up veto.

The purpose of the calorimeter trigger is to identify events with electrons, photons or hadrons of high transverse energy. There is a veto in the event, if the total transverse energy is less than 5.0GeV. The event is also vetoed if the SPD multiplicity is higher than 280. The E_T threshold is 2.6GeV for $e^{\pm}$, 2.3GeV for γs, 3.5GeV for charged hadrons and 4.0GeV for π^0s.

The muon trigger accepts the event if there is a single muon of $P_t > 1.3 GeV$ or if the sum of the two largest P_t muons is higher than 1.5GeV. All the L0 trigger information is combined by the L0DU and the event is accepted or

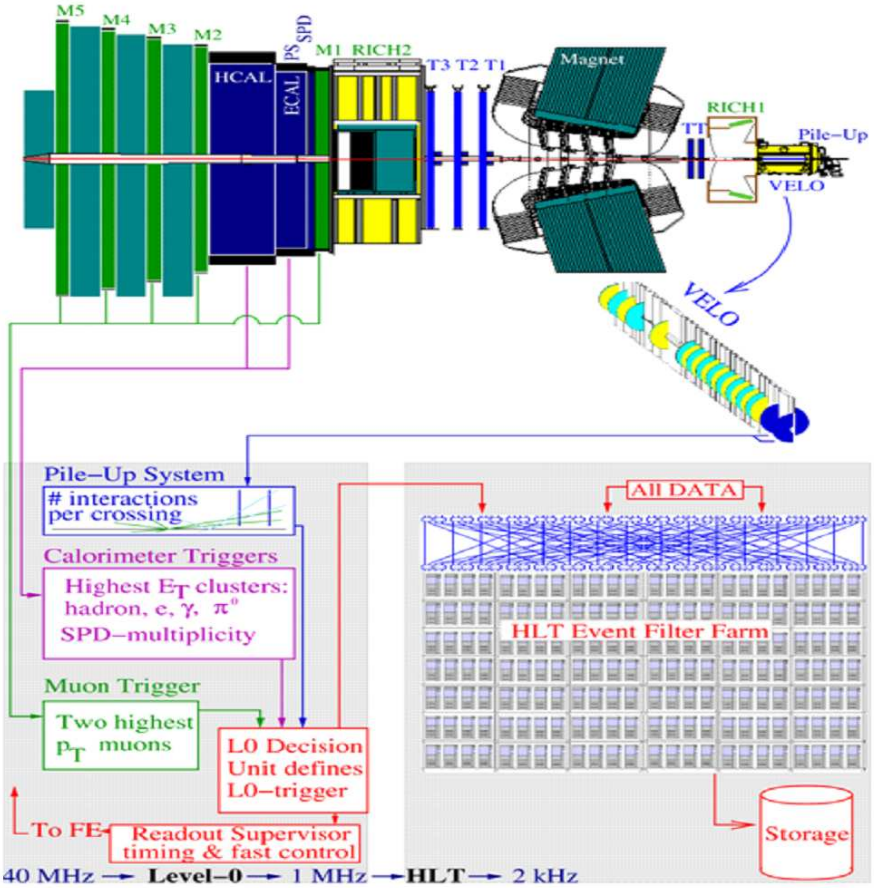


Figure 2.32: Schematic illustration of the LHCb trigger system.

rejected according to the above criteria.

2.10.2 High Level Trigger (HLT)

The HLT consists of a C++ application that will run over a CPU farm with about 2000 nodes. Even though the HLT application has access to all the data, because of the limited CPU power, it must reject the bulk of the events with only part of the available information. For this purpose a careful design of the so called “alleys” has been studied. The alleys’ task is to refine and confirm L0 decision. Most of the events are selected by only one L0 trigger channel and consequently they pass through only one alley.
However there are events (about 15%) which are selected by multiple triggers

and which pass through more than one alley. For these events care was taken to avoid the reconstruction of the tracks or of the primary vertex twice. The alleys operate independently and when an event is selected at least by one alley it is processed by the inclusive and exclusive trigger.
In Figure 2.33 the flow diagrams of the different trigger sequences is shown. The flow diagram of a single HLT alley is shown in Figure 2.34. Each alley starts with the L0 confirmation with improved resolution, using at least one tracking sub-detector. If the L0 confirmation is passed, additional B-decay track candidates are reconstructed.

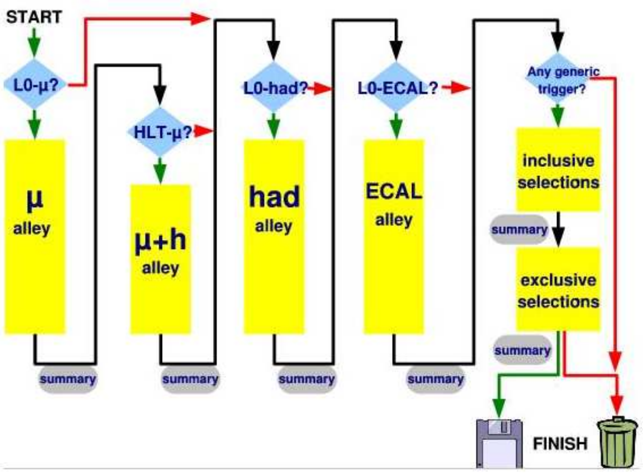


Figure 2.33: Schematic illustration of the flow diagram of the Alley structure in the HLT.

For high P_t muons, selected using M2-M5 stations, a fast 3D reconstruction is performed using also T1-T3 stations and a resolution of $\frac{\delta p}{p} \sim 6\%$ is achieved.
In the other alleys, a search in the T1-T3 tracking system cannot be done because it is too time consuming and the track reconstruction precision is $\frac{\delta p}{p} \sim 20\% - 40\%$. At pre-trigger level[1] the rate is sufficiently reduced to use T1-T3 information, the tracking reconstruction precision at this level is about $\frac{\delta p}{p} \sim 1\%$. In the alley trigger long tracks are used and impact parameter information added. Each alley ends with a summary in which all

---

[1] The pre-trigger is a particular stage within each alley, see Figure 2.34.

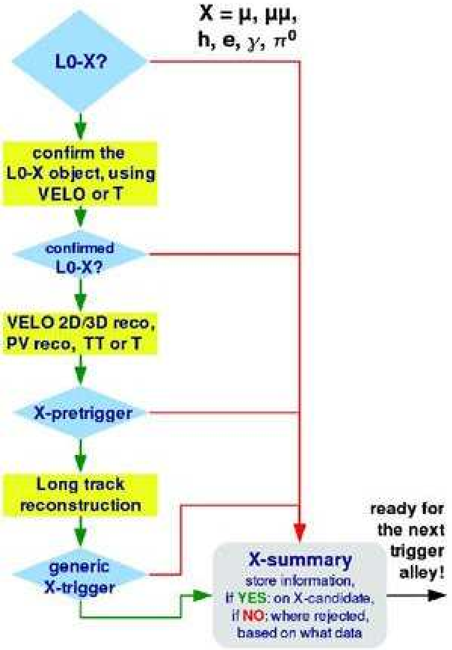


Figure 2.34: Schematic illustration of the flow diagram of a single HLT alley.

vertexes and track information, which could have triggered the event, are stored.

After the alleys' selection the rate is about $10kHz$. At this stage the all remaining 2D long tracks [2] are reconstructed and invariant mass cuts and precise pointing cuts are applied. The tracks are combined to reconstruct composite particles like K^*, Φ, D^0, D_s and J/ψ. The exclusive trigger has an output rate of about 200Hz. In the exclusive trigger different pre-selections are used to store various signals, side-bands and control channels (Figure 2.35).

The remaining 1.8kHz of data are divided into three streams: dimuon stream

---

[2] At this stage the tracks are reconstructed using only the VELO R-Sensor. This is due to the fact that the b-hadrons move predominantly along the beam axis. Therefore the impact parameter projection is large in the plane defined by the beam axis and the track, while in the perpendicular plane the secondary tracks are almost indistinguishable by tracks directly coming from the PV.

(600Hz), D^* candidates (300Hz) and b-inclusive (900Hz). The dimuon stream selects muon pairs with an invariant mass above 2.5GeV. This will be useful for studies on the systematic uncertainty in lifetime measurements. The D^* stream is useful for particle identification efficiency and misidentification studies, but can also be useful for CP violation measurements in charm physics.

The b-inclusive stream selects muons with high P_t and impact parameter, this stream can be used to study the trigger efficiency and data mining [3].

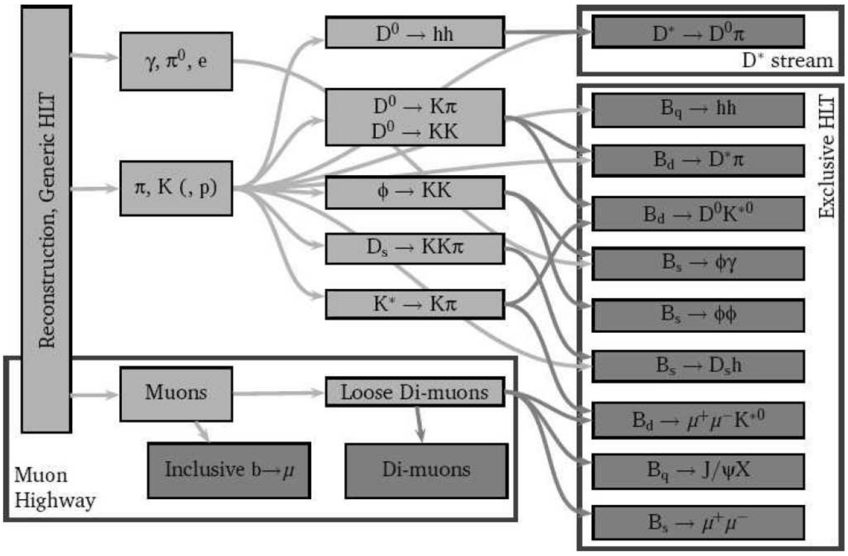


Figure 2.35: Simplified data flow in the HLT. Each Grey box is a set of algorithms.

2.11 The LHCb software

With the increasing complexity of the high energy physics experiments, the construction of the detector needs many years of research and development. Moreover the software systems have to process a very large amount of data. The study of the MonteCarlo (MC) simulation of physics and detector response is very important to understand the real data analysis, in particular

---

[3] At time of writing the HLT design was still not complete, this description must be thought as a general description, some details are still under discussion.

to test performance and robustness of hardware and software.

The LHCb software is written in Object Oriented (OO) language. This choice is motivated by the fact that OO languages are suitable when the software has a high level of complexity and when it must be developed by many physicists and must be high scalable. The LHCb software must be flexible and scalable to fit the experiment characteristics for a running period of 10-20 years. The main application is the framework GAUDI which is written in C++ language. Its purpose is to provide specific components for the interaction of the various tools and algorithms.

All the Gaudi classes are derived either from the Algorithm class or from the DataObject class[4]. The purpose of the Algorithm class is to manipulate data which interact with other specific Algorithms.

In the following, the main parts of the LHCb simulation software are discussed. In addition to the LHCb software packages the software ROOT and in particular the application TMVA (Toolkit for the Multivariate Analysis [62]) have been used in this thesis work.

2.11.1 Gauss

The Gauss package is a software devoted to physics simulation of proton-proton interactions, particle decays and interactions of stable particles with the different parts of the detector. This software is divided into two main phases: the generation and the simulation.

In the generation phase the proton-proton interaction and the unstable particle decays are simulated. Two main programs are used for this purpose: Pythia and EvtGen (originally developed by the *BABAR* collaboration). Pythia provides the description of the $p-p$ interaction and of the hadronization process. The heavy particle decays (including b-hadrons) are simulated by the EvtGen package. Within the EvtGen package there are many decay models written in the helicity amplitude formalism. Among these models there is the *phase space model* (*psp*) which takes into account just the phase space of the products for the computation of the kinematical variable distri-

---

[4] Among the base classes there is also the Converter class which is mainly inserted for technical reasons.

butions of the out-coming particles in the b-hadron rest frame. The daughter particles are then boosted in the lab frame. In this method no resonant structure in the decay is considered. In particular this is used for the $B_{s,d} \to l^+l^-$ decays.

The interaction of stable particles with the detector is simulated with the GEANT 4 package. The detector description is very accurate and was updated between the 2004 and 2006. For this reasons, in this study, two different data challenges known as DC04 and DC06, were used.

2.11.2 Boole

The output of the Gauss software consists of GEANT hits, which are the result of the interaction of the particles with the detector. This information is then processed by the Boole package. The purpose of this package is the simulation of the digitization procedure. This is the final step of the LHCb detector simulation.

The Boole software describes the readout electronics response to the particle interactions. Several detector effects like the spill-over effect, electronic noise and the LHC background are included in the simulation. In addition, the Boole software simulates the response of the L0 Trigger.

2.11.3 Brunel

The Brunel package is devoted to the reconstruction of LHCb events. It takes the Boole output as an input. Brunel provides track finding and fitting as described above. All the information of the tracking and identification detectors are combined to create proto-particles. A proto-particle is an object which contains all particle information: the particle trajectory, the momentum and the value of the identification likelihood for different particle hypotheses.

2.11.4 DaVinci

The main analysis software package is called DaVinci. Within DaVinci, particle objects are built from proto-particles, applying identification cuts.

The DaVinci package provides several methods for event selection. Within the DaVinci package there are two main objects: the MC truth and the reconstructed data. The reconstructed data are similar to what we would expect to have in the experiment. However, at this stage of the analysis, it is very interesting to verify the matching of the reconstructed particles with the MC truth, which can be done with various methods, as shown in Figure 2.36.

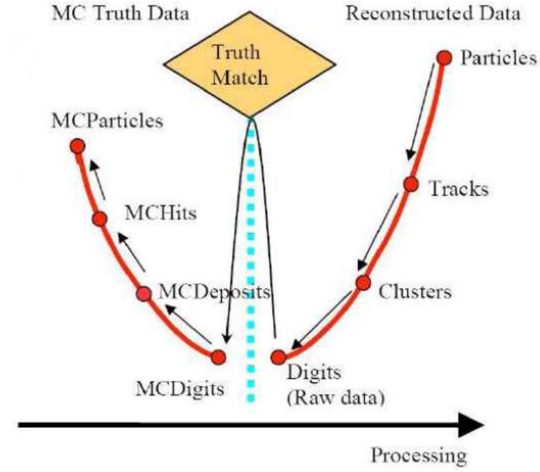


Figure 2.36: MC truth relation with the reconstructed data objects.

2.12 Other software

In addition to the LHCb software, other packages were used in this work. The ROOT software is a package developed at CERN for physics data analysis [66]. It is an object-oriented framework which provides various methods and tools for physics analysis.

The package TMVA is a toolkit for multivariate analysis. All the methods described in chapter 3, except for the PCA method, are present in this package.

Roofit is a software that was originally developed by the BABAR collaboration and it provides a toolkit for modeling expected distributions, for doing likelihood fits and for generating "Toy Monte Carlo" experiments when the variable distributions are given. This software was used in [69] to study the asymmetries in the $B_d^0 \to K^{*0}\mu^+\mu^-$ channel (see section 6.1).

2.13 Main physics goals of the LHCb experiment

The main goals of the LHCb experiment will be measurements of CP violation and rare decays.

Rare decays physics is extensively discussed in the context of this thesis.

Concerning CP violation measurements, the main targets of the LHCb experiment are the measurement of the angles of the unitarity triangles γ and ϕ_s.

The angle γ, which is involved in $b \to u$ transitions, can be measured from the time dependent CP asymmetry of the channel $B_s^0 \to D_s K$ and the $B_d^0 \to D^0 K$ decays. These channels can receive different contributions from NP. The LHCb experiment can do precise measurements of γ with a statistical uncertainty of about 5 degrees in $2fb^{-1}$ of integrated luminosity.

One of the most interesting CP-violating observable concerning the search of NP is the weak phase ϕ_s. The SM prediction of this quantity is $\phi_s \simeq -0.04$, however this value can be largely enhanced by NP contributions. The most sensitive channel to this measurement is the $B_s^0 \to J/\psi\phi$ decay. With a time dependent CP analysis of this decay a sensitivity of about 0.02 to ϕ_s can be achieved.

Chapter 3

Statistical tools

3.1 Multidimensional methods

Usually event selection is done using rectangular cuts with respect to geometrical and kinematical variables. However, even if this method it is the most intuitive and simple, it doesn't use all the information embedded in the data.

Moreover, it doesn't consider the correlation between the variables. In particular, in very rare decays we are interested in the branching ratio measurement. For such measurements it is very important to have high sensitivity but in general we don't care about possible biases in the variable distributions. In the following the multidimensional methods used in this thesis are briefly described. All these method, except the PCA method (Section 3.4) were present in the package TMVA.

3.2 Likelihood

Assuming that the variables x_i, used for signal selection, are independent from each others, we can define the likelihood function as follows:

$$L_k = \prod_{i}^{n} p_i(x_{ki}), \tag{3.1}$$

where $p(x_{ki})$ is the probability for the variable x_i to have the value x_{ki}.

The likelihood function of Eq.3.1 has a simple interpretation: it is the probability for a given set of data x_{ik} to be measured in the experiment, when the

$p_i(x)$ probability function is assumed. Since the maximum of a function corresponds to the maximum of the logarithm of that function, it is much more convenient for computational reasons to use the logarithm of the likelihood function (LL).

$$\ln L_k = \sum_{i}^{n} \ln p_i(x_{ki}). \tag{3.2}$$

The likelihood method can be used for parameter estimation. When a known theory with some free parameters is given, the maximum of the LL corresponds to the best estimation of the set of parameters.

When doing event selection we have to distinguish between two hypotheses: the signal hypothesis and the background hypothesis. In this case we have to compare the two functions L_{signal} and $L_{background}$ computed using the variable probability distribution for signal and background.

For this purpose the *difference in the logarithm of the likelihood* (DLL) is often used. However, in the signal selection, it is more convenient to use the likelihood ratio which is a function defined in the range $0 \leq R \leq 1$ as follows:

$$R = \frac{L_{signal}}{L_{signal} + L_{background}}. \tag{3.3}$$

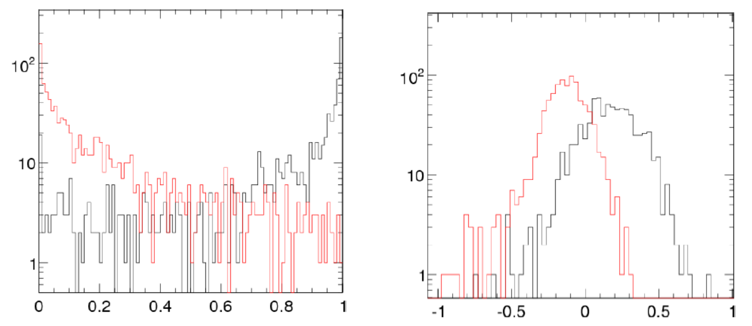
The background is therefore peaked at $R = 0$ and the signal at $R = 1$. Moreover, in very rare decay selections, the maximum sensitivity is usually achieved for low signal efficiency and high purity (i.e. high background rejection). For this reason the inverse sigmoid transformation of the variable R, which is a non linear transformation, is here used. This is defined as follows:

$$y \rightarrow -\frac{1}{\tau} ln \left(\frac{1}{R} - 1 \right), \tag{3.4}$$

where τ is a scale factor and R is the ratio computed as in Eq. 3.3. The transformation 3.4 has the advantage that the high and low efficiency regions are magnified. The result of the transformation is shown in Figures 3.1(a)-3.1(b).

3.3 Decorrelated Likelihood

In 3.1 it is implicitly assumed that the variables are independent and therefore the total probability is given by the product of the single variable distri-



(a) Typical likelihood ratio distribution.

(b) Likelihood ratio distribution after an inverse sigmoid transformation.

Figure 3.1: An example of likelihood ratio distribution before (a) and after (b) an inverse sigmoid transformation is shown. The background (black histogram) are $b\bar{b} \to e^{\pm}\mu^{\mp}$ events, the signal (red histogram) are $B_s^0 \to e^+\mu^-$ events.

butions, for all the variables. However, the correlation between the variables can be seldom neglected and so we pay a price in performance of the likelihood method.

In the hypothesis that the variables are linearly correlated it is possible to decorrelate the variables using the covariance matrix.

Σ being the correlation matrix constructed using the data:

$$\Sigma_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}(x_k^i - \overline{x}^i)(x_k^j - \overline{x}^j), \tag{3.5}$$

where $\overline{x}^i$ is the mean of the $i-th$ variable, y_i being a new set of variables such that:

$$y = C^{-1}x, \tag{3.6}$$

where C a symmetric real matrix such that

$$CC^{\dagger} = CC = \Sigma, \tag{3.7}$$

ee have:

$$x^{\dagger}\Sigma^{-1}x = y^{\dagger}C^{\dagger}\Sigma^{-1}Cy = y^{\dagger}y. \tag{3.8}$$

The set of variables y is therefore a set of decorrelated variables.

However this linear decorrelation process is valid only for linearly correlated and Gaussian distributed variables.

