UNIVERSITÀ DEGLI STUDI MILANO BICOCCA
Facoltà di Scienze MM. FF. NN.
Laurea Specialistica in Fisica

Measurement of $sin2\beta$ *from* $B^0_d \to J/\Psi K^0_S$ *with the LHCb detector*

Relatore: Prof.ssa Marta Calvi
Correlatore: Prof.ssa Clara Matteuzzi

Tesi di Laurea di
Nicola Mangiafave
Matricola: 080191

29 Aprile 2008
Anno Accademico 2006-2007

To my family

Contents

Preface

A matter–antimatter asymmetry is observed in the universe: the universe is made by matter, the antimatter is absent. CP violation is necessary to explain this asymmetry. The CP symmetry represents the invariance of physical laws under spatial coordinates inversion and particle–antiparticle exchange. It is experimentally violated in the weak interactions and it is conserved in the strong and electromagnetic interactions.

The theories that explain the interactions between particles and that predict their effects are called Standard Model. The Standard Model predicts CP violation in the weak interactions, at present there is no evidence from laboratory experiments that its predictions are wrong. The CP violation predicted by the Standard Model is not enough to explain the matter–antimatter asymmetry observed in the universe. For this reason new experiments are performed to look for CP violation effects beyond the Standard Model.

One of these experiments is LHCb. The detector is a forward spectrometer which will start to take data in summer 2008 at the LHC accelerator, in the CERN laboratories of Geneva. LHC is a ring in which two contrarotating proton beams are accelerated and then collide at an energy of 12 TeV in the centre of mass. LHCb will analyse the proton–proton collisions, to study the particles with a b quark through the reconstruction of their decay products. The b quark system is one of the most promising for the CP violation study.

The CP violation in the b quark mesons was observed for the first time measuring the $\sin 2\beta$ parameter. This parameter is non zero if and only if CP is violated. The first evidence of non zero $\sin 2\beta$ was shown in the 2001 at the Belle and BaBar detectors, from the decay mode $B^0_d \to J/\psi K^0_S$ (called golden channel).

In the first year of data taking at LHCb the $\sin 2\beta$ measurement will be performed in the same channel. The comparison with results obtained in previous experiments will show possible systematic errors due, for instances, to errors in

the calibration of the detector. Once the systematic errors will be understood, the possible discrepancies with the expected results will be the sign of a new physics beyond the Standard Model.

$\sin 2\beta$ is extracted from the asymmetry in the decay $B_d^0 \to J/\psi(\mu^+\mu^-)K_S^0(\pi^+\pi^-)$ and its CP conjugated $\overline{B}_d^0 \to J/\psi(\mu^-\mu^+)K_S^0(\pi^-\pi^+)$. To distinguish between B_d^0 and $\overline{B}_d$ flavour tagging algorithms are used. These introduce an additional parameter: the wrong tag fraction (ω_{tag}). As a consequence the measured parameter is $(1-2\omega_{tag})\sin 2\beta$. ω_{tag} is obtained from the asymmetry measure of the control channel $B_d^0 \to J/\psi(\mu^+\mu^-)K^{0*}(K^+\pi^-)$.

In this thesis the various aspects of the $\sin 2\beta$ measurement at LHCb are presented. The data from the LHCb simulation were analysed to tune the selection of the golden channel and the control channel. For the particles which compose the two decay modes a characterization of the kinematic variables distribution was made. Such distributions were compared with those of the background events, and cuts were applied on the relative variables. These cuts are conditions on the variables chosen in such a way that they are satisfied with an higher probability by the signal events, with respect to the background events. An appropriate choice of cuts brought to a huge reduction of the fraction of background events in the selected events, without reducing too much the fraction of selected events with respect to the produced events. The applied selection does not modify the decay time distribution of the B meson. In fact $\sin 2\beta$ is extracted from a decay time dependent asymmetry. If the distribution of this variable is modified, an acceptance function must be introduced. This function can't be extracted from experimental data. The acceptance function determined from the Monte Carlo simulation can differ significantly from the true one, and this introduces a systematic error.

The selection is also unified, which means that it is as similar as possible for the two channels. At the end of the selection the values of ω_{tag} for the two channels were compared, and it was verified that the results of the control channel can be applied to the golden channel.

On the selected events a model was studied to describe the distributions of mass and decay time of the B meson. This model is a set of probability density functions (*pdf*) which describe, with a very good approximation, the mass and decay time distributions for the signal events and the different background categories. The decay time *pdf* depends on the $\sin 2\beta$ parameter. The mass is used to discriminate between signal and background events. Combining these distributions a total *pdf*

with 16 parameters is obtained, and it will be used to fit real data.

To study the LHCb sensitivity to $\sin 2\beta$ a toy Monte Carlo was performed. A number of events equal to the one selected in one year of data taking at LHCb (10^7 s) were generated following the total *pdf* distribution. From these data and the total *pdf* a likelihood function was made. From the minimization of this function an estimation of $\sin 2\beta$ and its error was obtained. Moreover it was calculated the sensitivity varying the number of background events.

My study uses an accurate description of the background and a selection that does not modify the decay time distribution.

To conclude I tuned a selection for the channel $B_d^0 \to J/\psi(\mu^+\mu^-)K_S^0(\pi^+\pi^-)$ and its control channel $B_d^0 \to J/\psi(\mu^+\mu^-)K^{0*}(K^+\pi^-)$. This selection does not introduce any systematic error related to the acceptance function in the measurement of the $\sin 2\beta$ parameter. I made a model that extract the $\sin 2\beta$ from the real data and I estimated the LHCb sensitivity to this parameter.

Chapter 1

CP violation and b physics

CP violation is one of the main issues in modern physics. It is observed both directly in laboratories, and indirectly in the cosmological asymmetry between matter and antimatter. The two observations don't agree, in the sense that the "cosmological" CP violation is grater than the "laboratory" one. The CP violation observed in the experiments in laboratories is perfectly explained by the Standard Model (SM). This is the theory which describes and predicts the interactions between the particles. Until now there are no experiments which contradict the SM predictions, but there are several indications, one of these is the CP violation issue. The physicists think now that the SM is a low energy theory that hides a new high energy theory. An accelerator and four big experiments are under construction at the CERN laboratory to look for experimental proof of this new physics beyond the SM.

In the following section the CP violation theory and its challenges are described. The experiment which is going to study it will be showed in the following chapter.

1.1 The matter and antimatter asymmetry problem

In the universe a matter–antimatter asymmetry is observed. All things we see in nature are made of matter, and even in the deep universe, massive amounts of antimatter aren't seen. In case amounts of antimatter exist, they should interact with the surrounding matter, annihilating. The annihilation consequences are intense γ–ray emissions that should be visible and detectable from the earth, but

they have never been observed. Another proof of the absence of antimatter in our galaxy is the absence of antinuclei in the cosmic rays. However, the Big Bang should have created equal amounts of matter and antimatter since the Cosmic Microwave Background (CMB) radiation shows that in the early days of the cosmos particles and antiparticles were populating it on equal basis. Assuming that at the very early moments of the Big Bang baryons and antibaryons were created at an equal rate, the only survivors of this "hadron era" would be the nucleons and antinucleons; the rest disappear by decay[1]. The nucleons and antinucleons annihilated into two photons until the antinucleons can no longer find nucleons to annihilate with, and a residue of baryons and antibaryons is "frozen out". From this model the number of baryons and antibaryons left at the equilibrium is the same, and the ratio between the numbers of baryons (antibaryons) and photons can be estimated in:

$$\frac{N_B}{N_\gamma} = \frac{N_{\overline{B}}}{N_\gamma} \simeq 10^{-18} \tag{1.1}$$

The above conclusions should be valid today, but they contrasts with the observed matter–antimatter asymmetry and with :

$$\frac{N_B}{N_\gamma} \simeq 10^{-9}, \ \frac{N_{\overline{B}}}{N_B} < 10^{-18} \tag{1.2}$$

In 1966 Sakharov[2][3] proposed that, assuming initially the number of baryon equal to that of the antibaryon, the current matter–antimatter asymmetry can take place if the following three conditions subsist:

1. the baryon number conservation is violated,
2. deviation from the thermal equilibrium,
3. CP and C symmetry violation.

These three conditions must be valid in the first instants of the Big Bang in order to explain the current asymmetry. The necessity of the first condition is obvious, otherwise in each reaction the same number of baryon and antibaryon would be created (or destructed). The second requirement follows from the fact that, in thermal equilibrium, the density depends only on the temperature and the mass, which is the same for particles and antiparticles for the CPT theorem (see below). Thus at thermal equilibrium the particles and antiparticles density are the same. Thirdly, as it will be illustrated in the next section, C and CP violation are required

to generate a cosmological matter–antimatter asymmetry and they are necessary to distinguish unambiguously between matter and antimatter.

1.2 CP symmetry and CP violation

From the Noether theorem every symmetry of nature implies a conservation law and vice versa. For example the laws of physics are symmetrical with respect to the translation in time, the Noether theorem relates this invariance to conservation of energy. For this reason invariance under translation in space implies momentum conservation, and symmetry under spatial rotation implies angular momentum conservation. Similarly, invariance of electrodynamics under gauge transformations leads to conservation of charge. This is an internal symmetry and involves the fields, in contrast to the space–time symmetries which involve the coordinates. Among the space–time symmetries the rotations and translations are continuous transformations, but there are two space–time discrete symmetries: P and T parity[4]. The P parity consists in the invariance of physics under a discrete transformation which changes the sign of the spaces coordinates x, y and z. The time reversal transformation, usually called T, consists of changing the sign of the time coordinate t. In relativistic quantum field theory a transformation exists which transforms a field ϕ in a related field which has opposite charge, baryon number, lepton number, flavour and so on. If we say that ϕ is associated to a particle than we will say that the transformed field is associated with an antiparticle, and the transformation is called charge (or C) parity. If C symmetry holds, antiparticles behave in exactly the same way as the corresponding particles, and it is a mere matter of convention which of them we call particles and which antiparticles.