3.4 Projection and Correlation approximation (PCA)

The Projection and correlation approximation is described in [58]. Denoting with x_i the n-dimensional vector of the variables, in general these variables are not Gaussian distributed.
y_i being a set of variables obtained as follows:

$$y_i = \sqrt{2} erf^{-1}(2F(x_i) - 1), \tag{3.9}$$

where erf^{-1} is the inverse error function and the F is the cumulative distribution. The variables y_i are Gaussian distributed with respect to the projection planes. Assuming that the n-dimensional distribution of y_i is an n-dimensional Gaussian distribution we have:

$$P(x) = \frac{1}{(2\pi)^{n/2} \mid \Sigma \mid^{\frac{1}{2}}} e^{-\frac{1}{2} y^{\dagger}(\Sigma^{-1} - I) y} \prod_{i=1}^{n} e^{\frac{1}{2} y_i^2}, \tag{3.10}$$

with

$$\begin{aligned} P(x) &= \frac{1}{(2\pi)^{n/2} \mid \Sigma \mid} e^{-\frac{1}{2} y^{\dagger}(\Sigma^{-2}) y} = \\ &= \frac{1}{(2\pi)^{n/2} \mid \Sigma \mid e^{\frac{1}{2} y^{\dagger}(\Sigma^{-1} - I) y} \prod_{i=1}^{n} e^{-\frac{1}{2} y_i^2}}. \end{aligned} \tag{3.11}$$

By substituting the Equation 3.9 in the previous formula we have:

$$\begin{aligned} e^{-\frac{1}{2} y^2} &= exp\left\{ \frac{1}{2} (\sqrt{2} erf^{-1}(2F(x) - 1))^2) \right\} = \\ &= \frac{\sqrt{\pi}}{2} \frac{d}{dx} erf(erf^{-1}(2F(x) - 1)) = \\ &= \frac{\sqrt{\pi}}{2} \frac{d}{dx} (2 \frac{\int_{x_{min}}^{x} p(x')dx'}{\int_{x_{min}}^{x_{max}} p(x')dx'} - 1) = \\ &= \frac{\sqrt{\pi}}{\left(\int_{x_{min}}^{x_{max}} p(x')dx' \right)} p(x). \end{aligned} \tag{3.12}$$

Inserting 3.12 in 3.11 we obtain 3.10. Under this approximation we can use the probability ratio defined in the following Eq. 3.13, to separate signal and background.

$$R = \frac{P_{signal}(x)}{P_{signal}(x) + P_{background}(x)}. \tag{3.13}$$

The assumption that the $P(x)$ distribution of Equation 3.10 is an n-dimensional Gaussian distribution is not quite true. In fact, with this method, it is only guaranteed that the y_i projections are Gaussian distributions. However in many cases this approximation is good enough, but its validity should be controlled with a χ^2 test.

3.5 Artificial neural network (ANN)

Neural Networks are used in several branches of science. In physics the ANN are mainly used for classification tasks. The ANN purpose is to simulate some characteristics of the human brain:

-High parallelism;

-Non linearity;

-Adaptability.

In practice, the ANN is a set of interconnected neurons. Each neuron gives a certain response to a given signal in input. The neurons can be arranged into layers as in Figure 3.2, where an input layer which corresponds to the input variables, a hidden layer and an output layer can be recognized. The neurons are connected by links called synapses (in analogy with the human brain). Each neuron receives some input from all the other neurons of the previous layer and the synapses connect this neuron with all the neurons of the next layer. The response function consists of a synapse $k : \Re^n \rightarrow \Re$ function and a $\alpha : \Re \rightarrow \Re$ activation function.

The k function and the α function can have the following forms:

$$k : y^i \rightarrow \begin{cases} b_j + \sum_{j=1}^{n} y_i \omega_{ij} \\ b_j + \sum_{j=1}^{n} (y_i \omega_{ij})^2 \\ b_j + \sum_{j=1}^{n} \mid y_i \omega_{ij} \mid \end{cases} \tag{3.14}$$

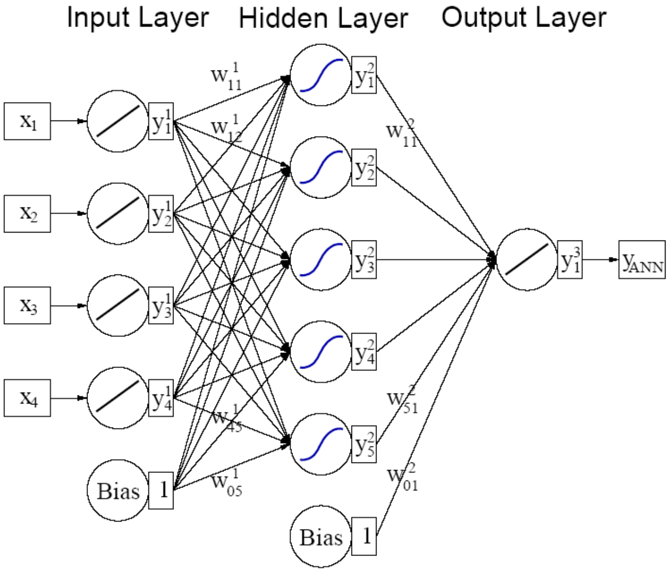


Figure 3.2: Schematic illustration of an artificial neural network with a single hidden layer.

$$\alpha : x \rightarrow \begin{cases} x \\ \frac{1}{1+e^{-kx}} \\ \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \\ e^{-\frac{x^2}{2}} \end{cases} , (3.15)$$

where y^i is the input of the previous layer as in Figure 3.2, the quantities ω_{ij} are the weights associated with the incoming connection and b_j are the biases.

The ANN used in this study consists of n input neurons which corresponds to the input variables, a single hidden layer and an output neuron. In the training procedure the weights are arranged to give the expected output (see section 3.8), 1 for the signal and 0 for the background.

When the network is properly trained, it can give an efficient response also to a statistically independent testing sample.

3.6 Boosted decision trees (BDT)

A decision tree is a method very similar to rectangular cuts. It consists of a sequence of binary decisions using a single variable at a time until a stop criterion is reached. In this way regions of signal and background are selected. A simplified scheme is presented in Figure 3.3.

The advantage of a decision tree, with respect to the standard rectangular cuts, is that this method allows us to select disconnected regions in the parameter space. In the boosting of decision trees, several decision trees are built from the same training sample by re-weighting the events.

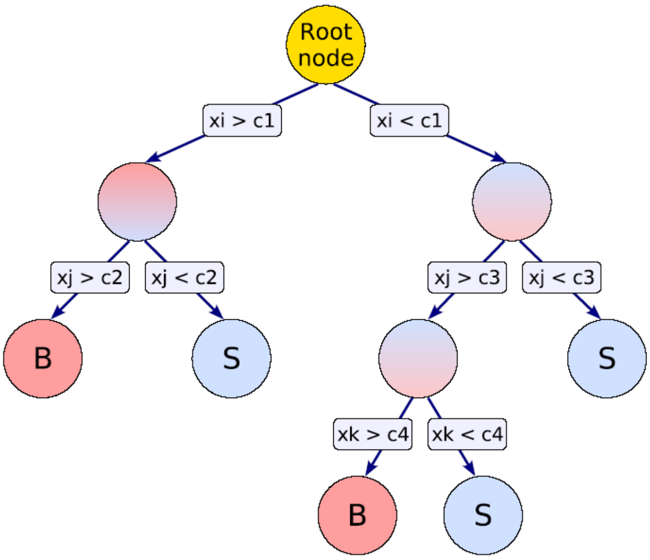


Figure 3.3: Schematic illustration of a decision tree.

The final classifier is given by a weighted majority vote of the different decision trees.

3.7 Fisher discriminant

The Fisher linear discriminant is a criterium able to choose a linear combination of the input variables that can be used for signal-background separation. It is described in [65].
The method consists of the choice of a particular axis in the iper-space of the variables. The Fisher function defined in 3.16 is a linear combination of the input variables.

$$F^k = \sum_{i=1}^{n} \lambda_i x_i^k \tag{3.16}$$

The difference between the mean of signal and background and the standard deviation with respect to the new axis are respectively given by the relations 3.17 and 3.18.

$$D = \sum_{i=1}^{n} \lambda_i d_i = \overrightarrow{\lambda} \cdot \overrightarrow{d} \tag{3.17}$$

$$S = \sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_i \lambda_j S_{ij} = \overrightarrow{\lambda} S \overrightarrow{\lambda} \tag{3.18}$$

The coefficients in 3.16 are chosen to maximize the significance of the difference between the two categories of events: $\frac{D^2}{S}$. It can be shown that the solution is given by the coefficients $\overrightarrow{\lambda}$ such that the relation 3.19 is satisfied.

$$S \overrightarrow{\lambda} = \overrightarrow{d}. \tag{3.19}$$

It is easy to show that the Fisher discriminant can exactly eliminate linear correlation in the input variables.

3.8 Training methods and the overtraining problem

Each multidimensional method needs at least two different data samples: a training sample for the parameter settings and a testing sample for the method performance evaluation.
The training phase is different for the different methods. In the likelihood and PCA methods the training sample is used for the construction of the *probability density functions (pdfs)* of the variables.

In the decorrelated likelihood method the training sample is used, in addition, for the construction of the decorrelation matrix.

Concerning ANN, there are several training methods. We have used the *back-propagation* method in which the output for each event is compared with the expected output and an error function is built. The weights are then adjusted to minimize the output error function.

In the Fisher discriminant method the training sample is used to extract the coefficients for building the Fisher function.

Finally in boosted decision trees the set of cuts that give the best signal-background separation are chosen with the training sample.

Testing is usually done with a statistically independent sample, called the testing sample.

The discrepancy between the performance in the training and testing samples is called over-training. In methods with few parameters, like the PCA or the decorrelated likelihood, the over-training is due to lack of statistics in the training sample. In methods with many degrees of freedom, like the ANN or the BDT, the training procedure is iterative, after a certain number of iterations, the performance improves with respect to the training sample and worsens with respect to a statistically independent sample. In non linear methods with many degrees of freedom, to avoid as much as possible the over-training, three statistically independent data sample are commonly used. The first sample is used for training, the second for the optimization of the training procedure, i.e. to choose the optimal point to stop the training, and a third sample is used for testing.

In this study two statistically independent samples for training and testing are always used. The over-training is then measured as the difference in performance between the two samples.

Even if over-training is a typical problem of multivariate methods, standard rectangular cuts can also be affected by it. In fact choosing the cut values in the training sample, it does not guarantee that you will have the same performance in a statistically independent sample.

3.9 The Punzi estimator

The optimization of the selection cuts is a complex problem. In general optimization depends on what we want to measure. In this study we are interested in optimizing the cuts for the search of NP.

The optimization of the experiment parameters for the search of new phenomena is a very common problem in physics. The cuts are usually chosen to maximize the significance with respect to the quantity we are interested in, i.e. branching ratio measurements.

However, several different definitions of significance can be found in literature. The most common definitions are those in Eq. 3.20 and 3.21.

$$\frac{S}{\sqrt{S+B}} \tag{3.20}$$

$$\frac{S}{\sqrt{B}}, \tag{3.21}$$

where S and B indicate the number of signal and background events respectively. Even if these quantities are often used to evaluated the significance, in [59] it was shown that both of them have problems when the experiment is designed for new phenomena searches.

The definition 3.20 has the problem that, being not linear in S, it is necessary to know the S value to maximize the sensitivity. However the signal level often depends on unknown parameters and sometimes we would like to carry out a blind analysis optimizing the cuts before the measurement.

The definition 3.21 is linear on S. However, it has the unwanted property that the definition breaks down at low values of B. For instance using the definition 3.21 we would prefer a situation in which $S = 0.1$ and $B = 10^{-6}$ than a situation in which $S = 10$ and $B = 1$.

In [59] it was proposed a sensitivity definition which is well behaved for low value of B and is independent on the quantity S. The Punzi criterium states that maximum sensitivity is obtained at the smallest cross-section measurable by the experiment. H^0 being the hypothesis that no signal is present and H^m the hypothesis that there is some signal, there is a minimum number of events n_{min} for which we would claim a discovery for the signal. The two hypotheses H^0 and H^m, in general predict different probabilities for the occurrence of n_{min}. In particular n_{min} is distant a and b standard deviations from the hypotheses H^0 and H^m respectively. Assuming that the two Poissonian distributions H^0 and H^m can be approximate with Gaussian distributions it can be shown that the maximum sensitivity is obtained for the maximum of the function:

$$\Pi = \frac{\epsilon(t)}{b^2 + 2a\sqrt{B(t)} + b\sqrt{b^2 + 4a\sqrt{B(t)} + 4B(t)}}. \tag{3.22}$$

A simpler expression for the 3.22 is obtained when the choice $a = b$ is made:

$$\Pi = \frac{\epsilon(t)}{a/2 + \sqrt{B(t)}}. \tag{3.23}$$

The Punzi criterium also can be applied when the maximization involves more than one parameter. In this case we have a set of observables X. The set of values of the X variables for which the H^0 hypothesis is rejected with a certain confidence level α, is called critical region. $1-\beta_\alpha(m)$ being the power of the test, i.e. the probability that X will fall in the critical region assuming H^m. The sensitivity can be defined as: *the region of the parameters m for which the power of the test chosen is greater or equal to the confidence Level for the limits in case of non discovery:* $1 - \beta_\alpha(m) > CL$. This definition is very general and can be applied to many physical problems.

In this thesis the choice of the optimal point for the cut on the multivariate output variable is done according with this criterium, using Eq. 3.23. It was also verified that this was the point with maximum sensitivity with the program *bayes.f* ([67]).

Chapter 4

$B^0_{s,d} \to e^{\pm}\mu^{\mp}$

4.1 Signal and Background

For this study the signal sample consists of the decays $B^0_{s,d} \to e^{\pm}\mu^{\mp}$, however for simplicity only the $B^0_s \to e^{\pm}\mu^{\mp}$ sample was generated assuming the efficiency to be the same for both the signal samples.
This decay was generated with the EvtGen package using the *psp* method. To simulate the physics process and the detector response the packages Gauss v15r19 and Boole v6r5 were used. The reconstruction was done with the software Brunel v24r5 and for the analysis the package DaVinci v12r18[1] was used. The signal invariant mass distribution is shown in Figure 4.1. For the momentum measurement, the bremsstrahlung recovering algorithm described in Section 2.9, was applied. A comparison between the signal invariant mass, obtained using the bremsstrahlung recovering algorithm and obtained without using it, is shown in Figure 4.2.

The main background is expected to come from $b\bar{b} \to e^{\pm}\mu^{\mp}$ events, which consists of an electron and a muon of opposite charges directly coming from the two b-hadrons, which decay semileptonically. This hypothesis was verified a posteriori (see Section 4.5). A large sample of these events was generated and analyzed. Moreover, this sample allowed also the evaluation of other potential sources of background taking into account the proper scale factors.

---

[1] At the time of the note of [5] the package DaVinci v12r17 was used, but this version was dismissed. The analysis was redone with the DaVinci v12r18 and was verified the agreement of the two software versions.

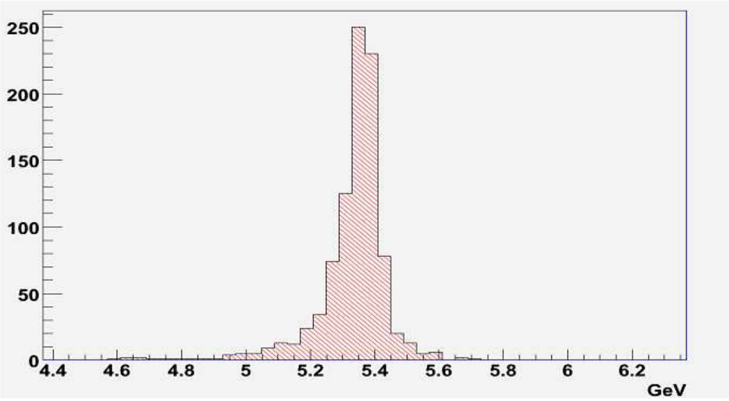


Figure 4.1: Invariant mass distribution for the $B^0_s \to e^\pm\mu^\mp$ channel.

-Events with one lepton coming from a b-hadron decaying semi-leptonically and the other lepton coming from a particle which doesn't come from a b-hadron decay chain, possibly also with misidentification.
-Events with both leptons coming from the same b-hadron decay chain, when the b-hadron decays with a double semi-leptonic decay.
-Events with both leptons coming from particles which come from the primary vertex, possibly with misidentification.
A more detailed discussion about these sources of background can be found in Section 4.5. In addition to these sources of background, some specific background samples not present in the $b\bar{b} \to e^\pm\mu$ sample, were generated. A potential background consists of $B(\Lambda_b) \to h_a^+ h_b^-$ decays with $h_{a(b)}^{+(-)} = \pi^{+(-)}, K^{+(-)}, p^{+(-)}$. These decays have the same topology as the signal and can be distinguished only with the invariant mass and using particle identification cuts.
Another potentially dangerous source of background is the sample $B^+ \to J/\psi K^+$ with the kaon misidentified as an electron or a muon and the J/ψ decaying into two leptons. Finally the possible background due to the $B_c \to J/\psi\mu\nu$ sample with the J/ψ decaying into two electrons was analyzed. In Table 4.1 all the events analyzed for this study are listed.
The background due to the Drell Yan process and due to leptonic and semileptonic decays of the Z^0 and $W^\pm$ bosons is considered negligible be-

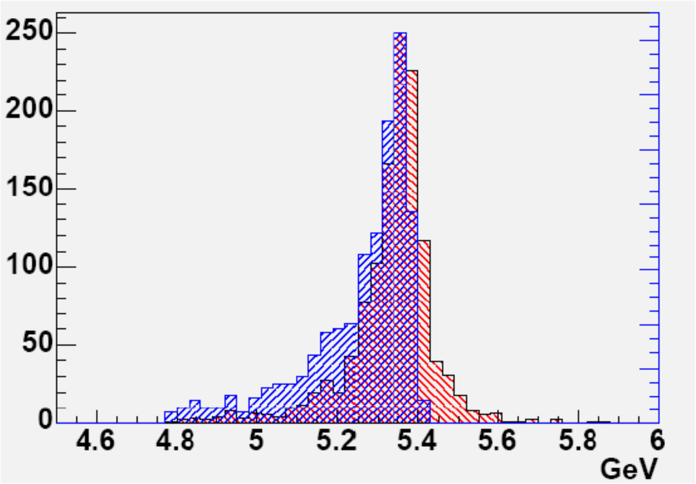
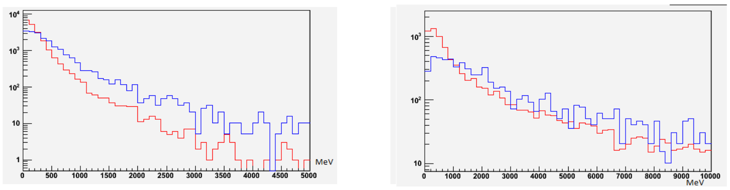


Figure 4.2: Invariant mass distribution for the $B_s^0 \to e^{\pm}\mu^{\mp}$ channel with (red histogram) and without (blue histogram) the bremsstrahlung correction.

cause of the small cross section. Moreover, the background of leptons coming from the PV is highly suppressed because of the cut on the impact parameter at the trigger level.

Concerning the background due to $c\bar{c}$ events, it was not studied in detail. However, it is expected to be less relevant than the background due to $b\bar{b}$ events. In Figure 4.3 the transverse momentum and the invariant mass distributions for $c\bar{c} \to \mu^+e^-$ and for $b\bar{b} \to \mu^+e^-$ events are shown. Cutting at $P_T > 1200MeV$ and at $Mass(\mu, e) > 4000MeV$ we expect the $c\bar{c}$ events to be about $\frac{1}{4}$ of the $b\bar{b}$ events (properly taking into account the cross sections). This is in agreement with [70], in which it was found that after the L0 and HLT single muon trigger 90% of the events contain b or c quarks, out of which 78% are $b\bar{b}$ events.

This study was originally done with the *DC04* detector description. The new sample generation with the *DC06* detector description has not been finished yet. However the trigger study discussed in Section 4.6 and the preselection cuts discussed in the next section were optimized on *DC06* data samples which are expected to be more similar to real data.



(a) P_T (MeV) (b) Invariant Mass (MeV)

Figure 4.3: Comparison between the P_T and the invariant mass distributions for $c\bar{c} \to \mu^+ e^-$ events (red lines) and $b\bar{b} \to \mu^+ e^-$ events (blue lines).

Table 4.1: Branching fractions, expected number of events in 1fb^{-1} at LHCb within the detector acceptance, number of analyzed events, equivalent integrated luminosity.

Channel	Branching ratio	events in $1fb^{-1}$	Analyzed events	L (pb^{-1})
B^0_d		$6.8 \cdot 10^{10}$		
B^0_s		$1.73 \cdot 10^{9}$		
$b\bar{b} \to e^-\mu^+$	$1.1 \cdot 10^{-2}$	$6.8 \cdot 10^{10}$	$5M$	2.1
$b\bar{b} \to e^+\mu^-$	$1.1 \cdot 10^{-2}$	$6.8 \cdot 10^{10}$	$5M$	2.1
$B^0_d \to \pi^+\pi^-$	$5.2 \cdot 10^{-6}$	$3.5 \cdot 10^{5}$	$40k$	113
$B^0_d \to K^+\pi^-$	$2 \cdot 10^{-5}$	$1.4 \cdot 10^{6}$	$8k$	5.8
$B^0_s \to K^+K^-$	$2.5 \cdot 10^{-5}$	$4.3 \cdot 10^{5}$	$8k$	18
$\Lambda_b \to p\pi^-$	$2.1 \cdot 10^{-5}$	$1.3 \cdot 10^{6}$	$8k$	5.9
$\Lambda_b \to pK^-$	$7.8 \cdot 10^{-5}$	$3.6 \cdot 10^{6}$	$8k$	2
$B_c \to J/\psi\mu\nu$	$2 \cdot 10^{-5}$	$2 \cdot 10^{6}$	$101k$	48
$B^+ \to J/\psi K^+$	$1 \cdot 10^{-3}$	$9.7 \cdot 10^{4}$	$200k$	1023
$B^0_s \to e^{\pm}\mu^{\mp}$			$102k$	

Table 4.2: $DC04$ soft-preselection cuts.

particle	variable	cut value
e	DLL	> -5
μ	DLL	> -15
e, μ	P_t	$> 1GeV/c$
e, μ	Vertex χ^2	< 49
B_s^0	P_t	$> 0.5GeV/c$
B_s^0	$\frac{IP}{\sigma_{IP}}$	< 5
e, μ	Invariant Mass	$> 4.0GeV/c^2 \quad , \quad < 7.0GeV/c^2$
B_s^0	$Angle(\overrightarrow{P}, \overrightarrow{r}_{B_s} - \overrightarrow{r}_{PV})$	$< 10mrad$

Table 4.3: $DC06$ soft-preselection and hard-preselection cuts. The soft cuts are used for the HLT exclusive selection, the hard cuts for the signal preselection before to apply the multivariate methods.

Variable	soft-preselection cut	hard-preselection cut
Vertex χ^2	< 14	< 10
$Angle(\overrightarrow{p}_B, \overrightarrow{r}_{BV} - \overrightarrow{r}_{PV})$	$< 32mrad$	$10mrad$
$IP_{B_s^0}/\sigma_{IP}$	< 6	< 6
$\mid m(e\mu) - m_{B_s^0} \mid$	$< 1GeV$	$< 1GeV$

4.2 Preselection

A soft set of preselection cuts was studied, which can be used for the HLT exclusive selection. At the time of note [5] a preselection with the $DC04$ data was studied. The cuts used are listed in Table 4.2.

However, with the $DC06$ data a new preselection was studied. The only requirement was to have a rejection of at least 10^{-3} on the $b\bar{b}$ inclusive sample [2]. The soft preselection cuts are listed in Table 4.3. These cuts are the χ^2 of the fit of the vertex between the two leptons, the significance of the impact parameter of the reconstructed B_s^0 meson with respect to the reconstructed primary vertex, the pointing angle, which is the angle between the direction from the primary vertex to the secondary B_s^0 decay vertex and the B_s^0 momentum and finally the difference between the invariant mass of the lepton pair and the invariant mass of the B_s^0 meson. In events with more

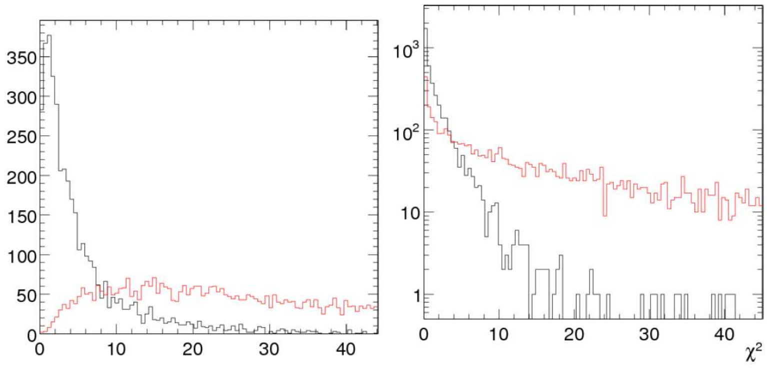
---

[2] Notice that at the moment of writing the HLT is still under development so it is possible that more requirements would be added in the future.

than one primary vertex, it was chosen the one associated with the smallest impact parameter significance ($sIPS$) of the B^0_s. With these cuts a rejection of $2 \cdot 10^{-4}$ in the inclusive $b\bar{b}$ sample was achieved.

The final analysis was done comparing various multidimensional methods. For some variables with high discriminating power we found that optimal performance could be achieved when a cut was directly applied. For this reason a preselection with harder cuts on the χ^2 of the fit of the vertex between the two leptons and on the pointing angle was applied (Table 4.3). These two variables are shown in Figure 4.4.

In the preselection, the *standard loose particle* identification cut was applied.



(a) Pointing angle (mrad). (b) χ^2 of the two leptons.

Figure 4.4: Additional variables used in the hard-preselection. The black histograms are signal events and the red histograms $b\bar{b} \to e^\pm\mu^\mp events$.

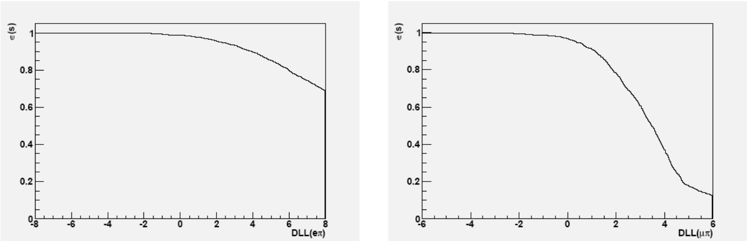
This corresponds to $DLL(\mu - \pi) > -10$ and $DLL(e - \pi) > -2$ [3]. In Figure 4.5, the efficiency and misidentification probability, for charged hadrons, as a function of the DLL are shown.

4.3 Choice of selection variables

For the final selection various multidimensional methods were compared. All these methods are constructed combining the information of the five

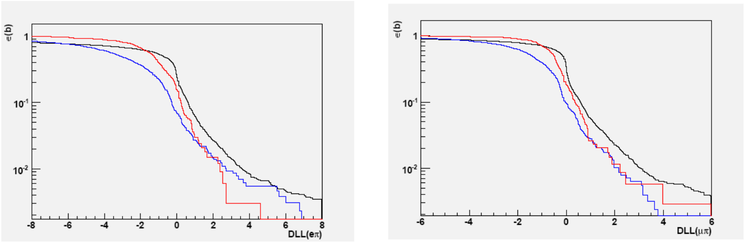
---

[3] With the DC06 data the *standard loose muons* are identified by the value 1 of the variable *IsMuon*. This variable assumes the value 1 when there are hits in each stations of the muon detector matched with a track without any other cut in the $DLL(\mu - \pi)$.



(a) Efficiency versus $DLL(e-\pi)$ for true electrons.

(b) Efficiency versus $DLL(\mu-\pi)$ for true muons.



(c) Efficiency versus $DLL(e-\pi)$ for charged pions (black line), charged kaons (blue line) and protons (red line)

(d) Efficiency versus $DLL(\mu-\pi)$ for charged pions (black line), charged kaons (blue line) and protons (red line

Figure 4.5: Efficiency as a function of the $DLL(\mu-\pi)$ and $DLL(e-\pi)$ for leptons and charged hadrons.

variables listed below. Because the main background is assumed to be the $b\bar{b} \to e^\pm\mu^\mp$, the selection was optimized to reject this kind of background. The term background will be used to indicate the $b\bar{b} \to e^\pm\mu^\mp$ sample when it is not explicitly indicated otherwise. A schematic illustration of signal and background topology is shown in Figure 4.11. The selection variables will be briefly described in the following.

-B^0_s **transverse momentum**: $P_t(B^0_s)$.

The transverse momentum of the reconstructed B-hadron is expected to be higher for the signal than for the background. The distributions are shown in Figure 4.6.

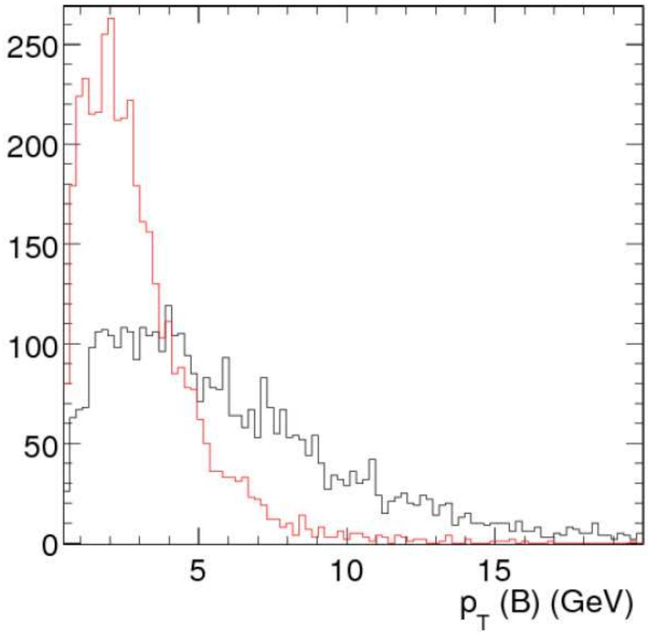


Figure 4.6: Transverse momentum of the B^0_s for the signal (black histogram) and for the $b\bar{b} \to e^\pm\mu^\mp$ background (red histogram).

-**Impact parameter significance of the** B^0_s: $IP_{B^0_s}/\sigma_{IP}$.

The significance of the impact parameter of the reconstructed B^0_s meson with respect to the primary vertex is expected to be peaked at low value for the signal as the B^0_s meson really comes from the primary vertex. For the background the vertex is an accidental crossing between two charged tracks. The $IP_{B_s^0}/\sigma_{IP}$ is therefore expected almost flat for the background. These distributions are shown in Figure 4.7.

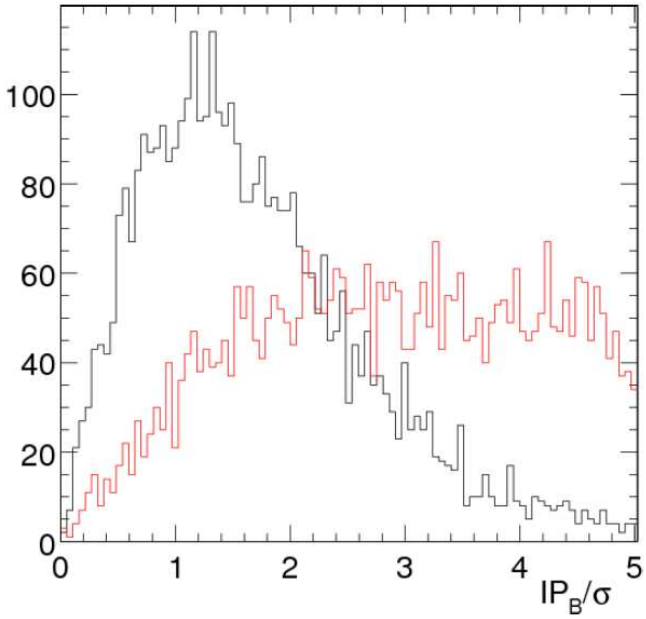


Figure 4.7: Impact parameter significance for the signal (black histogram) and for the $b\bar{b} \to e^{\pm}\mu^{\mp}$ background (red histogram).

-**Angular distance between the two leptons**: $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. This variable is essentially the Lorentz invariant angular distance between the two leptons. The distributions for the signal and background is shown in Figure 4.8. It can be seen that this variable is larger for the background than for the signal.

-**Flight distance significance**: $FS = \frac{r_B - \overrightarrow{r}_{PV}}{\sigma}$. This variable is the flight distance significance of the B_s^0 meson. The FS variable is expected to be smaller (on average) for the background than for the signal. This is due to the fact that this variable is proportional to the lifetime which is an exponential distribution. The flight distance (FD) probability distribution

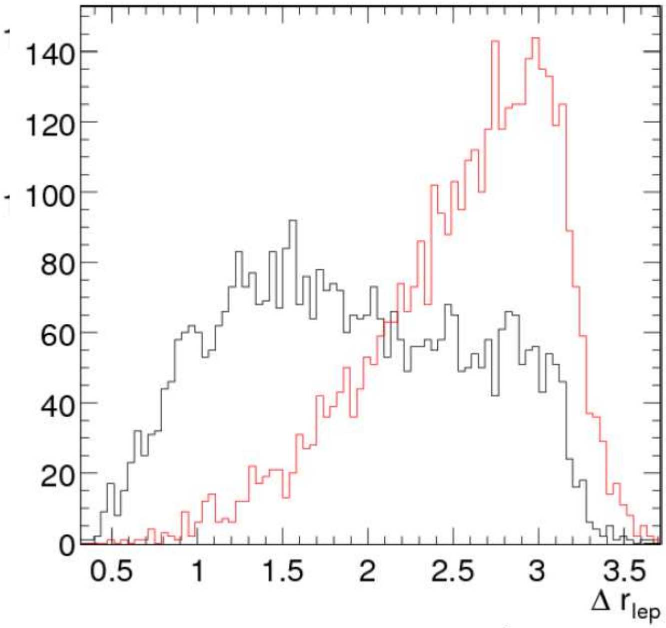


Figure 4.8: $\Delta R(e,\mu)$ for the signal (black histogram) and for the $b\bar{b} \to e^{\pm}\mu^{\mp}$ background (red histogram).

for the fake B^0_s meson in the background, is approximately given by the square of its probability distribution for the signal, integrating over FD. Therefore an average smaller FS for the background is expected. The FS distribution for signal and background is shown in 4.9.

-**Isolation** : $I_B = \frac{P_t(B^0_s)}{P_t(B^0_s)+\sum_i (P_t)_i}$.
This variable takes into account the isolation of the two leptons, weighting it with the momentum of the reconstructed B^0_s meson.

The sum in the formula of the isolation runs over the transverse momentum of all the particles that have a distance of minimum approach with respect to one of the two leptons smaller than 3σ. σ being the error on the measurement of the distance of minimum approach. The Isolation distributions for signal and background are shown in Figure 4.10. This definition of isolation was adopted to maximize the difference between the background and the signal. This choise is the result of a detailed study, comparing various methods ([37]). In Figure 4.12 the signal efficiency versus the background efficiency

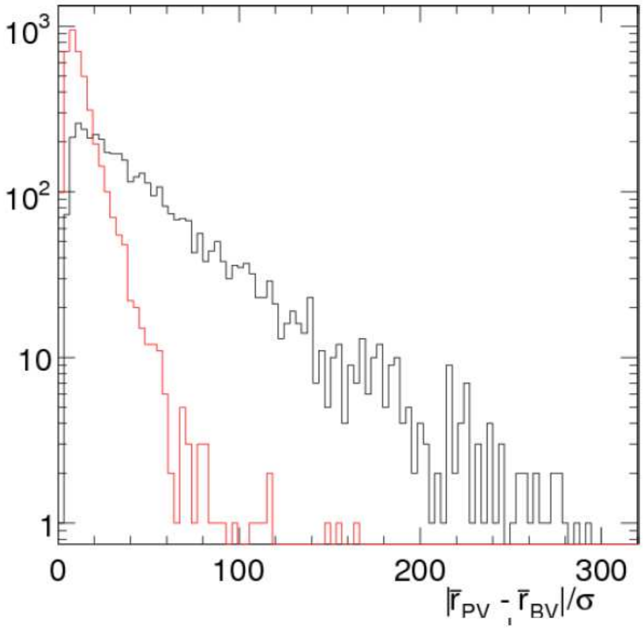


Figure 4.9: Flight distance significance for the signal (black histogram) and for the $b\bar{b} \rightarrow e^{\pm}\mu^{\mp}$ background (red histogram).

using different isolation definitions is shown. The blue curve is obtained summing over all the particles with $\frac{IP}{\sigma_{IP}} < 3$ with respect the B_s^0 decay vertex, the red curve is computed where the sum is extended to all the particles inside a cone of radius $\Delta R < 2$ with respect to the B_s^0 flight direction, finally the black curve is the chosen definition of isolation. The value of the parameters of the different methods used for the comparison were previously obtained optimizing the signal-background separation.

We found that in LHCb the best criterium is to use the lepton track isolation. This is due to the fact that the background is mainly constituted by semi-leptonic decays of b-hadrons in which other out-coming particles are emitted close to the lepton tracks, for the signal we expect lepton tracks isolated instead, because the b-hadron has a good fragmentation.

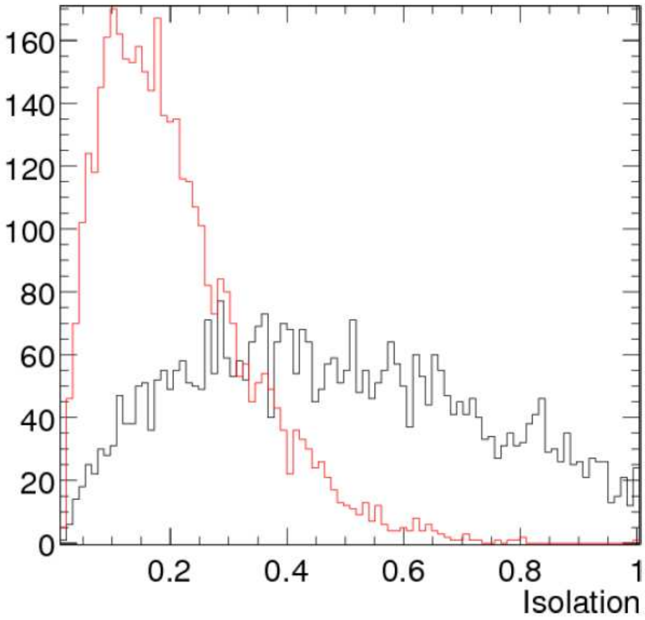
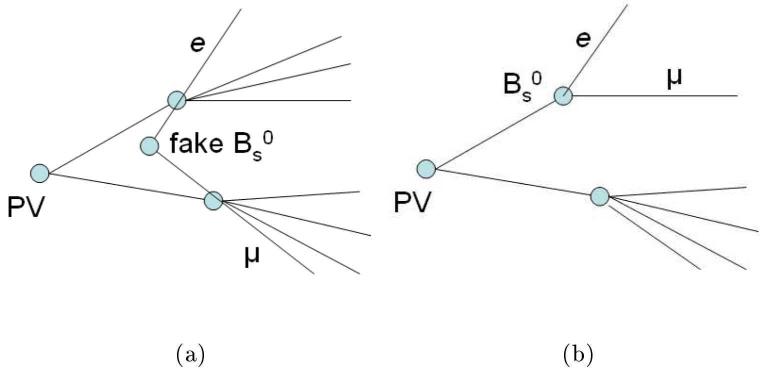


Figure 4.10: Isolation for the signal (black histogram) and for the $b\bar{b} \to e^{\pm}\mu^{\mp}$ background (red histogram).

4.4 Multivariate analysis

The multidimensional methods described in Chapter 3 were used for the signal selection of the $B^0_s \to e^{\pm}\mu^{\mp}$ channel. Once again the optimization is done looking at the $b\bar{b} \to e^{\pm}\mu^{\mp}$ which is the main source of background. In Figure 4.13 the signal efficiency versus the number of background events in 1fb^{-1} of integrated luminosity, for the different multivariate methods, is shown. As can be seen, the best method is the decorrelated likelihood, as this method has the best efficiency for a given level of background. It was also the best method concerning the over-training problem. In Table 4.4 the difference in the efficiency between training and testing samples for the different methods is listed.

In all the multidimensional methods the output is a single variable, in which all the information of the variables previously discussed is mixed. In Figures 4.14(a)-4.14(f) the outputs of the multidimensional methods for signal and background, after a sigmoid inverse transformation, are shown.



(a) (b)

Figure 4.11: Schematic illustration of the topology of a typical signal event (b) and of a typical background event (a).

Table 4.4: Estimation of the over-training for the different multivariate methods. The over-training is measured as the performance difference in the training and testing sample.

Method	Over-training
Likelihood	0.002
Decorralated Likelihood	-0.001
Fisher Discriminant	0.008
Neural Net	0.003
Decision Trees	0.089
PCA	0.003

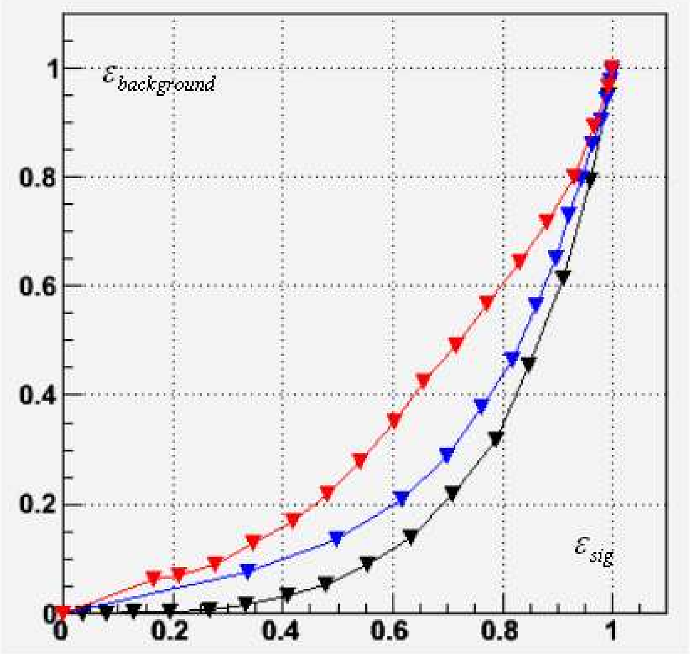


Figure 4.12: Efficiency on the signal versus the efficiency on the background for different Isolation definitions. The red line is the B^0_s flight direction isolation, the blue line is the B^0_s vertex isolation, the black line is the track isolation (used in this study).