Prior to 1956 it was taken for granted that the law of physics are invariant under P transformations. But in 1956 Lee and Yang, in order to solve the θ – τ problem, concluded that the P parity conservation in weak decay was "*only an extrapolated hypothesis unsupported by experimental evidence*" and suggested that P parity is broken in weak interactions. They suggested three experiments to prove this statement, one of them was performed by C. S. Wu[5] one year after. The Wu experiment proved that the P parity is maximally violated in the weak interactions. Soon after it was proved that also the C parity is violated in the weak interactions. In fact when C is applied to a left–handed neutrino, it gives a left–

handed antineutrino, which was never experimentally observed. In 1957, looking for a way to resurrect P-invariance, L.D. Landau stated that weak interactions should be invariant under the product of P reflection and C conjugation (CP symmetry). This theory resisted until the J. W. Cronin, V. L. Fitch[6] experiment performed in 1964 in the K^0 system. The K^0_L is a neutral particle with well defined mass and lifetime. There is no other particles with equal mass, therefore K^0_L must be its own antiparticle (for the CPT theorem, see below). If CP is conserved K^0_L is a CP eigenstate and must decay only in 3 pions, but Cronin and Fitch observed a tiny fraction of $K^0_L \to 2\pi$. A subsequent study of the semileptonic decays revealed an even more dramatic discrepancy between the Branching Ratio (BR) of the $K_L \to \pi^+ e^- \overline{\nu}_e$ and its CP-conjugate mode $K_L \to \pi^- e^+ \nu_e$. The K^0_L decays both to $\pi^+ e^- \overline{\nu}_e$ and to the CP conjugate mode $\pi^- e^+ \nu_e$. It was observed by Cronin and Fitch that it decays slightly less often to the first than to the second mode. This fact unequivocally establishes CP violation. From this experiment it is also clear how the CP violation establishes an unambiguous way to distinguish between matter and antimatter. We can say that the lepton in which a K^0_L decays less often is matter, the lepton in which a K^0_L decays preferentially will be considered as antimatter. Only thirty years later Cronin's and Fitch's discovery, the second major step was done: direct CP violation was observed in kaon decays[7]:

$$\frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_S \to \pi^+\pi^-)} \neq \frac{\Gamma(K_L \to \pi^0\pi^0)}{\Gamma(K_S \to \pi^0\pi^0)} \tag{1.3}$$

Finally, in the year 2001 CP violation was for the first time observed in the time dependent asymmetry of B^0 decays:

$$a(t) = \frac{N(B^0 \to J/\Psi K_{S(L)}) - N(\overline{B}^0 \to J/\Psi K_{S(L)})}{N(B^0 \to J/\Psi K_{S(L)}) + N(\overline{B}^0 \to J/\Psi K_{S(L)})} \neq 0 \tag{1.4}$$

Until now the C, P and CP symmetries are proved to be conserved in the electromagnetic and strong interactions and violated in the weak interactions. There aren't experiments which prove the violation of the T invariance in any type of interaction, but there is a compelling reason to believe that time reversal is violated in the weak interactions. It comes from the so called CPT theorem. Based only on the most general assumptions (Lorentz invariance, quantum mechanics and the idea that interactions are carried by fields), the CPT theorem states that the combined operation of charge conjugation, P parity, and time reversal (in any order) is an exact symmetry of all interactions. So if CP is violated there must be

a compensating violation of T. The CPT theorem has two important implications, that were proved experimentally at a very high level of precision: if the theorem is correct, every particle must have precisely the same mass and lifetime as its antiparticle.

1.3 CP violation in the Standard Model

In constructing the Standard Model Lagrangian the piece of the Lagrangian from which the up quarks (u) get their masses looks like:

$$\Delta\mathcal{L}_{\rm up} = f^{(u)}_{ik}\overline{Q}'^{i}_{L}u'^{k}_{R}H + \text{c.c.}\ , \quad i,k = 1,2,3 \tag{1.5}$$

where

$$Q'^{1}_{L} = \begin{pmatrix} u' \\ d' \end{pmatrix}_{L}\ , \quad Q'^{2}_{L} = \begin{pmatrix} c' \\ s' \end{pmatrix}_{L}\ , \quad Q'^{3}_{L} = \begin{pmatrix} t' \\ b' \end{pmatrix}_{L}\ ;$$
$$u'^{1}_{R} = u'_{R}\ , \quad u'^{2}_{R} = c'_{R}\ , \quad u'^{3}_{R} = t'_{R} \tag{1.6}$$

H is the Higgs doublet:

$$H = \begin{pmatrix} H^{0} \\ H^{-} \end{pmatrix} \tag{1.7}$$

and $f^{(u)}_{ik}$ is a coefficient. The piece of the Lagrangian which is responsible for the down quark (d) masses is similar:

$$\Delta\mathcal{L}_{\rm down} = f^{(d)}_{ik}\overline{Q}'^{i}_{L}d'^{k}_{R}\tilde{H} + \text{c.c.}\ , \tag{1.8}$$

where

$$d'^{1}_{R} = d'_{R}\ , \quad d'^{2}_{R} = s'_{R}\ , \quad d'^{3}_{R} = b'_{R} \ \text{ and } \ \tilde{H}_{a} = \varepsilon_{ab}H^{*}_{b}\ ,$$
$$\varepsilon_{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\ . \tag{1.9}$$

After $SU(2)\times U(1)$ spontaneous symmetry breaking, from formulas 1.5 and 1.8 two mass matrices emerge:

$$M^{ik}_{\rm up}\overline{u}'^{i}_{L}u'^{k}_{R} + M^{ik}_{\rm down}\overline{d}'^{i}_{L}d'^{k}_{R} + c.c. \tag{1.10}$$

The matrices $M_{\rm up}$ and $M_{\rm down}$ are arbitrary 3×3 matrices; their matrix elements are complex numbers. They can be written as a product of an hermitian and a unitary matrix:

$$M = UA\ , \ \text{ where } \ A = A^{+}\ , \ \text{ and } \ UU^{+} = 1\ , \tag{1.11}$$

It can be also shown that the matrix M can be diagonalized by 2 different unitary matrices acting from left and right:

$$U_L M U_R^+ = M_{\text{diag}} = \begin{pmatrix} m_u & & 0 \\ & m_c & \\ 0 & & m_t \end{pmatrix} \tag{1.12}$$

where m_i are real numbers (if M is hermitian ($M = M^+$) then we will get $U_L = U_R$). Having these formulas in mind the up-quarks mass term from eq. 1.10 can be written like:

$$\overline{u}_L'^i M_{ik} u_R'^k + c.c. \equiv \overline{u}_L' U_L^+ U_L M U_R^+ U_R u_R' + c.c. = \overline{u}_L M_{\text{diag}} u_R^+ + c.c. = \overline{u} M_{\text{diag}} u \ , \tag{1.13}$$

where we introduce the fields u_L and u_R according to the following formulas:

$$u_L = U_L u_L' \ , \quad u_R = U_R u_R' \ . \tag{1.14}$$

These fields are only the result of a change of basis performed by the U matrices. The change of basis is done to ensure a diagonal mass matrix; in fact, in the Lagrangian, the mass terms of the fields must be of the form $m\overline{\phi}\phi$ to have physical significance. Applying the same procedure to the matrix M_{down} we observe that it becomes diagonal as well in the rotated basis:

$$d_L = D_L d_L' \ , \quad d_R = D_R d_R' \ . \tag{1.15}$$

Thus we start from the primed quark fields and get that they should be rotated by 4 unitary matrices U_L, U_R, D_L and D_R in order to obtain unprimed fields with diagonal mass matrices. Since kinetic energies and interactions with the vector fields A_μ^3, B_μ and gluons are diagonal in the quark fields, these terms remain diagonal in a new unprimed basis. The only term in the Standard Model Lagrangian where matrices U and D show up is the charged current interaction with the emission of W-boson:

$$\Delta\mathcal{L} = g W_\mu^+ \overline{u}_L' \gamma_\mu d_L' = g W_\mu^+ \overline{u}_L \gamma_\mu U_L^+ D_L d_L \ , \tag{1.16}$$

and the unitary matrix $V \equiv U_L^+ D_L$ is called the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix.

A unitary $n \times n$ matrix can be parametrized by n^2 independent real quantities ($2n^2$ real parameters minus n^2 unitarity conditions). $n(n-1)/2$ of these are real,

and are called Euler angles because they are associated with rotations in a n–dimensional space. The remaining $n(n+1)/2$ are complex and are called phases. Not all phases have physical meaning, as some may be removed by a new definition of the quark fields. In fact the Lagrangian is invariant under a transformation of the quarks phases. Of these $2n$ field phases (n from the up–type quarks and other n from the down–type quarks), $(2n-1)$ are arbitrary. Thus, the number of measurable phases in the CKM matrix is $(n-1)(n-2)/2$. Thus the 3×3 CKM matrix is specified by three Euler angles and one phase angle, say δ. The phase will enter the wave function as $e^{i(\omega t+\delta)}$ and clearly this is not invariant under time reversal $t \to -t$. So this phase introduces the important possibility of T–violating, or equivalently, CP–violating amplitudes in the Standard Model.