4.5 Background composition

A study of the background categories according to the MC associators of the DaVinci package was carried out in [5]. This study had shown that the main background consists of $b\bar{b} \to e^{\pm}\mu^{\mp}$ events. With the new preselection, the sensitivity study was completely carried out again comparing various multidimensional methods, assuming that the main background consists of $b\bar{b} \to e^{\pm}\mu^{\mp}$ events.

Table 4.5 shows the background composition after the $DC04$ preselection cuts for the $b\bar{b} \to e^{\pm}\mu^{\mp}$ sample. As can be seen, the event with the two leptons coming directly from the b-hadron (or through bremsstrahlung for the electrons) are largely dominant. The contribution indicated as $PV - PV$, in which none of the two leptons come from the b-hadron decay chain, where the background due to the misidentification of $\pi^{\pm}$ or $K^{\pm}$ coming from the

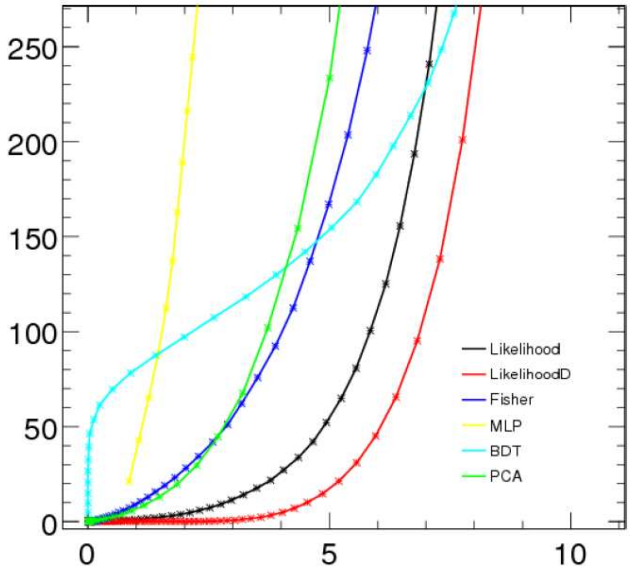


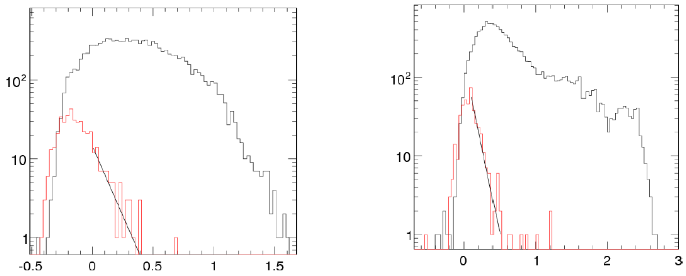
Figure 4.13: Signal efficiency versus the number of background events in 1fb^{-1} for various multidimensional methods.

primary vertex is also included, has a scaling factor 100. Even taking into account the scaling factor the $PV - PV$ background is negligible.

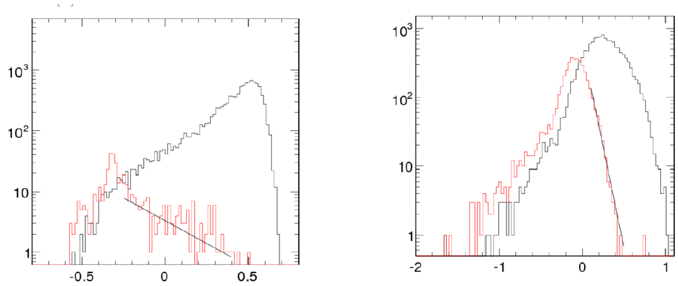
Concerning the background due to one of the two leptons not coming from the b-hadron decay chain ($b - PV$ and $PV - b$), or due to the two leptons coming from the same b-hadron decay chain ($b - c$ same), the scaling factor is about 20.

Among these sources of background the only relevant source is that of electrons coming from converted photons which don't come from the b-hadron decay chain ($PV - b$). However these events are concentrated at low value of the likelihood ratio ($LR < 0.1$) [5]. Therefore it is expected that this source of background is negligible after the final selection.

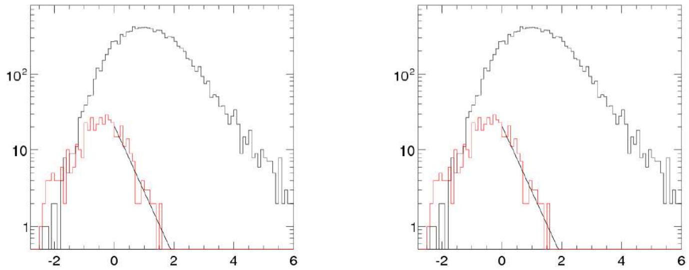
The background due to $B(\Lambda_b) \rightarrow h_a^+ h_b^-$ (two body charmless) events is topologically and kinematically very similar to the signal. The main kinematical difference is the invariant mass distribution, in fact when a pion or a kaon is misidentified for an electron or a kaon for a muon the invariant mass is systematically lower than the mass of the decaying particle. The only specific



(a) Likelihood. (b) Decorrelated likelihood.



(c) Boosted decision trees. (d) PCA method.

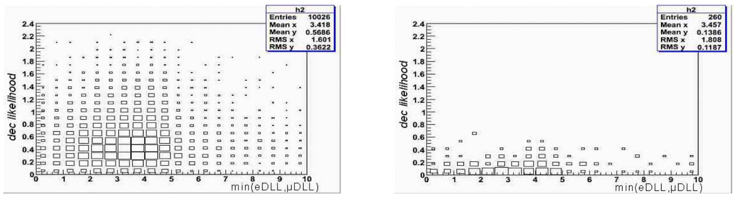


(e) Fisher discriminant. (f) Artificial neural net.

Figure 4.14: Multivariate method output distributions for the signal (black histogram) and the background (red histogram).

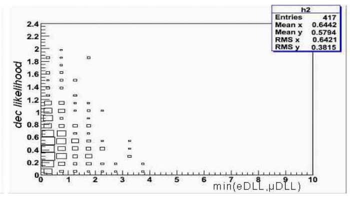
Table 4.5: Background categorization for $b\bar{b} \rightarrow e^{\pm}\mu^{\mp}$ sample after the preselection cuts studied for the *DC04* data challenge.

MCID	origin	soft-preselection
e, μ (or h)	$b - b$	2877
e, μ (or h)	$b - PV$	16
e, μ (or h)	$PV - b$	62
e, μ (or h)	$PV - PV$	1
e, μ (or h)	$b - c$(same)	23
μ, 0	ghost	139
e, 0	ghost	8
total		3105



(a) $B_s^0 \rightarrow e^{\pm}\mu^{\mp}$ decay.

(b) $b\bar{b} \rightarrow e^{\pm}\mu^{\mp}$ events.



(c) $B_d^0 \rightarrow \pi^{+}\pi^{-}$ decay.

Figure 4.15: Decorrelated likelihood versus the minimum of the DLL for muon and electrons.

cuts applied to this kind of background were muon and electron identification cuts. Concerning muon identification, a $DLL(\mu - \pi) > 2$ cut was applied. This cut allow us to have a pion misidentification probability of about 1%, with 78% muon identification efficiency. For the electrons a $DLL(e - \pi) > 4$ cut, corresponding to 90% electron identification efficiency and to about 1% pion misidentification probability, was applied. To evaluate the number of remaining events after the cuts, the likelihood efficiency and the misidentification probability were factorized. This can be done as the DLL and the decorrelated likelihood are, in good approximation, independent from each others, as can be seen in the Figure 4.15.
After the identification cuts, the $B \to h_a^+ h_b^-$ background was found negligible with respect to the main background. In Figure 4.16, the invariant mass distribution, for the $b\bar{b} \to e^\pm \mu^\mp$ sample and for the various $B \to h_a^+ h_b^-$ decays, after the preselection and the application of the L0 trigger, but before the application of the particle identification cuts, are shown.

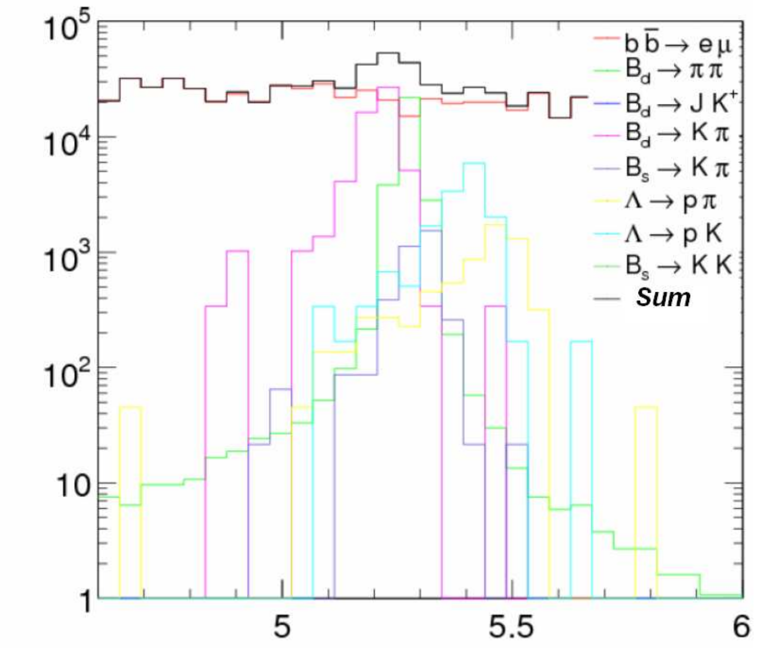


Figure 4.16: Invariant mass distribution for $b\bar{b} \to e^\pm \mu^\mp$ and for $B(\Lambda_b) \to h_a^+ h_b^-$ events in $1\mathrm{fb}^{-1}$.

Table 4.6: L0 Trigger efficiency after the soft preselection for the signal. The efficiency is listed for the various L0 channels.

Trigger Channel	Accepted events
Total L0	90.50%
Electron L0	29.75%
Photon L0	5.37%
Hadron L0	15.43%
π^0 Local L0	20.11%
π^0 Global L0	22.04%
π^0	24.52%
Muon L0	59.64%
Sum Muon L0	83.06%

4.6 Trigger

The Level 0 trigger was studied for the signal with $DC06$ data sample. In Table 4.6, all the $L0$ channels with the relative efficiency, after the soft preselection cuts, are listed. The total efficiency is 90.5%. However, this efficiency increase up to the 97% after the selection. At the time of writing, the HLT was not yet ready, particularly concerning the electron part. However we expect that the HLT performance will not be too different from the old $L0+L1$. In this study was found that the $L0 + L1$ efficiency for $DC04$ data sample was about 86% [5].

4.7 How to deal with real data

4.7.1 Training sample for the background

In the present study, the LHCb sensitivity to the $B^0_{s,d} \to e^{\pm}\mu^{\mp}$ decay is evaluated with the decorrelated likelihood method. However to use the multivariate methods, it is necessary to have a training sample for both signal and background. The training sample is needed to built the *pdf*s and the decorrelation matrix in the decorrelated likelihood and for the neural net training in the artificial neural net method. In any case, training samples are necessary for all the multivariate methods. In this study half of the samples of signal and background were used for the training procedure and

Table 4.7: Contamination of the sidebands from two body charmless decays for various DLL cuts.

$DLL(\mu-\pi)$	$DLL(e-\pi)$	Contamination
> -10	> -2	4%
> 0.1	> 1.5	1%
> 2	> 4	0.6%

the other half for testing. Moreover, it is necessary to train the methods knowing before which is signal and which is background.

Obviously in real data we don't know which is signal and which is background. For the background we can train the multivariate methods by looking at the sidebands.

As can be seen in Figure 4.16 the sidebands are contaminated by $B(\Lambda_b) \to h_a^+ h_b^-$ events. This is a problem because the $B(\Lambda_b) \to h_a^+ h_b^-$ decays have a likelihood distribution similar to the signal. In Table 4.7 three DLL cuts, with the correspondent likelihood contamination of two body charmless decays, are listed.

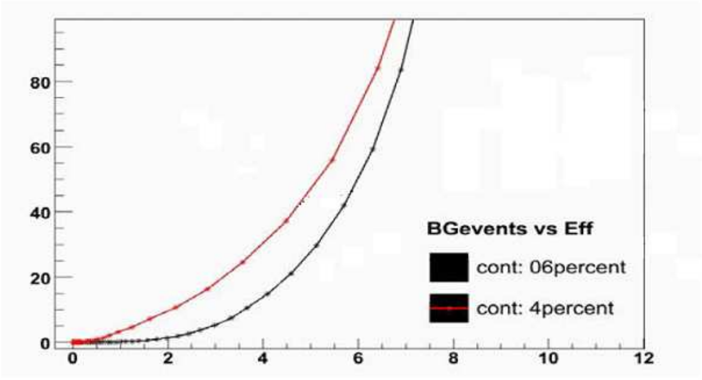


Figure 4.17: Signal efficiency versus number of background events in 1fb^{-1} for two levels of contamination of the sidebands from two body charmless decays. The red line is computed for 0.6% of contamination, the red line for 4% of contamination.

In Figure 4.17, the decorrelated likelihoods, for 0.6% contamination of the side bands (red curve) and for 4% contamination of the sidebands, are shown. As can be seen the effect of the contamination of the sidebands is to reduce the power of discrimination of the decorrelated likelihood[4]. This problem is also present in other analysis of LHCb such as the $B_s^0 \rightarrow \mu^+\mu^-$ channel. We have proposed two possible solutions. One possibility is to eliminate the part of the sidebands where the peaks are expected. For instance we can take the

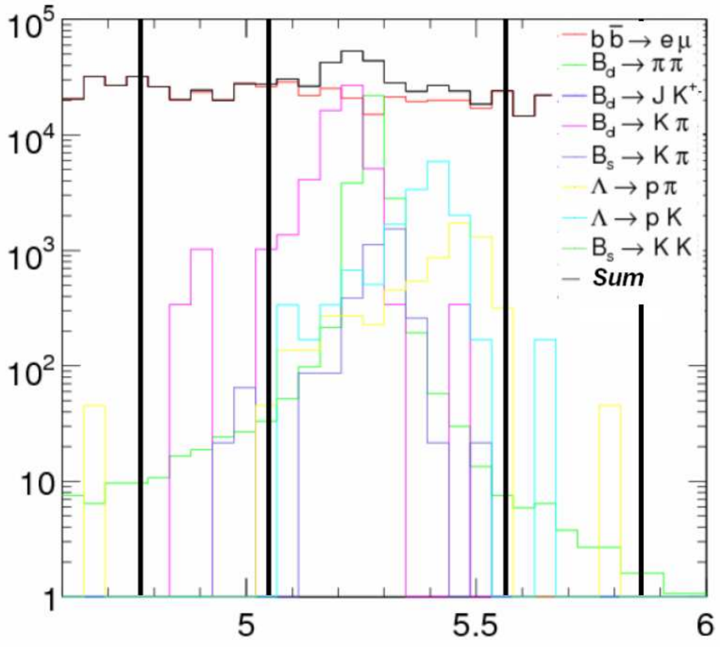


Figure 4.18: Invariant mass distribution for $b\bar{b} \rightarrow e^\pm\mu^\mp$ and for $B(\Lambda_b) \rightarrow h_a^+ h_b^-$ events in 1fb^{-1}. Possible sidebands that can be used to extract the background variable distributions for the training of the multidimensional methods in the $B_s^0 \rightarrow e^\pm\mu^\mp$ event selection are also shown.

sidebands as in Figure 4.18. Another solution could be to use a stronger particle identification cut in the sidebands to have lower contamination and to use a looser particle identification cut in the final analysis. This is the approach here adopted.

This approach is justified by the fact that the decorrelated likelihood and the particle identification likelihood are not correlated (Figure 4.15)[5].

---

[4] It can be shown that the same effect is induced in all the multivariate methods.

[5] This hypothesis which is valid in the MC simulation should be verified in the real data

4.7.2 Training sample for the signal

Another problem is the training sample for the signal. We have to verify that the variable distributions, used for the signal selection, are the same as in the MC simulation and in real data. Some variable distributions can be simply recovered by tuning the MC using the real detector resolution.

However, some other variables depend on parameters for which a MC tuning procedure is more difficult. The isolation variable, for instance, depends on the number of particles coming from the PV and on their P_t, these parameters are difficult to re-tune. In addition the Isolation variable depends also on the other signal kinematical variables, as for instance the P_t of the B-meson.

The channels $B(\Lambda_b) \to h_a^+ h_b^-$ have very similar distributions of kinematical and geometrical variables to the signal and also the isolation variable is expected to be very similar. Moreover, as discussed in Section 6.3, an ideal control channel should have the same trigger as the signal.

Therefore a good control channel could be the $B^0_d \to \pi^+\pi^-$, which has a branching ratio of $Br(B^0_s \to \pi^+\pi^-) = (4.9 \pm 0.4) \cdot 10^{-6}$. A comparison between the variable distribution for the signal and the $B^0_d \to \pi^+\pi^-$ decay is shown in Figure 4.19. These distributions are obtained after a selection cut in the decorrelated likelihood, the value of the cut is the same used for the computation of the limit in Section 4.8.

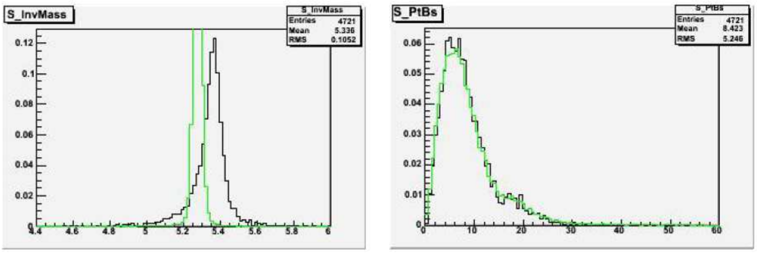
4.7.3 Trigger in real data

To correctly evaluate the limit in case that we will not observe any signal event, or to measure the branching ratio in the case that we observe signal events, we have to know the efficiency of the selection. For this purpose we must use control channels.

However, the $B^0_d \to \pi^+\pi^-$ decay, discussed in the previous Section, is mainly triggered by the hadron trigger, while the Signal is mainly triggered by muon and electron triggers. Moreover, variables like the efficiency of the particle identification for electrons and muons, or effects related to the electron bremsstrahlung, must be taken under control.

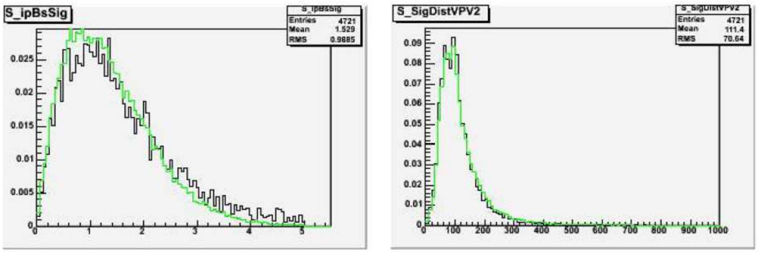
---

with the use of control channels.



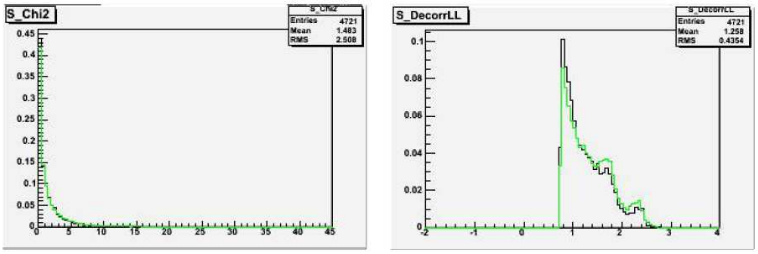
(a) Invariant mass (GeV/c^2)

(b) Transverse momentum of the $B_{s,d}$ meson $P_t(B)$ GeV/c



(c) Impact parameter significance of the $B_{s,d}$ meson

(d) Flight distance significance of the $B_{s,d}$ meson



(e) Vertex χ^2

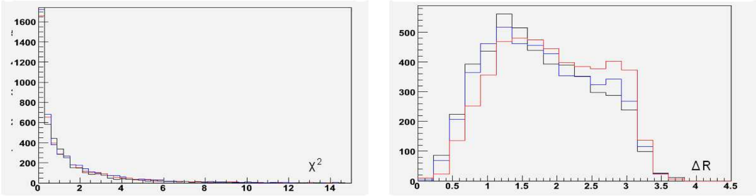
(f) Decorrelated likelihood

Figure 4.19: Comparison of kinematical variable distributions for the signal (black line) and the $B_d^0 \to \pi^+\pi^-$ (green line).

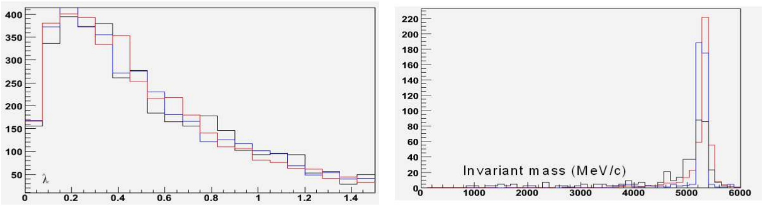
The channels $B^+ \to J/\psi(\to e^+e^-)K^+$ and $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ have a trigger similar to the signal. They are mainly triggered by the single muon and the single electron trigger. Therefore they can also be used for the normalization of the branching ratio in case of discovery. The comparison between the variable distributions for the signal and these control channels is shown in Figure 4.20.