Historically Kobayashi and Maskawa in 1973[10] proposed a third quark generation before the second was complete, and long before any experimental evidence for a third generation of leptons or quarks was found. They were wondering how to explain the CP violation within the Cabibbo–GIM scheme, and they understood that a complex number was needed in the quark mixing matrix. For the reasons explained above, such a term can always be eliminated in a 2×2 matrix, with a suitable redefinition of quarks phases. So the minimal dimension of the rotation matrix, in order to have CP violation, is 3×3, and hence 3 must be the minimal number of quark families.

1.3.1 CKM matrix

The CKM matrix describes the rotation between the weak eigenstates (d', s', b') and the mass eigenstates (d, s, b), and gives the weak charged current coupling between quarks of different flavour (see section 1.3):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} , \tag{1.17}$$

where V_{ij} is the matrix element coupling the i^{th} up-type quark to the j^{th} down-type quark. The standard parametrization of the CKM matrix is:

$$V_{CKM} = \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & s_{13}e^{-i\delta} \\ -c_{23}s_{12} - s_{23}s_{13}c_{12}e^{i\delta} & c_{23}c_{12} - s_{12}s_{13}s_{23}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -s_{23}c_{12} - c_{23}s_{13}s_{12}e^{i\delta} & c_{23}c_{13} \end{pmatrix} , \tag{1.18}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ $(i,j = 1,2,3)$. However it is useful to reparametrize V_{ik} with the help of the so-called Wolfenstein parametrization[11]. We can define:

$$\lambda \equiv s_{12}\ ,\quad A \equiv \frac{s_{23}}{s_{12}^2}\ ,\quad \rho = \frac{s_{13}}{s_{12}s_{23}}\cos\delta\ ,$$
$$\eta = \frac{s_{13}}{s_{12}s_{23}}\sin\delta\ , \tag{1.19}$$

where $\lambda \sim 0.22$ is the Cabibbo angle and η represents the CP violating phase. Then we can write the V_{CKM} expanded in terms of λ:

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\overline{\rho}-i\overline{\eta}) \\ -\lambda - iA^2\lambda^5\overline{\eta} & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\overline{\rho}-i\overline{\eta}) & -A\lambda^2 - iA\lambda^4\overline{\eta} & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^6\right)\ , \tag{1.20}$$

where

$$\overline{\rho} \equiv \rho\left(1-\frac{\lambda^2}{2}\right)\ ,\ \overline{\eta} \equiv \eta\left(1-\frac{\lambda^2}{2}\right)\ . \tag{1.21}$$

This last form of CKM matrix is very convenient for qualitative estimations, and is numerically accurate as much as the degree of expansion. It also makes manifest the hierarchy of the matrix elements: the diagonal elements are of order one, the elements 12 and 21 are of order λ , the elements 32 and 23 are of the order of λ^2 (~ 0.04) and the other are of the order of λ^3 (~ 0.008).

1.3.2 Unitarity Triangle

The unitarity of the matrix V_{CKM} leads to the following six equations (under each term in these equations the power of λ entering it, is shown):

$$\underset{\sim\lambda}{V_{ud}^*V_{us}} + \underset{\sim\lambda}{V_{cd}^*V_{cs}} + \underset{\sim\lambda^5}{V_{td}^*V_{ts}} = 0 \qquad (ds) \tag{1.22}$$

$$\underset{\sim\lambda^3}{V_{ud}^*V_{ub}} + \underset{\sim\lambda^3}{V_{cd}^*V_{cb}} + \underset{\sim\lambda^3}{V_{td}^*V_{tb}} = 0 \qquad (db) \tag{1.23}$$

$$\underset{\sim\lambda^4}{V_{us}^*V_{ub}} + \underset{\sim\lambda^2}{V_{cs}^*V_{cb}} + \underset{\sim\lambda^2}{V_{ts}^*V_{tb}} = 0 \qquad (sb) \tag{1.24}$$

and

$$\underset{\sim\lambda}{V_{ud}V_{cd}^*} + \underset{\sim\lambda}{V_{us}V_{cs}^*} + \underset{\sim\lambda^5}{V_{ub}V_{cb}^*} = 0 \qquad (uc) \tag{1.25}$$

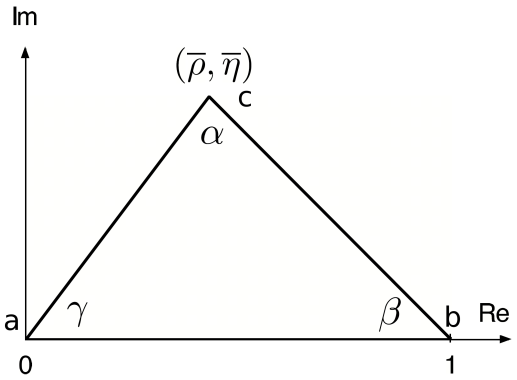

Figure 1.1: The unitarity triangle.

$$\underset{\sim\lambda^3}{V_{ud}V_{td}^*} + \underset{\sim\lambda^3}{V_{us}V_{ts}^*} + \underset{\sim\lambda^3}{V_{ub}V_{tb}^*} = 0 \qquad (ut) \tag{1.26}$$

$$\underset{\sim\lambda^4}{V_{cd}V_{td}^*} + \underset{\sim\lambda^2}{V_{cs}V_{ts}^*} + \underset{\sim\lambda^2}{V_{cb}V_{tb}^*} = 0 \qquad (ct) \tag{1.27}$$

These are the results of the orthogonality between any pair of columns or any pair of rows. The (ij) represents the i^{th} and the j^{th} column or row. Each equation requires the sum of three complex numbers to vanish and so they can be represented geometrically in the complex plane as a triangle. Among these triangles, known as unitarity triangles, four are almost degenerate: one side is much shorter than the two others. Only two triangles, expressed by equations (1.23) and (1.26), have all three sides of comparable lengths (order of λ^3). The two triangles are identical to $\mathcal{O}(\lambda^3)$ and differ only by $\mathcal{O}(\lambda^5)$ correction. So, we can consider only one nondegenerate unitarity triangle; it is usually described by a complex conjugate of the equation (1.23):

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \ . \tag{1.28}$$

It is redrawn by choosing a phase convention such that $V_{cd}V_{cb}^*$ is real and dividing the length of all the sides by $|V_{cd}V_{cb}^*|$.

The unitarity triangle (db) is commonly referred as *the* unitarity triangle, and its angles are commonly called β, α and γ (according to western notation) or ϕ_1, ϕ_2 and ϕ_3 (according to eastern notation). As shown in figure 1.1 it is a triangle in the $(\overline{\rho},\overline{\eta})$ complex plane with two verteces at $\overline{\eta} = 0$. So, if and only if CP is violated: the $\overline{\eta}$ parameter is different from 0, the triangle is non degenerate and the angles γ and β are nonzero. From the unitarity triangle one can easily obtain

the following definitions:

$$\alpha = \arg\left(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}}\right), \quad \beta = \arg\left(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}}\right), \quad \gamma = \arg\left(-\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}}\right). \tag{1.29}$$

The multiplication of any up quark field by a phase does not change the unitarity triangle. While multiplying d- or b-quark field on a phase we will rotate it as a whole not changing its angles (which are physical observables). So we can conclude that four quantities are needed to specify CKM matrix: s_{12}, s_{13}, s_{23} and δ, or λ, A, ρ and η, and the full description of the CP violation in the Standard Model is given by the parameters of the unitarity triangle.

1.4 CP violation in mesons decays

1.4.1 Neutral meson mixing

Due to the non–conservation of flavour in the weak interactions, the oscillation (also called mixing) between neutral meson particle and antiparticle can be observed. In order to mix, the mesons must be neutral. They must not coincide with their antiparticles and they must decay through weak interactions. There are four such pairs:

$$K^0(\bar{s}d) - \overline{K}^0(s\bar{d}) \ , \quad D^0(c\bar{u}) - \overline{D}^0(\bar{c}u) \ ,$$
$$B_d^0(\bar{b}d) - \overline{B}_d^0(b\bar{d}) \ \text{ and } \ B_s^0(\bar{b}s) - \overline{B}_s^0(b\bar{s}) \ .$$

Mixing in the Standard Model occurs in the second order weak interactions through the box diagram which is shown in figure 1.2 for $B^0 - \overline{B}^0$ pair.

Being $\mid X^0 >$ a neutral meson state and $\mid \overline{X^0} >$ the antimeson, the time development of an arbitrary state $a(t) \mid X^0 > + b(t) \mid \overline{X^0} >$ is given by the Schrodinger equation:

$$i\frac{\partial}{\partial t}\begin{pmatrix} a \\ b \end{pmatrix} = H\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \tag{1.30}$$

where H is the effective 2×2 Hamiltonian used to describe the meson-antimeson mixing. This effective Hamiltonian is not hermitian since it takes mesons decays into account. It can be written as: $H = M - \frac{i}{2}\Gamma$, where both M and Γ are hermitian matrices. As a consequence of the CPT theorem (particle and antiparticle masses

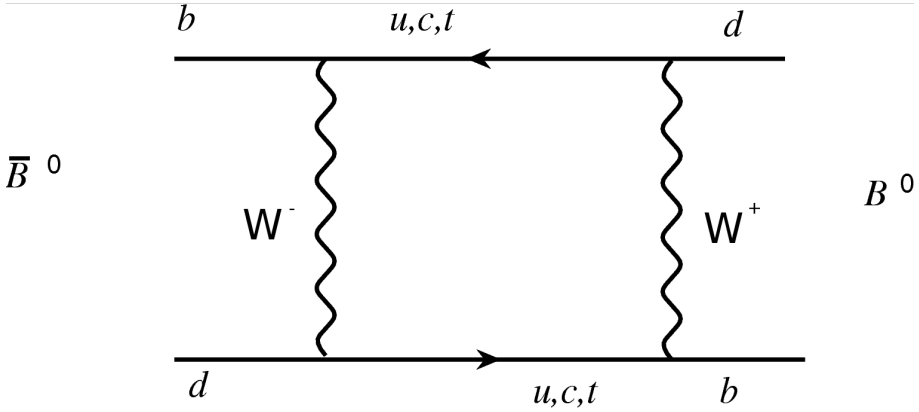

Figure 1.2: $B^0 - \overline{B}^0$ mixing diagram.