4.8 Sensitivity to the $B^0_{s,d} \to e^{\pm}\mu^{\mp}$ decays

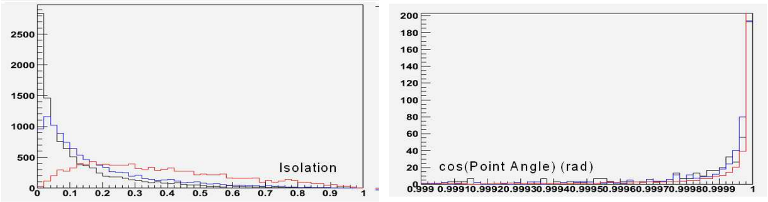
Taking into account all the background sources listed in Table 4.1 it is possible to evaluate the LHCb sensitivity to the $B^0_{s,d} \to e^{\pm}\mu^{\mp}$ channels. For this computation the decorrelated likelihood method, which gives the best result was chosen. The optimal cut point in the decorrelated likelihood was chosen according to the Punzi criterium. In Figure 4.21, the behavior of the Punzi variable as a function of the signal efficiency, is shown. The maximum corresponds to the region of maximum sensitivity. It can be seen that this point is different for the different methods.
In addition to likelihood and identification cuts, a cut on the invariant mass of $\mid m(e\mu) - m_{B^0_s} \mid < 100MeV$, was applied.
Because the expected SM prediction of the branching ratio is zero and the new physics branching ratio is a function of unknown parameters, we give here the upper limit to the branching ratio, in the hypothesis that only background is observed. The limit was computed with the Bayesian method described in [67]. A flat prior *pdf* is assumed for the signal and Gaussian *pdf*s for the acceptance and the background.
The decorrelated likelihood background distribution was fitted with an exponential in the region of high signal purity to extract the expected number of background events, as shown in Figure 4.14(b).
The limits in one year of LHCb running at nominal luminosity are $Br(B^0_s \to e^{\pm}\mu^{\mp}) < 1.3 \cdot 10^{-8}$ and $Br(B^0_d \to e^{\pm}\mu^{\mp}) < 3.2 \cdot 10^{-9}$ at 90% CL. Therefore the LHCb experiment has the possibility to improve on the present limits ([26] , [27]) on these branching ratios, by more than two orders of magnitude just in one year of running at nominal luminosity.
Concerning the usual Pati-Salam model, these limits can be directly trans-



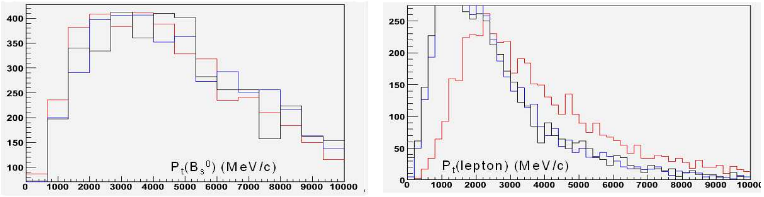
(a) Vertex χ^2. (b) $\Delta R(e, \mu)$.



(c) $\lambda = \frac{P}{m} \cdot c\tau$. (d) Invariant Mass.



(e) Isolation. (f) $\cos(PointAngle)$



(g) Transverse momentum of B^0_s (h) Transverse momentum of leptons.

Figure 4.20: Comparison between the variable distributions for the signal (red histograms) and for the control channels $B^+ \to J/\psi(\to e^+e^-)K^+$ (black histograms) and $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ (blue histograms).

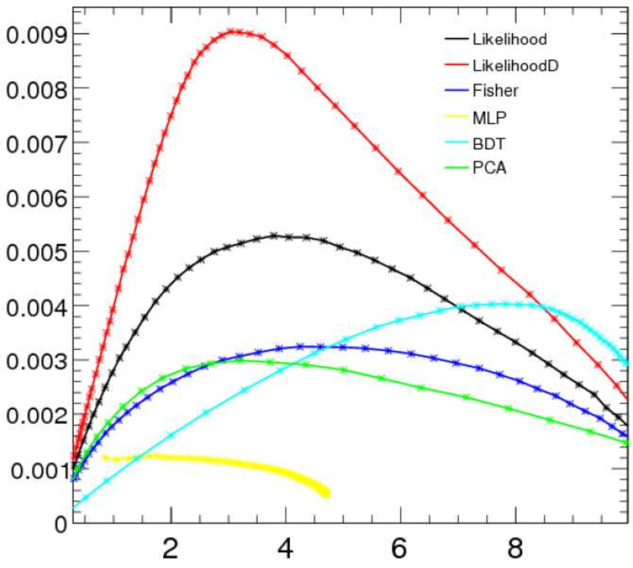


Figure 4.21: Punzi variable as a function of the signal efficiency for the various multidimensional methods. The maximum corresponds to the point of maximum sensitivity in agreement with the Punzi criterium.

lated into limits on LQ masses, as shown in Figures 4.22-4.23, where the present limits are indicated by red lines and the expected LHCb limits in 2fb^{-1} by blue lines.

However, when the LQ mixing with the W boson is considered, an additional F_{mix} factor, according to Eq.1.20, must be taken into account.

Finally, concerning the Pati-Salam $SU(4)_C \times SU(2)_L \times SU(2)_R$ alternative model, using Eq. 1.39, the limits on these LFV branching ratios, for the B^0_s and B^0_d mesons, can be translate into the two lower bounds on the LQ masses: $M_{LQ} > 3TeV$ and $M_{LQ} > 4TeV$. These limits are very interesting if compared to the prediction $M_{LQ} \sim TeV$ given by the author of [11].

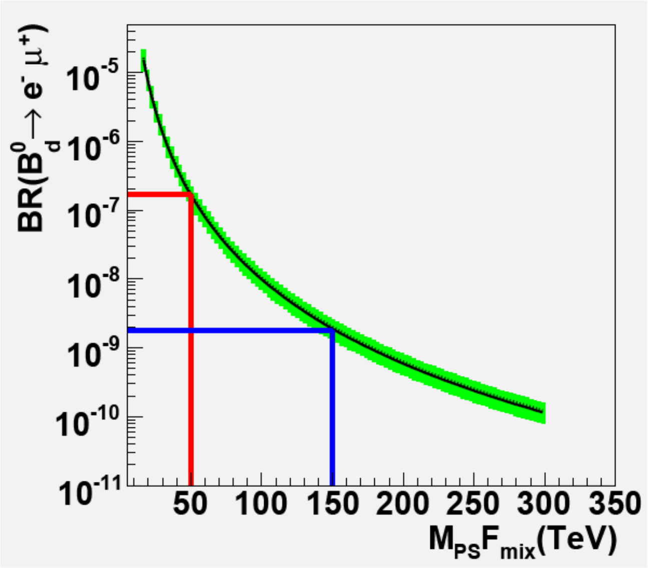


Figure 4.22: LQ mass $M_{LQ} {\cdot} F_{mix}$, according to the 1.20, versus the branching ratio of the $B^0_d \to e^{\pm}\mu^{\mp}$ decay. The present limit is indicated with the red line, the limit in one year of LHCb running at nominal luminosity is indicated with the blue line. The green band shows the effect of the effect of the error on $F_{B(d,s)}$.

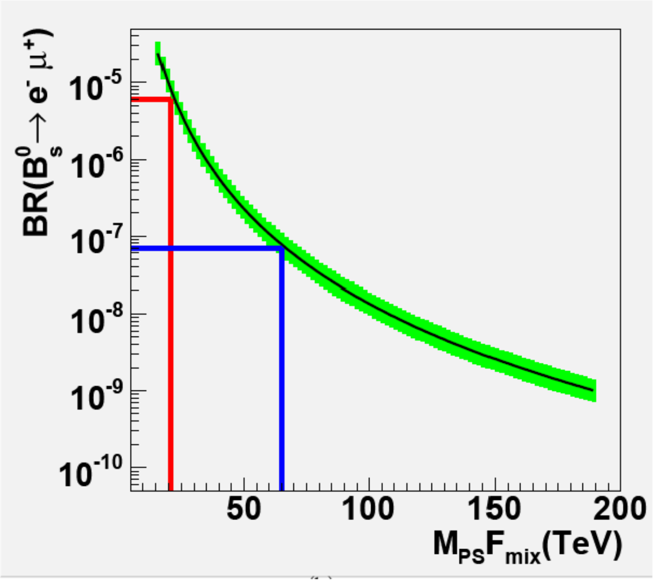


Figure 4.23: LQ mass $M_{LQ} \cdot F_{mix}$, according to the 1.20, versus the branching ratio of the $B^0_s \to e^\pm \mu^\mp$ decay. The present limit is indicated with the red line, the limit in one year of LHCb running at nominal luminosity is indicated with the blue line. The green band shows the effect of the effect of the error on $F_{B(d,s)}$.

Chapter 5

$B^0_{s,d} \to \mu^+\mu^-\gamma$

5.1 Signal and background

A study of the LHCb sensitivity to the $B^0_{s,d} \to \mu^+\mu^-\gamma$ channel is presented. The signal was generated with the software: Gauss v25r10, Boole v13r3 and Brunel v31r8. For simplicity only the sample $B^0_s \to \mu^+\mu^-\gamma$ was generated, assuming the same efficiency for the $B^0_d \to \mu^+\mu^-\gamma$ sample.

This decay was not present in the EvtGen package, moreover none of the models already present in the package were suitable to simulate the $B^0_s \to \mu^+\mu^-\gamma$ decay. Therefore we have generated the decay with the *psp* model.

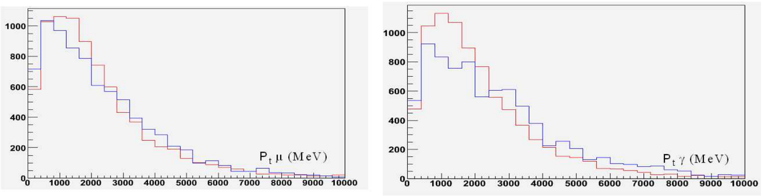
To estimate the systematic uncertainty of the decay model, we have compared the variable spectra obtained with the psp model and the spectra obtained with a software interfaced with Pythia and written by the authors of [42][1].

The comparison between the kinematical variable spectra obtained using the two different programs is shown in Figures 5.1(a)-5.1(d). As can be seen, the main difference between the two simulation is the dimuon invariant mass distribution and the P_t of the photon.

In first approximation the dimuon invariant mass is irrelevant for the sensitivity study, while for the photon P_t, we re-weight the events a posteriori taking into account the difference in the kinematical distributions between

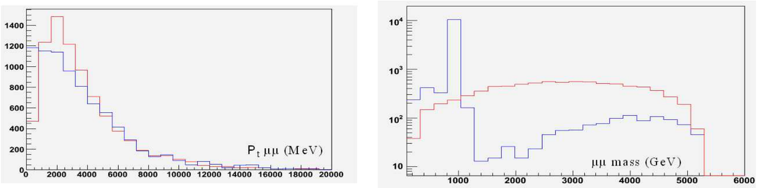
---

[1] This software was used for the computation of the branching ratio and the dimuon spectrum in [42]



(a) P_T of the muons. The blue line is computed according to the software used in [42], the red line is computed according to the psp model in Gauss.

(b) P_T of the photons. The blue line is computed according to the software used in [42], the red line is computed according to the psp model in Gauss



(c) P_T of the di-muons. The blue line is computed according to the software used in [42], the red line is computed according to the psp model in Gauss

(d) Invariant mass of the di-muons. The blue line is computed according to the software used in [42], the red line is computed according to the psp model in Gauss

Figure 5.1: Comparison between the kinematical variables, at generation level, for the simulation with the psp model of Gauss (red histograms) and for the simulation with the software of [42] (blue histograms) in the $B^0_s \to \gamma\mu^+\mu^-$ decay. Concerning Figure (d), the cuts on the dimuon invariant mass listed in Table 5.4, were applied.

Table 5.1: Number of analyzed, preselected and truth matched events for signal and background.

Channel	evts (comb)	Presel evts (comb)	Presel Truth Matched
$B_s^0 \to \mu^+\mu^-\gamma$	6877(16455)	3155(3509)	2761
$b\bar{b} \to \mu^+\mu^-$	14907042	26967(30235)	

the two programs.

This study was done with the *DC06* data, $b\bar{b} \to \mu^+\mu^-$ events were assumed to be the main source of background. This sample consists of $b\bar{b}$ events with at least two muons in the final state[2]. It was computed that this sample has a cross section 2.6% of the inclusive $b\bar{b}$ cross section. The background was generated with the software: Gauss v25r10, Boole v12r10 and Brunel v30r17. For both signal and background samples the package DaVinci v19r2 was used in the analysis. The samples analyzed are listed in Table 5.1.

5.2 Preselection

To reduce the rate on tape, a preselection was studied. This can be used at the HLT level as an exclusive selection.

The photon momentum was computed correcting the energy measurement in ECAL, assuming that it comes not from the PV, but from the dimuon intersection vertex.

Concerning photon identification, the *standard loose all photons* identification cut was used. This includes not only neutral clusters in the calorimeters which are identified as photon with the criteria of Section 2.9, but also electron pairs compatible with a photon conversion in the material. One of the difficulties of this analysis is the large number of photons present in the events. In fact there is often more than one candidate in the same event. In Figure 5.2 the P_t spectrum of the photons in signal events is shown. The red line represents truth matched signal events $B_s^0 \to \gamma\mu^+\mu^-$, the black line

---

[2] This sample doesn't consist only of muon directly coming from the two b-hadrons, as for the $b\bar{b} \to e^\pm\mu^\mp$. It is a sample extracted from the $b\bar{b}$ inclusive sample, with at least two muons in the final state.

Table 5.2: Preselection cut for the $B^0_s \to \gamma\mu^+\mu^-$ decay.

Cut Variable	Cut Value
$P_t(\mu)$	$> 500MeV$
$E_t(\gamma)$	$> 1GeV$
$Angle(\overrightarrow{P}_{B^0_s}, \overrightarrow{r}_{B^0_s} - \overrightarrow{r}_{PV})$	$< 14mrad$
$\mid M(\mu\mu\gamma) - M_{B^0_s} \mid$	$< 600MeV$

consists of photons coming from π^0 decays which come from the primary vertex, the blue line consists of photons coming from the other b-hadron. As can be seen, the P_t of the photon is a very sensitive variable to distinguish a signal photon from a background photon. The preselection cuts are listed in Table 5.2, with these cuts a rejection of about $3 \cdot 10^{-4}$ on the $b\bar{b}$ inclusive sample can be achieved. After the preselection, the number of combination per event is reduced to 1.1 for both signal and background samples. In Figures 5.3(a)-5.3(d) the preselection variables for signal and background are shown.

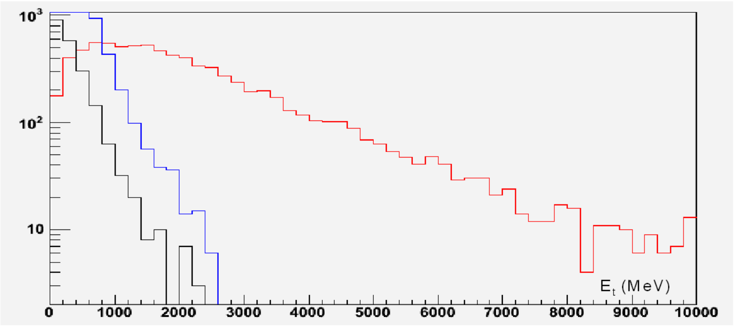
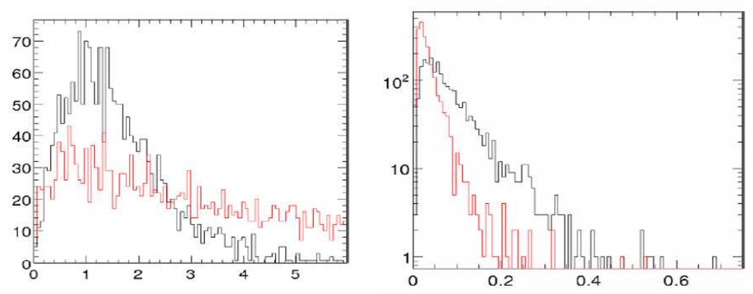


Figure 5.2: Transverse momentum of the γ in signal events with truth match (red line), from the other b-hadron (blue line) and from π^0 (black line).

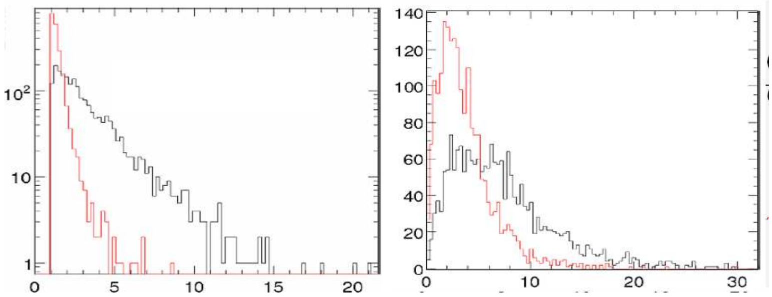
5.3 Multivariate analysis

This channel being a very rare decay, it is very important to have high efficiency with a high level of background rejection. Therefore this is the typical analysis in which the multivariate methods give better results with respect



(a) $\frac{IP}{\sigma_{IP}}$ of the B_s^0.

(b) Isolation of the photon.



(c) E_t of the photon.

(d) Transverse momentum of the B_s^0 ($P_t(B_s^0)$).

Figure 5.3: Distribution of some of the input variables used with multidimensional methods. The black histograms are signal events and the red histograms are background events.

to an analysis done with standard rectangular cuts. For this channel three multivariate methods were compared: the Fisher discriminant, the likelihood ratio and the decorrelated likelihood ratio. The following five input variables were used for signal selection.

-The smallest impact parameter significance of the B^0_s meson with respect to the primary vertexes in the event: $\frac{IP_{B^0_s}}{\sigma}$;

-The pointing angle of the B^0_s meson: $Angle(\overrightarrow{P}_{B^0_s}, \overrightarrow{r}_{PV} - \overrightarrow{r}_{B^0_s})$;

-The transverse momentum of the B^0_s meson : $P_t(B^0_s)$;

-The transverse momentum of the photon: $P_t(\gamma)$;

-The isolation of the photon:$\frac{P_t(\gamma)}{P_t(\gamma)+\sum_{\Delta R<2}(P_t)_i}$.

The sum in the isolation variable runs over the particles within a cone of $\Delta R < 2$ with respect to the photon flight direction. The selection variables for this channel are very similar to the variables used for the $B^0_s \to e^\pm\mu^\pm$ selection, except for $P_t(\gamma)$ and the definition of the Isolation.

This is due to the fact that the main source of background consists of the combinatorial background of photons. Signal events have in general a more isolated photon with a higher P_t with respect to the background.

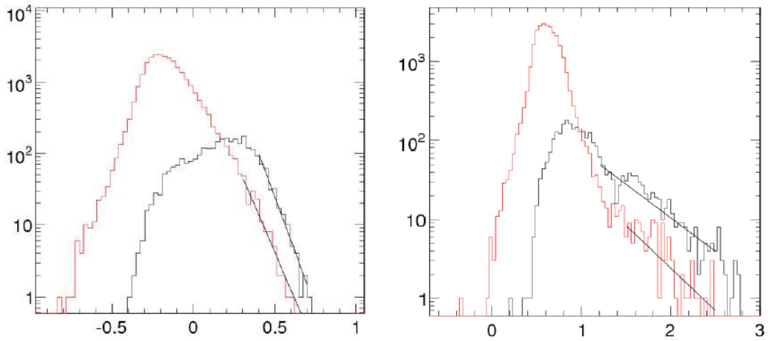
In Figures 5.4(a)-5.4(c) the outputs of the three multidimensional methods used are shown. The comparison of the efficiency as a function of the background events can be seen in Figure 5.5. This time the best method seems to be the likelihood ratio. In addition to the likelihood ratio cut, an invariant mass cut of $\mid Mass(\mu^+\mu^-\gamma) - Mass(B^0_s) \mid < 210 MeV/c^2$ was applied for the final selection. The invariant mass distributions for signal and background are shown in Figure 5.6. The red histogram is the signal with the MC truth match, the black histogram consists of signal events for which the photon doesn't come from the $B^0_s \to \gamma\mu^+\mu^-$ decay, the blue histogram is the $b\bar{b} \to \mu^+\mu^-$ invariant mass distribution [3].

5.4 Trigger

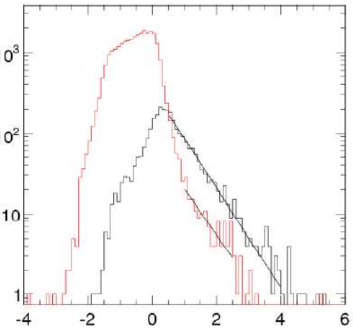
The first level of trigger (*L0* Trigger) was studied. The efficiency for signal events after the preselection is about 89%. In Table 5.3 the efficiency for

---

[3] The normalization between the signal and the background is arbitrary.



(a) Likelihood ratio (b) Decorrelated likelihood ratio



(c) Fisher discriminant

Figure 5.4: Distributions of the multidimensional output for the three different methods for signal (black histogram) and background (red histogram).