and lifetimes are identical) $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. The eigenstates of H, are not $\mid X^0 >$ and $\mid \overline{X^0} >$, but $\mid X_1 >$ and $\mid X_2 >$ defined as:

$$\begin{aligned} \mid X_1 >= p \mid X^0 > +q \mid \overline{X^0} > \ , \\ \mid X_2 >= p \mid X^0 > -q \mid \overline{X^0} > \ . \end{aligned} \tag{1.31}$$

p and q represent the amount of meson state mixing and are complex numbers with $|p|^2 + |q|^2 = 1$. Therefore the weak physical states are a mixing of the mass eigenstates. This is not true for the electromagnetic and strong interactions in which the mass and interaction eigenstates are the same. The corresponding eigenvalues are:

$$E_{1,2} = \left(M \mp \frac{\Delta M}{2} \right) - \frac{i}{2} \left(\Gamma \mp \frac{\Delta \Gamma}{2} \right) \ , \tag{1.32}$$

where if $M_{1,2}$ and $\Gamma_{1,2}$ are the masses and decay widths of the two eigenstates, we have: $\Delta M = M_2 - M_1$ and $\Delta \Gamma = \Gamma_2 - \Gamma_1$.

Substituting the time evolution of the mixing eigenstates

$$\mid X_{1,2}(t) >= e^{-iE_{1,2}t} \mid X_{1,2} >$$

in equation 1.31 and inverting, we can find the time evolution for the mass eigenstates:

$$\begin{aligned} \mid X^0(t) >= f_+(t) \mid X^0 > + \frac{q}{p} f_-(t) \mid \overline{X^0} > \ , \\ \mid \overline{X^0}(t) >= f_+(t) \mid \overline{X^0} > + \frac{p}{q} f_-(t) \mid X^0 > \ , \end{aligned} \tag{1.33}$$

where

$$f_{\pm}(t) = \frac{1}{2}\left[e^{-iE_1 t} \pm e^{-iE_2 t}\right] \quad . \tag{1.34}$$

This means that if we produce a $\mid X^0 >$ at the time 0, at a time $t > 0$ it will be observed a $\mid X^0 >$ with a probability $|f_+(t)|^2$ and a $\mid \overline{X^0} >$ with a probability $|\frac{q}{p} f_-(t)|^2$. In particular:

$$|f_{\pm}(t)|^2 = \frac{1}{4}\left[e^{-\Gamma_1 t} + e^{-\Gamma_2 t} \pm 2e^{-\left(\frac{\Gamma_1+\Gamma_2}{2}\right)t}\cos\left(\Delta M t\right)\right] \quad . \tag{1.35}$$

The last term in equation 1.35 describes the oscillation character of the X^0 and $\overline{X^0}$ content in an initial pure X^0 beam. This means that if we take an initial pure X^0 sample and we observe the number of particles which decay as an X^0 as a function of time, we will see an exponential distribution with a superimposed oscillation of frequency ΔM.

1.4.2 CP violation types

The time dependent decay rates for the neutral mesons X^0 and $\overline{X^0}$ in a final state f are:

$$\begin{aligned} \Gamma_f(t) &\equiv \Gamma\left(X^0(t) \to f\right) \\ &= |A_f|^2 \left[|f_+(t)|^2 + \left(\frac{q}{p}\frac{\overline{A_f}}{A_f}\right)^2 |f_-(t)|^2 + 2Re\left\{\frac{q}{p}\frac{\overline{A_f}}{A_f} f_+^*(t) f_-(t)\right\}\right] \end{aligned} \tag{1.36}$$

$$\begin{aligned} \overline{\Gamma_f}(t) &\equiv \Gamma\left(\overline{X^0}(t) \to f\right) \\ &= |A_f|^2 \left[\left|\frac{\overline{A_f}}{A_f}\right|^2 |f_+(t)|^2 + \left(\frac{p}{q}\right)^2 |f_-(t)|^2 + 2\left(\frac{p}{q}\right)^2 Re\left\{\frac{q}{p}\frac{\overline{A_f^*}}{A_f} f_+^*(t) f_-(t)\right\}\right] \end{aligned} \tag{1.37}$$