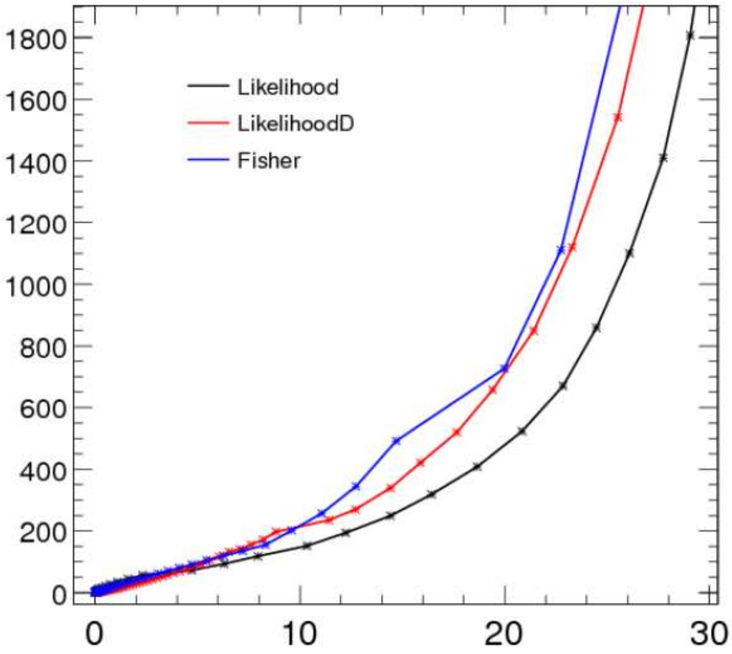


Figure 5.5: Signal efficiency versus the number of background events in 2fb^{-1} for the three multidimensional methods. The black line is computed with the likelihood method, the red line with the decorrelated likelihood method and the blue line with the Fisher discriminant method.

the various L0 trigger channels is listed. However, the *L0* dimuon trigger is under development and we would expect an increasing efficiency when this algorithm is ready. The HLT was not available at the moment of the study. However we expect a very high total efficiency close to 90% because this channel will be triggered predominantly by the double muon trigger.

5.5 Sensitivity to the $B^0_{s,d} \to \gamma\mu^+\mu^-$ decays

The LHCb sensitivity to this channel was studied in the hypothesis that the main background consists of $b\bar{b} \to \mu^+\mu^-$ events.

The behavior of the Punzi variable with respect to the efficiency of the signal for the three multivariate methods is shown in Figure 5.7. The background

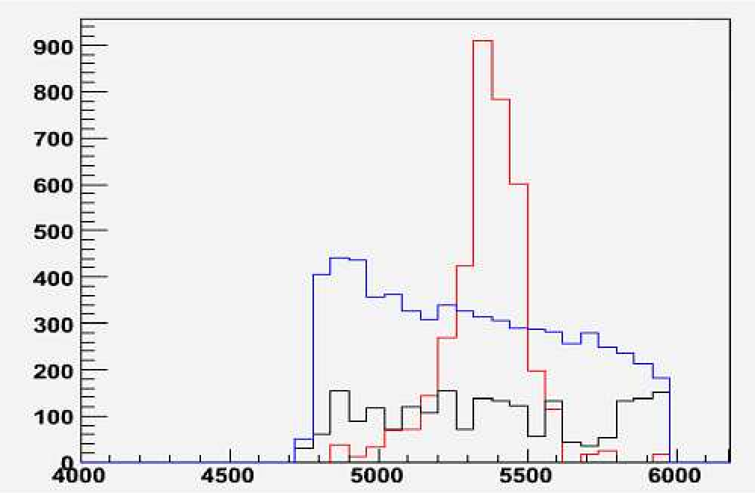


Figure 5.6: Invariant mass distribution for signal and background ($b\bar{b} \to \mu^+\mu^-$). The red histogram is the signal with truth match, the black line are signal events in which the photon doesn't come from the $B^0_s \to \gamma\mu^+\mu^-$ decay, the blue histogram is the background ($b\bar{b} \to \mu^+\mu$). Signal and background distributions are not normalized.

due to $B \to h_a^+ h_b^- + (\gamma)$ was considered negligible if compared to the larger background due to $b\bar{b} \to \mu^+\mu^-$ events. However, a cut on $DLL(\mu - \pi) > -10$ was applied to reject the background due to pions coming from the PV.
Since the resonance regions are affected by large theoretical uncertainties, these regions are not interesting for the search of NP. Therefore cuts on dimuon invariant mass to reject the long distance contribution of known resonances, were applied. These cuts are listed in Table 5.4.

Using the likelihood ratio method, the limits in one year of running of the LHCb experiment are: $Br(B^0_s \to \mu^+\mu^-\gamma) < 5.8 \cdot 10^{-7}$ and $Br(B^0_d \to \mu^+\mu^-\gamma) < 1.4 \cdot 10^{-7}$ at 90% CL.
Unfortunately, the LHCb experiment will not be able to measure the predicted SM branching ratio. Therefore the LHCb experiment will be sensitive only to large NP effects.
However, this measurement could be interesting for some NP models, in which operators which are not present in the $B^0_s \to \mu^+\mu^-$ decay, appear. The fact that the standard likelihood ratio has better performance with re-

Table 5.3: L0 Trigger efficiency for the signal. The efficiency for the various L0 channels are listed.

Trigger Channel	Accepted events
Total L0	88.46%
Electron L0	6.14%
Photon L0	16.63%
Hadron L0	13.12%
π^0 Local L0	11.49%
π^0 Global L0	17.68%
π^0	17.68%
Muon L0	60.49%
Sum Muon L0	85.83%

Table 5.4: The excluded dimuon invariant mass regions for the $B^0_s \to \gamma\mu^+\mu^-$ decay are listed.

excluded $\mu\mu$ invariant mass regions	$J/\psi(1S, 2S)$ mass rejection $2900 - 3200 MeV/c^2$ and $3650 - 3725 MeV/c^2$, $Mass(\mu^+\mu^-) < 1200 MeV/c^2$

spect to the decorrelated likelihood ratio has not yet been quite understood and will be subject of further investigation. However, even without other improvements, the limit on the B^0_d, in just one year of data taking, is equal to the world best result already set by BABAR ([48]). Moreover, even if the limit for the $Br(B^0_s \to \mu^+\mu^-\gamma)$ is of the same order as the present limit on the B^0_d of [48], the $B^0_s \to \gamma\mu^+\mu^-$ decay is much more interesting concerning the NP searches. In fact by looking at Eq. 1.43, it can be seen that both SM and NP operators are multiplied by the factor $V_{tq}V^*_{tb}$. Whichever operators in Eq. 1.43 contribute to the B^0_s and B^0_d decays, we would expect to have $\frac{Br(B^0_d \to \mu^+\mu^-\gamma)}{Br(B^0_s \to \mu^+\mu^-\gamma)} \sim \frac{1}{20}$ because of Cabibbo suppression. Hence, the sensitivity to NP operators, which can be achieved in the LHCb experiment with this channel, is 20 times higher than the sensitivity achievable with the same limit on the $B^0_d \to \mu^+\mu^-\gamma$ [4] decay.

---

[4] A Similar argument is not valid for the LFV decays $B^0_{s,d} \to e^\pm\mu^\mp$ in context of Pati-Salam like models. This is due to the fact that these decays can be represented by tree level diagrams.

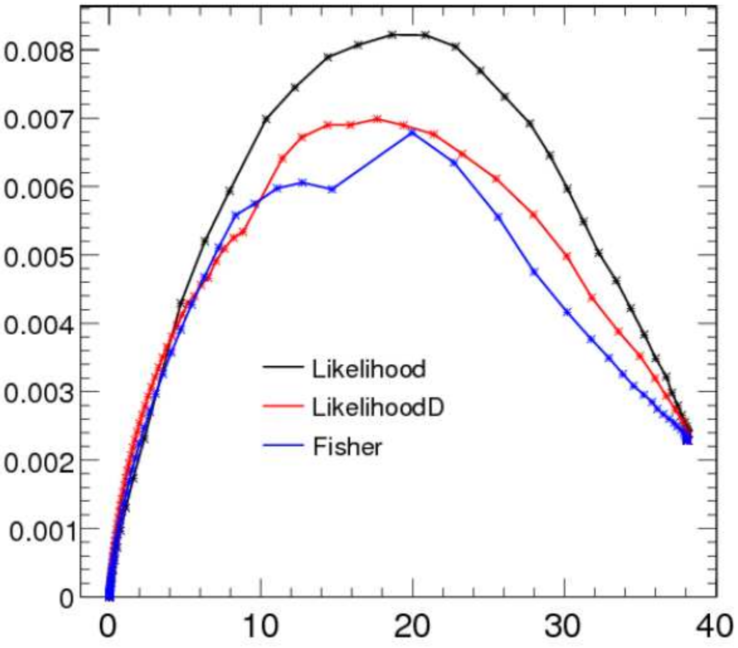


Figure 5.7: Punzi variable as a function of the signal efficiency for the different muldimensional methods. The maximum is the point of maximum sensitivity.

Chapter 6

Angular distributions in $B_d^0 \to K^{*0}(892)\mu^+\mu^-$

6.1 Observables in the $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ channel

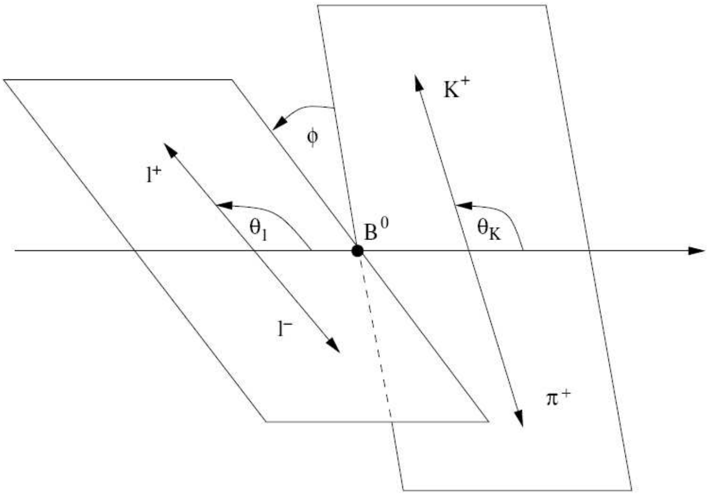


Figure 6.1: Kinematic variables for the $B_d^0 \to K^{*0}\mu^+\mu^-$ decay.

In Section 1.7 it was explained that the asymmetries $A_T^{(1),(2)}$ and the FBA in the $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ decay are predictable with a very small theoretical error within the SM and are sensitive to NP, in particular to right-handed currents. However, it is not possible to measure directly these asymmetries except for the FBA. Projecting the angular distribution onto the three angles (θ_K, θ_l, ϕ) it is possible to obtain the following distributions [60]:

$$\begin{aligned}
\frac{d\Gamma}{d\phi} &= \frac{\Gamma}{2\pi}\left(1+\frac{1}{2}(1-F_L)A_T^{(2)}\cos 2\phi + A_{Im}\sin 2\phi\right) && (6.1)\\
\frac{d\Gamma}{d\theta_l} &= \Gamma(\frac{3}{4}F_L\sin^2\theta_l + \frac{3}{8}(1-F_L)(1+\cos^2\theta_l) + && (6.2)\\
&+ A_{FB}\cos\theta_l)\sin\theta_l &&\\
\frac{d\Gamma}{d\theta_K} &= \frac{3\Gamma}{4}\sin\theta_K\left(2F_L\cos^2\theta_K + (1-F_L)\sin^2\theta_K\right) && (6.3)
\end{aligned}$$

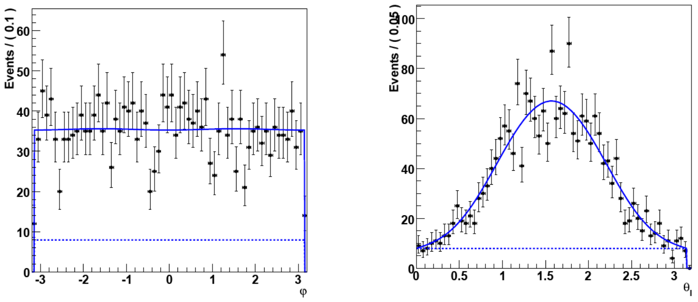
The LHCb sensitivity to the FBA in the $B_d^0 \to K^{*0}\mu^+\mu^-$ channel was studied in detail in [54] with a *counting experiment*. The acceptance and background effects were taken into account. To measure the $A_T^{(2)}$ asymmetry it is necessary to fit the distributions 6.1 and 6.3.
In Figure 6.2 the possible measurements in 2fb^{-1} of integrated luminosity of the LHCb experiment, simulated with the software RooFit, are shown. The theoretical distributions are superimposed (blue lines). The sensitivity to the observables $A_T^{(2)}$, F_L and A_{FB} in 2fb^{-1} of integrated luminosity of the LHCb experiment are shown in Table 6.1, for different s regions.

These results were obtained in [60] with *RooFit* under the following conditions:
-The distributions are given in the range 1GeV$^2 < s <$ 6GeV^2, which is the range in which the theoretical predictions are valid;
-Resolutions and acceptance effects were neglected;
-The background distributions are considered flat in all the observables.
The results are listed in Table 6.1 and Table 6.2[1].

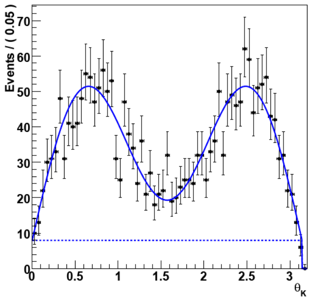
The resolution in the A_{FB} and F_L measurement is very good. However, even if the parameter F_L is in principle sensitive to NP effects, these effects are comparable to the theoretical uncertainties (Figure 6.3). The resolution on the $A_T^{(2)}$ parameter is not good. This is due to the fact that the $(1-F_L)$ function has a minimum in the middle region of the variable s ($1.00 < s < 6.00$) which is the most interesting for the NP measurement and that there is correlation between the two parameters in the fit (see equation 6.1). However

---

[1] This tables are taken from [60]



(a) ϕ distribution. (b) θ_l distribution.



(c) θ_K distribution.

Figure 6.2: Possible measurements in 2fb^{-1} of integrated luminosity at LHCb for the angles ϕ (a), θ_l (b) and θ_K (c). The theoretical distributions are also shown (blue lines). The expected background is represented by the dashed lines.

Table 6.1: Sensitivity to $A_T^{(2)}$, F_L and A_{FB} in $2fb^{-1}$. The results are given for three regions of s, being s the dimuon invariant mass square.

s region (GeV$^2/c^4$)	$A_T^{(2)}$	A_{FB}	F_L
0.05-1.00	0.14	0.034	0.027
1.00-6.00	0.42	0.020	0.016
6.00-8.95	0.28	0.022	0.017

Table 6.2: Sensitivity to $A_T^{(2)}$, F_L and A_{FB} in $10fb^{-1}$. The results are given for three regions of s, being s the dimuon invariant mass square.

s region (GeV$^2/c^4$)	$A_T^{(2)}$	A_{FB}	F_L
0.05-1.00	0.007	0.017	0.011
1.00-6.00	0.16	0.008	0.007
6.00-8.95	0.13	0.010	0.008

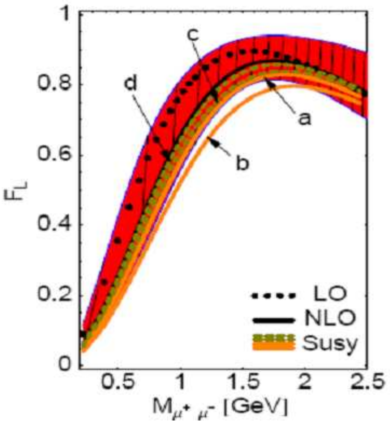


Figure 6.3: SM prediction for the observable F_L as a function of the dimuon mass. The leading order and next-to-leading order prediction are shown. The band represents the theoretical uncertainty. Predictions of some SUSY models are also shown.

this precision improves with the statistics.

It is not possible to measure the $A_T^{(1)}$ asymmetry with a projection study. For this measurement a full angular fit is necessary, which has not been yet implemented. The study of [60] was done without taking into account acceptance and resolution effects. Concerning the resolution effects, these are negligible as long as the variables are slowly varying with respect to the experimental resolution. Because the angular resolution is quite good this is actually the case. However, biases coming from the detector acceptance can no longer be neglected in this kind of measurement.

In this thesis work the possibility to correct the angular biases in the $B_d^0 \to K^{*0}\mu^+\mu^-$ channel with the control channel $B_d^0 \to J/\psi K^{*0}$ was studied.

6.2 Signal selection

The $K^{*0}(892)$ is an unstable particle which decays with a branching ratio of almost 100% in $K\pi$, with an amplitude of $\Gamma = 50.3 \pm 0.6 MeV$. For the strangeness conservation, the $K^{*0}(892)$ decays with a K^+ in the final state and the $\overline{K}^{*0}(892)$ has a K^- in the final state. Therefore this decay is flavor specific, i.e. B_d^0 and $\overline{B}_d^0$ can be recognized by the decay products.

Table 6.3: Selection and preselection cuts for the $B_d^0 \to K^{*0}\mu^+\mu^-$ channel.

Cut Variable	Preselection cut value	Selection cut value
B_d^0 mass window	$\pm 500 MeV/c^2$	$\pm 50 MeV/c^2$
B_d^0 P_t	$> 250 MeV/c$	$> 250 MeV/c$
B_d^0 vertex χ^2	< 50	< 20
B_d^0 Pointing angle	$< 141 mrad$	$< 22 mrad$
B_d^0 FS	$-$	> 6
B_d^0 sIPS	< 8	< 5
B_d^0 Vertex Isolation	$-$	11
$K^{*0}(892)$ mass window	$\pm 300 MeV/c^2$	$\pm 100 MeV/c^2$
$K^{*0}(892)$ P_t	$> 300 MeV/c$	$> 300 MeV/c$
$K^{*0}(892)$ vertex χ^2	< 30	< 30
$K^{*0}(892)$ FS	$-$	> 1
$K^{*0}(892)$ sIPS	> 1.5	> 1.5
$K^\pm$ P	$> 2000 MeV/c$	$2000 MeV/c$
$K^\pm$ P_t	$250 MeV/c$	$400 MeV/c$
$K^\pm$ sIPS	> 1.5	> 3
$\mu\mu$ excluded mass window	$-$	$J/\psi(1S, 2S)$ mass rejection $2900 - 3200 MeV/c^2$ and $3650 - 3725 MeV/c^2$
$\mu\mu$ vertex χ^2	$-$	< 15
$\mu\mu$ flight-distance	$-$	$> 1mm$
$\mu^\pm$ P	$> 4000 MeV/c$	$4000 MeV/c$
$\mu^\pm$ P_t	$300 MeV/c$	$500 MeV/c$
$\mu^\pm$ sIPS	> 0.5	> 2.0

The selection and preselection cuts were studied in [55], these cuts are listed in Table 6.3 [2]. With these cuts a $\frac{B}{S} = 0.5 \pm 0.2$ after the Trigger can be achieved, with a signal efficiency of $1.11 \pm 0.03\%$.

For this kind of analysis it is not appropriate to use multivariate methods for signal selection, for two main reasons:

-The signal rate is sufficiently high and it is less important with respect to very rare decays to have high efficiency;

-We want to have under control all the possible sources of bias due to signal selection.

---

[2] The vertex Isolation variable is defined as the number of particles which have a IPS< 3 with respect to the B_d^0 decay vertex.

6.3 Choice of the control channel

If we neglect resolution effects, the differential decay rate of Equation 1.44 can be written as:

$$I_{measured}(\theta_K, \theta_l, \phi, s) = I_{true}(\theta_K, \theta_l, \phi, s) \times A_{det}(\theta_K, \theta_l, \phi, s) \tag{6.4}$$

where $I_{measured}$ is the 4-dimensional measured distribution of the observables, I_{true} is the true distribution of the observables and A_{det} is the detector acceptance. The function A_{det} depends not only on detector efficiency effects but also on selection cuts.

Obviously knowing the function A_{det} it is possible to obtain the function I_{true} in which we are interested in, from the function $I_{measured}$.

In principle it would be possible to measure the function A_{det} if we had an *ideal control channel*. An ideal control channel has the following characteristics:

-It is well known, i.e. it is a channel for which we know the observable distributions with sufficient precision;

-It has very high statistics;

-It has the same final state as the signal;

-It has the same observable spectra as the signal;

-It is triggered by the same channels as the signal.

When there is no ideal control channel the function A_{det} cannot be measured directly. In this work the use of the control channel $B_d^0 \to J/\psi K^{*0}(892)$ is proposed. This channel is very well known, its angular distributions were measured by the *BABAR* experiment in a very clean environment [61]. It has sufficiently high statistics, its branching ratio is $Br(B_d^0 \to J/\psi K^{*0}) = (1.33 \pm 0.06) \cdot 10^{-3}$ and $Br(J/\psi \to \mu^+\mu^-) = (5.96 \pm 0.06) \cdot 10^{-2}$.

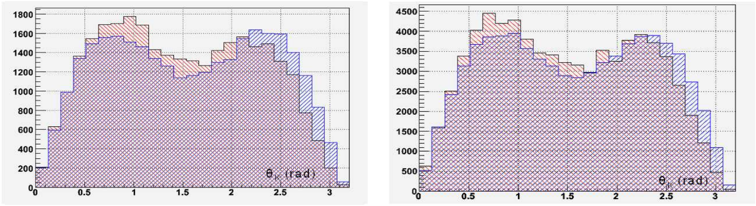
Moreover signal and the control channel have the same final state and both of them are triggered by the same channel (the dimuon trigger).

The only problem is that for the control channel the s variable is almost a Dirac δ function peaked around the J/ψ invariant mass. The kinematical spectrum of the muon pair is, therefore, completely different in the two channels and for this reason it is hard to use this channel for the θ_l recovering (see below).

For this study full simulated data of the *DC06* data challenge for both signal and control channel were used. For the MC simulation and the data analysis the following packages were used: Gauss v25r12, Boole v12r10, Brunel v30r17 and DaVinci v17r8. For simplicity the background is neglected, assuming for it a flat distribution in all the observables.

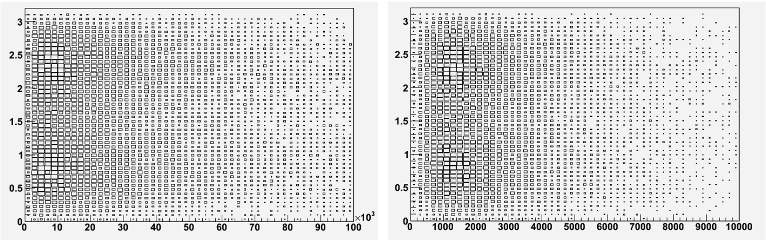
6.4 θ_K recovering

In Figures 6.4(a)-6.9(c) the comparison between the true MC θ_K distribution (blue histogram), and the θ_K distribution for preselected events with truth match with respect to the MC truth (red histogram), is shown. The total acceptance effect is similar for both signal and control channel. In Figures 6.5(a)-6.5(b) the scatter plots of the θ_K angle versus the muon momentum and the muon transverse momentum are shown. It can be seen that the θ_K angle is quite independent on the muon spectrum. This allows us to disen-



(a) $B_d^0 \to J/\psi K^{*0}(892)$ (b) $B_d^0 \to K^{*0}(892)\mu^+\mu^-$

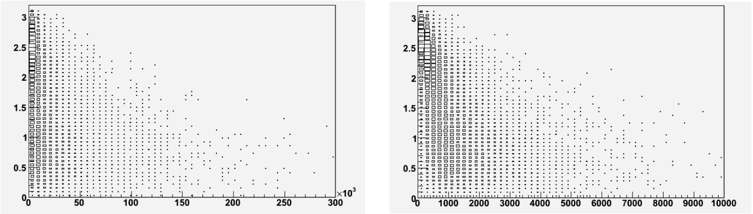
Figure 6.4: θ_K distribution for the reconstructed preselected events (red histogram) and for the MC truth (blue histogram).



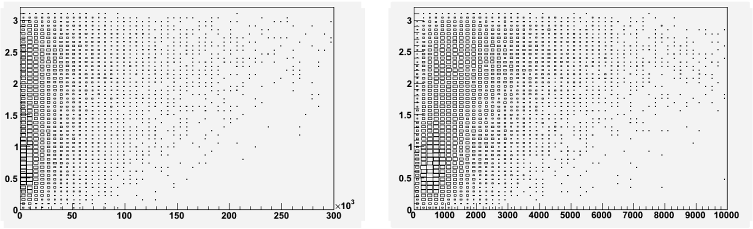
(a) θ_K VS $P(\mu^{\pm})$ (b) θ_K VS $P_t(\mu^{\pm})$

Figure 6.5: θ_K distribution versus the P and P_t for the muons.

tangle the θ_K recovering procedure from the muon kinematical variables. However the θ_K angle distribution is strongly correlated with the final hadron momentum (Figures 6.6(a)-6.6(d)). In the following, with the term "measured distribution", we indicate the variable distribution of the preselected events with truth match. The choice to take the truth matched is justified by the fact that we would like to separate the effect of fake combinations and the effect due to the detector acceptance. Here two methods for the θ_K recovering, using the control channel $B_d^0 \to J/\psi K^{*0}$, are presented [7]. These method are the correction using P or P_t information and the correction based on the direct θ_K recovering.



(a) θ_K VS $P(\pi^\pm)$ (b) θ_K VS $P_t(\pi^\pm)$



(c) θ_K VS $P(K^\pm)$ (d) θ_K VS $P_t(K^\pm)$

Figure 6.6: θ_K distribution versus the P and P_t of the hadrons $\pi^\pm$ and $K^\pm$.

6.4.1 θ_K correction using P or P_t information

The detector acceptance must depend in some way on the kinematical variables (P, P_t) of the particles in the final state. Here an event-by-event effective correction is proposed. We assume that the detector acceptance depends only on the transverse momentum of the particles in the final state.

Table 6.4: χ^2 of the difference between the measured distribution and the MC truth distribution. The χ^2 is computed with respect to the straight line $\Delta\theta_K = 0$.

Channel	χ^2	χ^2 per degree of freedom	Method
$B_d^0 \to J/\psi K^{*0}(892)$	351	14.1	No recovering
$B_d^0 \to J/\psi K^{*0}(892)$	33.1	1.32	Using P_t
$B_d^0 \to K^{*0}(892)\mu^+\mu^-$	348	13.9	No recovering
$B_d^0 \to K^{*0}(892)\mu^+\mu^-$	35.2	1.40	Using P_t
$B_d^0 \to K^{*0}(892)\mu^+\mu^-$	22.0	$8.80 \cdot 10^{-1}$	direct

In this case we have:

$$\epsilon_{tot} = \epsilon_{tot}(P_t(\mu^+), P_t(\mu^-), P_t(\pi^\pm), P_t(K^\pm)). \tag{6.5}$$

Moreover, assuming that the efficiency as a function of P_t for the different particles is not correlated, 6.5 becomes:

$$\epsilon_{tot}(P_t) = \epsilon_{\mu^+}(P_t) \times \epsilon_{\mu^-}(P_t) \times \epsilon_\pi(P_t) \times \epsilon_K(P_t) \tag{6.6}$$

These hypotheses can be verified a posteriori. Because the θ_K distribution is uncorrelated with the muon spectrum (Figures 6.5(a)-6.5(b)) we can assume that $\epsilon_{\mu^\pm}(P_t) = cost.$ We can extract $\epsilon_{K(\pi)}(P_t)$ from the control channel $B_d^0 \to J/\psi K^{*0}(892)$. Than the signal events can be weighted by the function:

$$W = \{\epsilon_\pi(P_t) \cdot \epsilon_K(P_t)\}^{-1}. \tag{6.7}$$

In Figure 6.7 the ratio of the kaon P_t distribution for the measured particles and for the MC truth is shown. This plot is by definition the efficiency as a function of P_t, except for a scaling factor. The distribution in Figure 6.7 is fitted with the function of Eq. 6.8 for pions and kaons:

$$\epsilon_{\pi(K)}(P_t) = 1 - e^{-\alpha(\beta)\frac{P_t}{GeV}}. \tag{6.8}$$

The coefficients α and β (for the $\pi^\pm$ and $K^\pm$ respectively) are extracted from the channel $B_d^0 \to J/\psi K^{*0}(892)$. In Figures 6.8(a)-6.8(d) the P_t distribution for the signal, before and after the recovering procedure, compared with the MC truth distribution is shown. The efficiency extracted from the control channel can be therefore used to recover the P_t distribution of the signal.

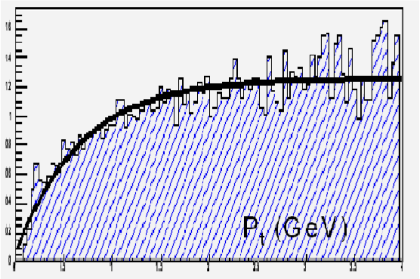


Figure 6.7: Efficiency as a function of P_t. The function used for the fit (Eq. 6.8) is also shown.

The main result of this procedure is that it allows θ_K recovering. This is shown in Figures 6.9(b)-6.9(d). These Figures show the comparison between the MC truth distribution and the measured distribution before (Figure 6.9(c)) and after (Figure 6.9(d)) the recovering procedure.

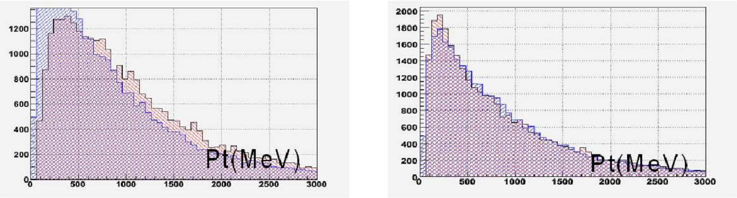
In Figure 6.10(b), the absolute difference between the MC truth distribution and the measured distribution, is shown. The red points are the experimental data before the recovering procedure, the blue points are obtained with the recovering procedure, the straight line represents the MC distribution.

To measure the goodness of the recovering procedure we can compute the χ^2 of the experimental points with respect to the theoretical distribution (the straight black line in 6.10(a)-6.10(c)) of Figure 6.10(b). The result is shown in Table 6.4. For both signal and control channel, the χ^2 per degree of freedom, after the recovering procedure, is significantly smaller than before and it is close to one.

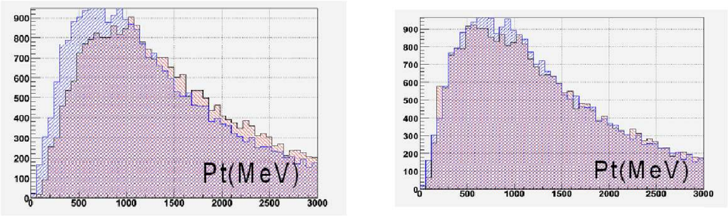
Unfortunately, in the real data we cannot directly extract $\epsilon_{\pi(K)}(P_t)$ by looking at the P_t distributions of the control channel. In fact the P_t distribution is affected by large uncertainties due to non perturbative QCD effects.

However, the two parameters α and β, needed for the computation of $\epsilon_{tot}(P_t)$, can be extracted minimizing the χ^2 of the difference between the θ_K distribution in the data and in the MC for the control channel. The θ_K distribution is in fact very well known for the control channel [61] and it can be directly taken from *BABAR* measurements.

This procedure is less efficient in the tails of the θ_K distribution. This is due

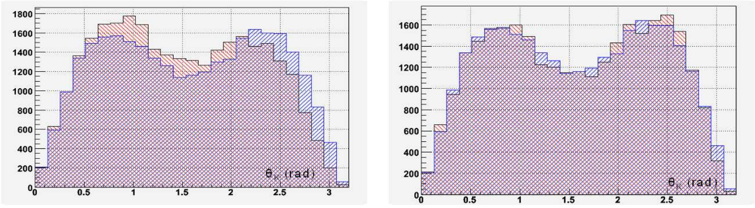


(a) $\pi^{\pm}$ before recovering (b) $\pi^{\pm}$ after recovering

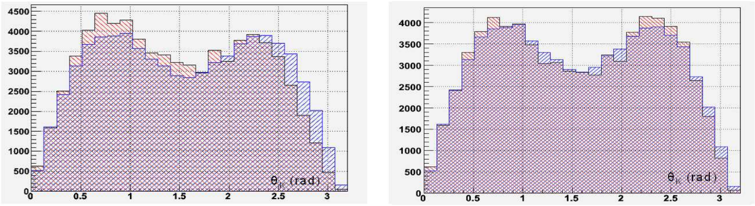


(c) $K^{\pm}$ before recovering (d) $K^{\pm}$ after recovering

Figure 6.8: P_t distribution before and after the recovering procedure for the signal. The $\epsilon_{\pi,K}(P_t)$ used for the recovering procedure was extracted from the control channel. The red histograms are full simulated reconstructed and preselected events, the blue histograms are MC truth events.



(a) θ_K $B_d^0 \to J/\psi K^{*0}(892)$ before recovering (b) θ_K $B_d^0 \to J/\psi K^{*0}(892)$ after recovering



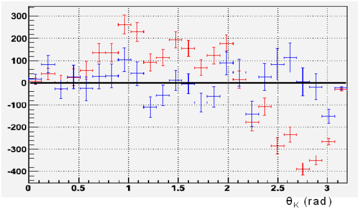
(c) θ_K $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ before recovering (d) θ_K $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ after recovering

Figure 6.9: θ_K distribution before and after the event-by-event recovering procedure for both signal and control channel. The red histograms are full simulated reconstructed and preselected events, the blue histograms are MC truth events.

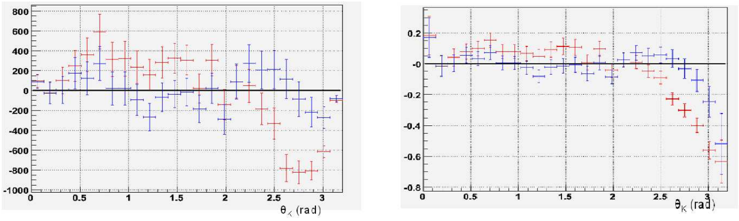
to the fact that the tails are more sensitive to the low P_t distributions of the particles and it is impossible to recover the events removed from the P_t cuts. Because many preselections have P_t cuts we have also tried to recover the θ_K distribution using the P information instead of the P_t. However the efficiency function for the single particle is not dependent on a single parameter any more. The efficiency $\epsilon_{\pi(K)}(P)$ was fitted with the function of Eq. 6.9:

$$\epsilon_{\pi(k)}(P) = e^{-\gamma(\delta)\frac{P}{GeV}}\{1 - e^{-\alpha(\beta)\frac{P}{GeV}}\}. \tag{6.9}$$

The signal events can be weighted by the function $W = \{\epsilon_\pi(P) \cdot \epsilon_K(P)\}^{-1}$. Using this method it is still possible to recover the θ_K distribution (Figures 6.11(a)-6.11(b)) but the result is worse than the previous result obtained using the P_t information. This is due to the fact that the information of the P_t is now diluted by the boost. Moreover there are twice as many parameters as before and it could be difficult to extract the four needed parameters from the θ_K distribution of the control channel. In conclusion, weighting the signal events with $\epsilon_{tot}(P_t)$, it is possible to recover the θ_K distribution. In this case



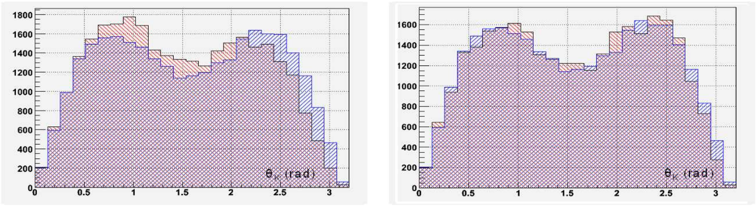
(a) Absolute difference between the measured distribution and the MC θ_K distribution for the $B_d^0 \to J/\psi K^{*0}(892)$ channel. The blue points are computed with the recovering procedure, the red points are the original data.



(b) Absolute difference between the measured distribution and the MC θ_K distribution for the $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ channel. The blue points are computed with the recovering procedure, the red points are the original data.

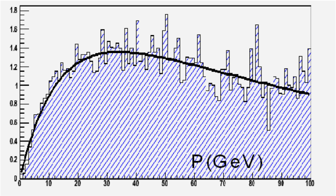
(c) Relative difference between the measured distribution and the MC θ_K distribution for the $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ channel. The blue points are computed with the recovering procedure, the red points are the original data.

Figure 6.10:



(a) θ_K distribution of the $B_d^0 \to J/\psi K^{*0}(892)$ channel before recovering. The red histogram consists of full simulated reconstructed and preselected events, the blue histogram consists of MC truth events.

(b) θ_K distribution of the $B_d^0 \to J/\psi K^{*0}(892)$ channel after recovering. The red histogram consists of full simulated reconstructed and preselected events, the blue histogram consists of MC truth events.



(c) Efficiency as a function of P for kaons in the final state in the control channel.

Figure 6.11:

it is very reasonable to assume that $\epsilon_{\pi(K)}(P_t)$ is the same for both signal and control channel. In fact we expect that the efficiency as a function of P_t doesn't depend on the particular channel.

6.4.2 Correction based on the direct θ_K recovering

Assuming that the efficiency as a function of θ_K is the same for both signal and control channel, we can directly recover the θ_K distribution. This can be done by measuring the efficiency $\epsilon(\theta_K)$ in the control channel and weighting the signal events with the function $\epsilon^{-1}(\theta_K)$. In Figure 6.12 the weighting function, extracted from the comparison between the true θ_K distribution and the measured θ_K distribution in the control channel, is shown. This is fitted with the function in Eq. 6.10.

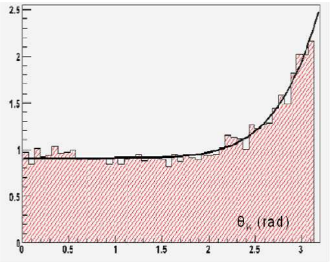
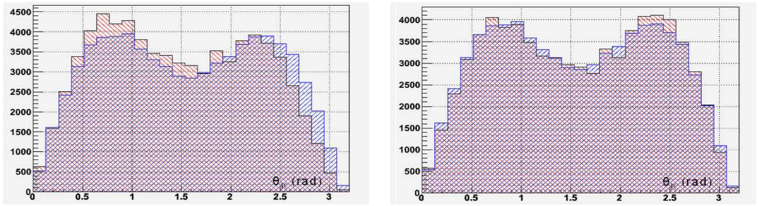


Figure 6.12: Weighting function for direct θ_K recovering. The $W = \epsilon^{-1}(\theta_K)$ function of 6.10 is also shown.

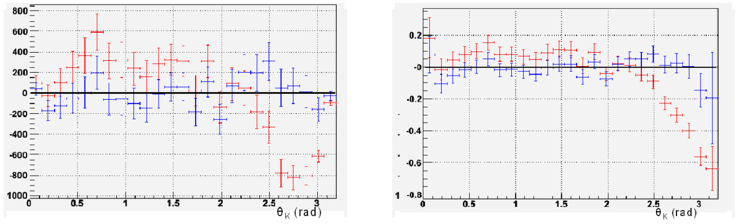
$$W = \epsilon^{-1}(\theta_K) = \alpha(x^{\beta} + \gamma) + \delta. \tag{6.10}$$

The recovered θ_K distribution is shown in Figure 6.13(b). The absolute difference between the measured θ_K distribution and the MC truth θ_K distribution before (red points) and after the recovering procedure (blue points) can be seen in Figure 6.13(c). The χ^2 of the experimental data with respect to the theoretical curve are listed in Table 6.4 for the original and recovered data. As can be seen, the correction based on the direct θ_K recovering gives the best result, however this procedure is based on the hypothesis that the $\epsilon(\theta_K)$ function is the same for both signal and control channel. It is difficult to verify this hypothesis in real data. Moreover the θ_K observable is not a *good* variable to express the detector efficiency. For this reason the other method was studied.



(a) θ_K $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ before recovering. The red histogram consists of full simulated reconstructed and preselected events, the blue histogram consists of MC truth events.

(b) θ_K $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ after recovering. The red histogram consists of full simulated reconstructed and preselected events, the blue histogram consists of MC truth events.



(c) Absolute difference between the measured θ_K distribution and the θ_K theoretical distribution. The red points are computed before the recovering procedure, the blue points after the recovering procedure.

(d) Relative difference between the measured θ_K distribution and the θ_K theoretical distribution. The red points are computed before the recovering procedure, the blue points after the recovering procedure.

Figure 6.13:

6.5 ϕ distribution

The ϕ angle is the angle between the planes of the z-axis and the $K^{*0}(892)$ decay and the muon pair in the B-meson rest frame (6.1). Because of the symmetry of the system it is hard to think of a detector effect which can distort this distribution. Therefore we would expect a flat ϕ acceptance function. In Figure 6.14, the MC truth ϕ distribution (blue histogram) and the measured ϕ distribution after a very strong P_t artificial distortion for the particle in the final state, are shown. Moreover, the ϕ distribution is completely independent on all the kinematical variables of the particles in the final state (Figures 6.15(a)-6.15(f)).

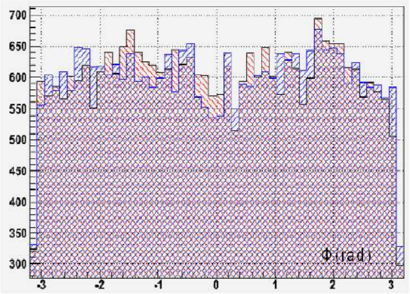


Figure 6.14: ϕ distribution for the MC truth (blue histogram) and after a very strong artificial distortion of the P_t of particles in the final state (red histogram).

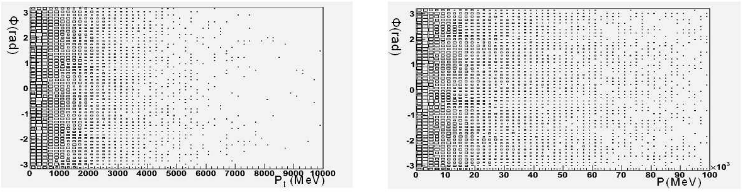
6.6 θ_l recovering

6.6.1 Effect of selection cuts

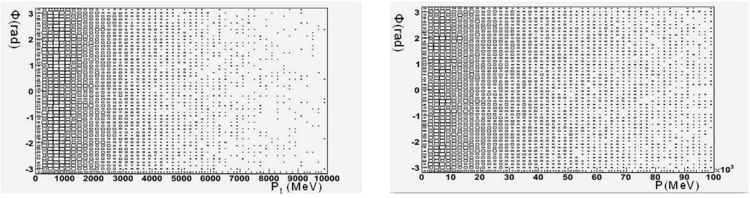
Among the preselection and selection cuts there are cuts that do not give a θ_l flat acceptance function. These effect is predominantly due to muon P_t cuts. In Figures 6.16(a)-6.16(d), it can be seen the θ_l efficiency in different dimuon invariant mass bins, when a cut of $P_t > 300MeV$ is applied. However, also $sIPS$ cuts have a small effect on the shape of θ_l acceptance function (Figure 6.17(a)).

The other cuts, like cuts on the vertex χ^2 and on the FS, are not correlated to θ_l angle (see for instance Figure 6.17(b)).

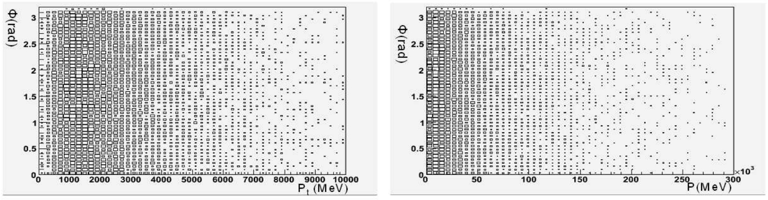
Even if it is better to avoid cuts that give a non flat acceptance function with respect to the angles, it is difficult to avoid P_t cuts which are also present at the Trigger level. However, this channel will be triggered predominantly by the dimuon Trigger. It can be seen from Figures 6.18(a)-6.18(b) that the cut on the dimuon P_t is not correlated with the θ_l variable. The dimuon P_t is also not correlated with the other angles (Figures 6.19(a)-6.19(c)).



(a) ϕ Vs $P_t(\pi^{\pm})$ (b) ϕ Vs $P(\pi^{\pm})$

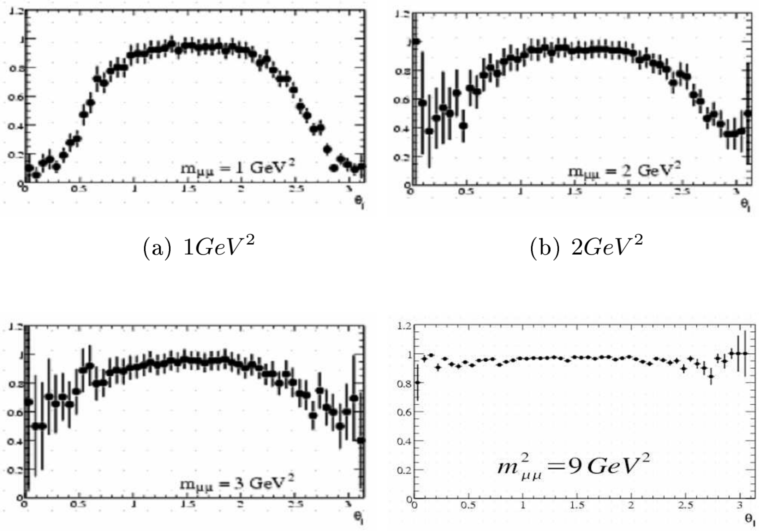


(c) ϕ Vs $P_t(K^{\pm})$ (d) ϕ Vs $P(K^{\pm})$



(e) ϕ Vs $P_t(\mu^{\pm})$ (f) ϕ Vs $P(\mu^{\pm})$

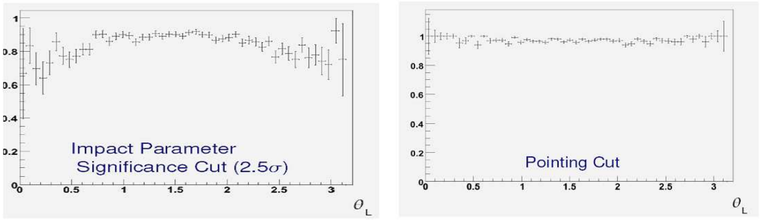
Figure 6.15: correlation of the ϕ angle with the P and P_t of the particles in the final state.



(a) $1GeV^2$ (b) $2GeV^2$

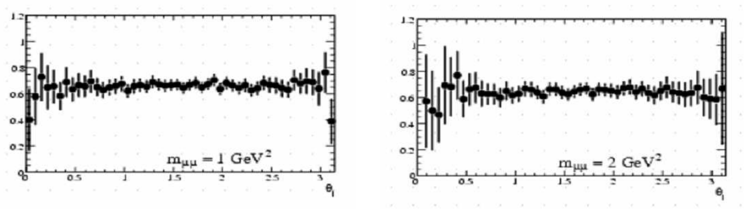
(c) $3GeV^2$ (d) $9GeV^2$

Figure 6.16: θ_l acceptance function for various muon P_t cuts. Note that the acceptance function is flat for high dimuon mass.



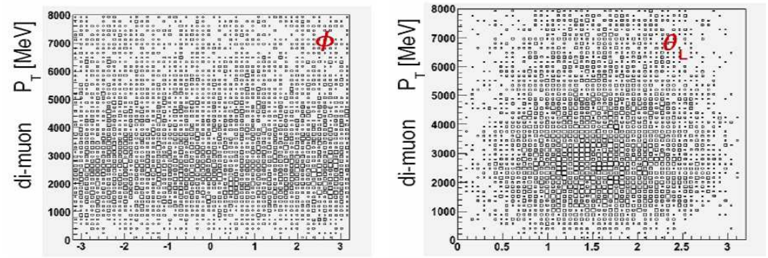
(a) low s (b) high s

Figure 6.17: θ_l acceptance when a $sIPS$ cut ($sIPS > 3\sigma$) on the muons is applied for a low dimuon mass a) ($s < 4GeV^2/c^4$) and for high dimuon mass b) ($s > 4GeV^2/c^4$).

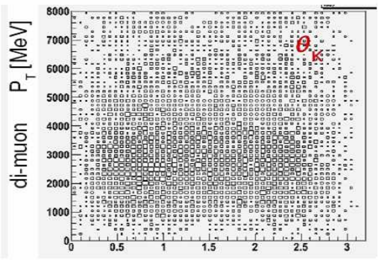


(a) $s = 1GeV^2$ (b) $s = 2GeV^2$

Figure 6.18: θ_l acceptance for two values of the dimuon mass, with a dimuon P_t cut at $1GeV$.



(a) Dimuon P_t VS ϕ. (b) Dimuon P_t VS θ_l

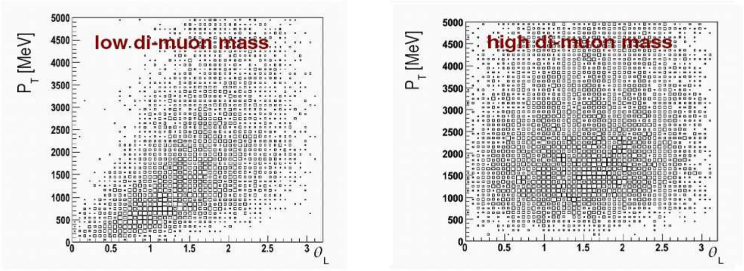


(c) Dimuon P_t VS θ_K

Figure 6.19: Dimuon P_t versus the ϕ, θ_l and θ_K angles.

6.6.2 Problems in θ_l recovering

The θ_l distribution is very strongly correlated to the muon transverse momentum for low dimuon invariant masses (Figure 6.20(a)) but it is almost independent of it for high dimuon invariant masses.
Since the efficiency with respect to θ_l is flat for high dimuon invariant masses, in the region of the control channel (the J/ψ invariant mass is about $3.1 GeV/c^2$), we expect a flat θ_l acceptance function for the $B_d^0 \rightarrow J/\psi K^*$ decay. Therefore we cannot hope to recover the θ_l distribution with this control channel.
However, the θ_l recovering was studied with the more promising control channel $B_d^0 \rightarrow K^+\mu^+\mu^-$ for which the dimuon invariant mass spectrum is the same as the signal and for which no FBA is expected.
This study is not part of this thesis work.



(a) low s (b) high s

Figure 6.20: Correlation of the θ_l angle with the muon P_t in the $B_s^0 \rightarrow K^{*0}\mu^+\mu^-$ decay for low dimuon mass ($s < 4GeV^2/c^4$) and high dimuon mass ($s > 4GeV^2/c^4$).

6.7 Conclusions

The $B_d^0 \rightarrow K^{*0}\mu^+\mu^-$ is a very promising channel for the search of NP, not only for the FBA but also for the other asymmetries.
To measure the $A_T^{(1)}$ asymmetry a full 3-dimensional angular analysis with 10 free parameters is necessary. This fit has not yet been implemented.
The $A_T^{(2)}$ asymmetry and the FBA can be extracted by a projective analysis.

The LHCb sensitivity to the FBA in the $B_d^0 \to K^{*0}\mu^+\mu^-$ was already studied in detail in [54]. The very promising control channel $B_d^0 \to K^+\mu^+\mu^-$ was already proposed for θ_l recovering. However this channel cannot be used for θ_K recovering because it is not possible to define the θ_K angle.

For the θ_K recovering the $B_d^0 \to J/\psi K^{*0}(892)$ channel is here proposed. Two recovering methods for the θ_K angle were studied in this thesis (the correction using P_t information and the correction based on the direct θ_K recovering).

This study has also highlighted that the selection cuts must be chosen very carefully to avoid cuts that do not give a flat acceptance function with respect to the angles, like P_t or IPS cuts.

Conclusions

In this thesis three analysis are presented: The sensitivity of the LHCb experiment to the $B^0_{s,d} \to e^{\pm}\mu^{\mp}$ decays, to the $B^0_{s,s} \to \gamma\mu^+\mu^-$ decays and a study of angular biases correction in the $B^0_d \to K^*\mu^+\mu^-$ decay.

The selections for the $B^0_{s,d} \to e^{\pm}\mu^{\mp}$ and $B^0_{s,s} \to \gamma\mu^+\mu^-$ channels were carried out for the first time in LHCb, using a full MC simulation. For both channels, after a preselection done using standard rectangular cuts, various multidimensional methods were compared for the final selection.

Concerning the $B^0_{s,d} \to e^{\pm}\mu^{\mp}$ decays, the upper bound on the branching ratios are $Br(B^0_s \to e^{\pm}\mu^{\mp}) < 1.3 \cdot 10^{-8}$ and $Br(B^0_d \to e^{\pm}\mu^{\mp}) < 3.2 \cdot 10^{-9}$ at 90% CL, in 1 year, running at nominal luminosity. These results can improve substantially the limits from past experiments. These limits were interpreted in the context of various beyond the SM scenarios.

Concerning the $B^0_{s,d} \to \gamma\mu^+\mu^-$, the limits are $Br(B^0_s \to \mu^+\mu^-\gamma) < 5.8 \cdot 10^{-7}$ and $Br(B^0_d \to \mu^+\mu^-\gamma) < 1.4 \cdot 10^{-7}$ at 90% CL, in 1 year, running at nominal luminosity. These limits would represent the world best limits on these branching ratios. Particularly interesting is the $B^0_s \to \mu^+\mu^-\gamma$ which is enhanced by a factor 20 with respect the $B^0_d \to \mu^+\mu^-\gamma$ and for which there is no measurement available.

The $B^0_d \to K^*\mu^+\mu^-$ decay is one of the most promising channel for early measurements in LHCb. The asymmetry $A_T^{(2)}$, which is an observable sensitive to various NP effects, can be measured with a projection study.

This measurement is not straightforward, as the θ_K distribution, necessary to extract $A_T^{(2)}$, is very sensitive to the acceptance. Here two recovering methods (the recovering using P_t information and another one based on direct θ_K recovering) were studied.

Both of them can be used to recover the θ_K distribution. The one based on the direct θ_K recovering gives the best result. However, θ_K is not directly related to the detector acceptance. For this reason also another method were studied. A comparison between the results obtained with the two methods can be done with real data.

Bibliography

[1] K.Riesselmann, arXiv:hep-ph/9711456.

[2] M.J.G.Veltman, Acta Phys. Polon. B **12** (1981) 437.

[3] J.A.Casas, J.R.Espinosa, I.Hidalgo, arXiv:hep-ph/0607279.

[4] R.Barbieri, A.Strumia, arXiv:hep-ph/0007265.

[5] W.Bonivento, N.Serra, LHCb-2007-028, (CERN-LHCb-2007-028).

[6] W.Bonivento, F.Dettori, N.Serra
LHCb sensitivity to $B_s^0 \to \gamma\mu^+\mu^-$,
in preparation.

[7] T.Blake, U.Egede, P.Koppemburg, N.Serra
Correcting biases in angular distribution measurement of the $B_d^0 \to K^{*0}\mu^+\mu^-$,
in preparation.

[8] D.I.Kazakov, arXiv:hep-ph/0012288v2.

[9] R.N.Mohapatra, M.K.Parida, Phys. Rev. **D47**, 264 (1993).

[10] D.G.Lee et al., Phys.Rev. **D51**, 229 (1995).

[11] R.Foot, Phys. Lett. **B420**, 333-339 (1998).

[12] R.Foot, H.Lew, Phys. Rev. **D41** 3502 (1990).

[13] R.Foot, H.Lew, R.R.Volkas, Phys. Rev. **D44** 1531 (1991).

[14] R.Blumenhagen et al., arXiv:hep-ph/050205.

[15] G.Aldazabal et al., Nucl. Phys. **B452**, 3 (1995).

[16] K.Kakushadze, S.H.H.Tye, Phys. Rev. **D54**, 7520 (1998).

[17] K.Kakushadze et al., Int. J. Mod. Phys. **A13**, 2551 (1998).

[18] J.Pati and A.Salam, Phys.Rev.D 10,275(1974).

[19] G.Valencia and S.Villenbrock, Phys.Rev.D 50,6843(1984); the formula reported by us corrects for a misprint already noted by variuos authors e.g. [21] and [20].

[20] M.C.Chang et al.,(Belle Collaboration), Phys.Rev.D 68 (2003)111101.

[21] F.Abe et al.,(CDF Collaboration), Phys.Rev.Lett.81 (1998)5742.

[22] A.Kuznetsov and M.Mikheev, Phys.Lett.B 329(1994)295.

[23] A.Sher et al., (BNL 865 Collab.), Phys. Rev. **D72** 012005 (2005).

[24] R.Appel et al., (BNL 865 Collab.), Phys. Rev. Lett. **85** 2877 (2000).

[25] D.Ambrose et al., (BNL E871 Collab.), Phys. Rev. Lett. **81** 5734 (1998).

[26] F.Abe et al., (CDF Collab.), Phys. Rev. Lett. **81** 5742 (1998).

[27] M.C.Chang et al., (Belle Collab.), Phys. Rev. **D68** 111101 (2003).

[28] W.P.Lam, K.W.Ng, Phys. Rev. **D44** 3345 (1991).

[29] S.Davidson, D.Bailey, B.A.Campbell, Z.Phys **C61** 613 (1994).

[30] M.Tanabashi, PDG *"LeptoQuark quantum number"* (2007).

[31] J.D.Lykken, Phys. ReV. **D54** 3693 (1996).

[32] N.Arkani-Hamed, S.Dimopoulos, G.Dvali, Phys. Lett. **B429** 263 (1998).

[33] S.Davidson et al.,Z.Phys.C 61, 613(1994).

[34] LHCb Collaboration, CERN/LHCC 2003-030, LHCb TDR 9 (2003).

[35] J.Ellis, S.Heinemeyer, K.A.Olive, G.Weiglein, JHEP 0502:013 (2005).

[36] B.Aubert et al., (D0 Collab.) http://www-d0.fnal.gov/Run2Physics/WWW/result/b.htm .

[37] A.Ambulencia et al., (CDF Collab.) Note 8956, http://www-cdf.fnal.gov/physics/new/bottom/bottom.html .

[38] D.Martinez, J.A.Hernando, F.Teubert, LHCb-2007-033, (CERN-LHCb-2007-033).

[39] S.Chen et al., (CLEO Collab.), Phys. Rev. Lett. **81** (2001).

[40] P.Koppenburg et al., (BELLE Collab.) Phys. Rev. Lett. **93** (2004).

[41] A.Aubert et al., (BaBar Collab.), arXiv:hep-ex/0207076.

[42] D.Melikhov, N.Nikitin, Phys. Rev. **D70** 114028 (2004).

[43] T.M.aliev, A.Ozpineci, M.Savci, Phys. Lett. **B520** 69 (2001).

[44] S.R.Choudhry, A.S.Cornell, N.Gaur, G.C.Joshi Int. J. Mod. Phys. **A21** 2617 (2006).

[45] G.Eilam C. Lu, D. Zhang, Phys. Lett. **B391** 461 (1997).

[46] T.M.Aliev, A.Ozpineci, M.Savci, Phys. Rev. **D55** 7059 (1996).

[47] C.Q.Geng, C.C.Lih, W.Zhang, Phys. Rev. **D62** 074017 (2000).

[48] B.Aubert et al., (BaBar Collab.) BaBar-PUB-07/033 (SLAC-PUB-12594), arXiv:hep-ph/07062870v1
Submitted to Phys. Rev. Lett..

[49] F.Kruger, J.Matias, Phys. Rev. **D71** 094009 (2005).

[50] A.Ali, P.Ball, L.T.Handoko, G.Hiller, Phys. Rev. **D61** 074024 (2000).

[51] C.S. Kim, C.D.Lu, T.Morozumi, Phys. Rev. **D62** 034013 (2000).

[52] F.Kruger, L.M.Sehgal, N.Sinha, R.Sinha, Phys. Rev. **D61** 114028 (2000).

[53] E.Lunghi, J.Matias, JHEP **0704** 058 (2007).

[54] J.Dickens, V.Gibson, C.Lazzeroni, M.Patel LHCb-2007-039, (CERN-LHCb-2007-039).

[55] J.Dickens, V.Gibson, C.Lazzeroni, M.Patel LHCb-2007-038, (CERN-LHCb-2007-038).

[56] A.Ishikawa et al., (Belle Collab.), Phys. Rev. Lett. **96** 251801 (2006).

[57] B.Aubert et al., (BABAR Collab.), Phys. Rev. Lett. **91** 221802 (2003).

[58] D.Karlen, Computer in Physics **12:4** 380 (1998).

[59] G.Punzi, PhysStat2003, arXiv:physics/0308063v2.

[60] U.Egede, LHCb-2007-057, (CERN-LHCb-2007-057).

[61] B.Aubert et al., arXiv:hep-ex/07040522v2 .

[62] A.Hocker, P.Speckmayer, J.Stelzer, F.Tegenfeldt, H.Voss, K.Voss TMVA software http://tmva.sourceforge.net/

[63] A.Hocker et al., TMVA User Guide, arXiv:physics/0703039 (2007).

[64] M.C.Chang et al., (Belle Collab.) Phys. Rev. **D68** 111101 (2003).

[65] A.R.Fisher, Annals of Wugenics **7** 179 (1936).

[66] F.Rademakers, R.Brun, ROOT software http://root.cern.ch/.

[67] Software *bayes.f* http://www.cdf.fnal.gov/physics.

[68] W.Verkerke, D.Kirkby RooFit Manual v2.07 arXiv:physics/0306116.

[69] Software *RooFit* http://foofit.sourceforge.net/ .

[70] L.Fernandez, P.Koppenburg, LHCb-2005-047, (CERN-LHCb-2005-047) *Revision 2: August 9, 2006.*

[71] R.K.Rllis, W.J.Stirling, B.R.Webber, *QCD and Collider Physics*, Cambridge University Press, Cambridge (1996).

Ringraziamenti

Questo lavoro non avrebbe mai visto la luce senza l'apporto di alcune persone. Vorrei ringraziare in particolar modo Prof. Biagio Saitta, non solo per l'apporto scientifico ma anche per il suo esempio di correttezza e professionalita', oltre che amore per la scienza.
Ringrazio Walter, senza di lui questa tesi non sarebbe mai stata scritta. Vorrei inoltre ringraziarlo per la sua comprensione, la sua pazienza e il suo supporto in questi anni di lavoro assieme, oltre che per le innumerevoli discussioni di fisica.
Ringrazio di cuore Rudolf per le interessanti discussioni e per i suoi consigli e suggerimenti.
Ringrazio inoltre tutti i componenti del Gruppo I dell'INFN di Cagliari, in particolare Alessandro.
Ringrazio Gino Isodori, per avere avuto la pazienza di leggere questo lavoro e di inviarmi i suoi commenti.
I also want to expecially thank Ulrik Egede for his contribution in improving this thesis. He showed me what being a good physicist means.
Poiche' non ho avuto il dono di esprimermi in versi e poiche' la mia prosa non e' sufficiente ad esprimere tutta la gratitudine per tutte le persone che mi sono state vicine o mi hanno aiutato, in particolare nell'ultimo periodo di tesi, mi limito ad un ingiusto elenco alfabetico dei loro nomi.
Alessandro A., Alessandro C., Alessandro F., Alessandro R., Amit, Andrea, Antonio Sanna, Antonio Serra, Barbara, Battistina, Carla, Carmelo, Carmen, Chiara, Cristian, Davide, Dominik, Emiliano, Enrico, Fabio P., Fabio U., Fabrizio, Federico, Francesco D., Francesco P., Gino, Giovanni, Giulia, Jenny, Jose', Kalina, Margherita, Nicola, Nilu, Pang Pang, Patrick, Raf-

faella, Riccardo, Rita, Roberto C., Roberto M., Rosalba, Rudolf, Sabrina, Salvatore, Sara, Sergio, Silvia, Thomas, Ulrik, Valentina, Violetta, Walter.
Tutte queste persone sappiano che la mia gratitudine nei loro confronti e' sincera e sentita.
Infine ringrazio Alessandra, a cui dedico questa tesi.