Where $A_f = \langle f|H|X^0\rangle$ and $\overline{A_f} = \langle f|H|\overline{X^0}\rangle$ are the decay amplitudes. Any difference between these two rates is a clear proof of CP violation. The first terms in Γ_f and $\overline{\Gamma_f}$ show that CP violation is generated if:

$$|A_f| \neq |\overline{A_f}| \quad . \tag{1.38}$$

This is called *CP violation in the decay amplitudes*. From the second term in the rates it's clear that CP is violated if:

$$\left|\frac{q}{p}\right| \neq 1 \quad . \tag{1.39}$$

In this case the oscillation rate for $X^0 \to \overline{X^0}$ is different from that for $\overline{X^0} \to X^0$. This type of CP violation is called *CP violation in the mixing*, and results in two physical eigenstates which are different from CP eigenstates. If the first and the second terms are identical between the two rates, CP violation can still arise due to the third term. If the first and the second terms are identical the two amplitudes differ if and only if:

$$Im\left\{\frac{q}{p}\frac{\overline{A_f}}{A_f}\right\} \neq 0 \quad . \tag{1.40}$$

Since this process involves both the decay amplitudes and mixing, it is commonly referred to as *CP violation in the interference between the mixing and decay amplitudes*.

Another widely used notation is the following. When the CP violating effects are entirely independent from the final state, the CP violation is called *indirect*. This is the case of the CP violation in the mixing. Conversely the process is called a *direct* CP violation, and this is the case of the CP violation in the decay. The interference contains aspects of both indirect and direct CP violation.

1.4.3 CP violation in the B system

As shown in section 1.3.2 the CP violation parameter is well represented in the unitarity triangle, by the position of the vertex C. It corresponds to the unitarity condition applied to the d and b column of the CKM matrix. Thus decays which involves b quarks are favourite to show CP violating effects. There are several other advantages in looking for CP violation in the b–system. First of all the b quarks are the heaviest quarks which hadronize. Therefore the Standard Model predictions are calculated with lower theoretical uncertainties. Moreover the B mesons have many decay modes, that allow to measure the angles of the unitarity triangle in several ways. These independent measurements can show the presence of new physics in case they don't agree each other.

1.5 Experimental status of CP violation in the B system

In 2001 CP violation was observed for the first time measuring a non–zero time dependent asymmetry of $B^0_d \to J/\psi K_s$[8][9]. It was a clear sign of non–zero β

angle in the unitarity triangle. This observation was done in the BaBar and Belle experiments: which worked at e^+e^- colliders at the energy of the $\Upsilon(4S)$, i.e. as B–factories. A big contribution to the B physics come by CDF and D0, the experiments which work at the Tevatron, a p^+p^- accelerator.

As already stated the physicists are interested in measuring the unitarity triangle in as many different ways as possible. From different measures of the same parameters can emerge a possible inconsistency which would be a clear sign of physics beyond the Standard Model. Until now all the experiments seem to confirm the Standard Model predictions, even in the CP–violation sector. The problem is that the CP violation predicted in the Standard Model is too small to explain the matter–antimatter asymmetry experimentally observed in the universe. This induces the physicists to search new source of CP violation, and hence new physics. This new physics can manifest itself in different ways. For example, the presence of a particle in a loop can change the Branching Ratio of a certain decay, even if the particle is too heavy to be produced directly. So there is a great interest in looking for rare decays, even non CP–violating, and this is another field in which the b–system is competitive. For the reason showed above a dedicated experiment is under construction at CERN: it is called LHCb, and will be described in the next chapter.

Chapter 2

The LHCb experiment at LHC

To study the physics beyond the Standard Model particles accelerators with high energy and high luminosity are needed. With this goal the Large Hadron Collider (LHC) is under construction at CERN. This accelerator will have the highest luminosity and energy in the centre of mass ($\sqrt{s} = 14$ TeV). At this energy the $b\bar{b}$ production cross section is 500 μb. The b quarks will be analysed by one of the four LHC experiments: LHCb. LHCb will study the CP violation in the b system. In this chapter the LHC and LHCb will be described.

2.1 The Large Hadron Collider

The LHC [12] will produce head–on collisions between two beams of either protons or lead ions. It is a 27 km ring made of 1232 dipole superconductive magnets 14.2 m long, which surround a double vacuum pipe. The ring is installed in a tunnel 3.8 m in diameter, buried 50 to 175 m below ground. Located between the Jura mountain range in France and Lake Geneva in Switzerland, the tunnel was built in the 1980s for the previous big accelerator, the Large Electron Positron collider (LEP).

The beams will be created in CERN's existing chain of accelerators and then injected into the LHC. At the last state of preacceleration, two counter rotating beams will be injected at 450 GeV into the LHC from the SPS accelerator (the Super Proton Synchrotron). The superconducting niobium-titanium magnets are positioned inside a cryostat containing superfluid helium at a temperature of 1.9 °K. They will provide a magnetic field up to 8.34 Tesla and they will guide the

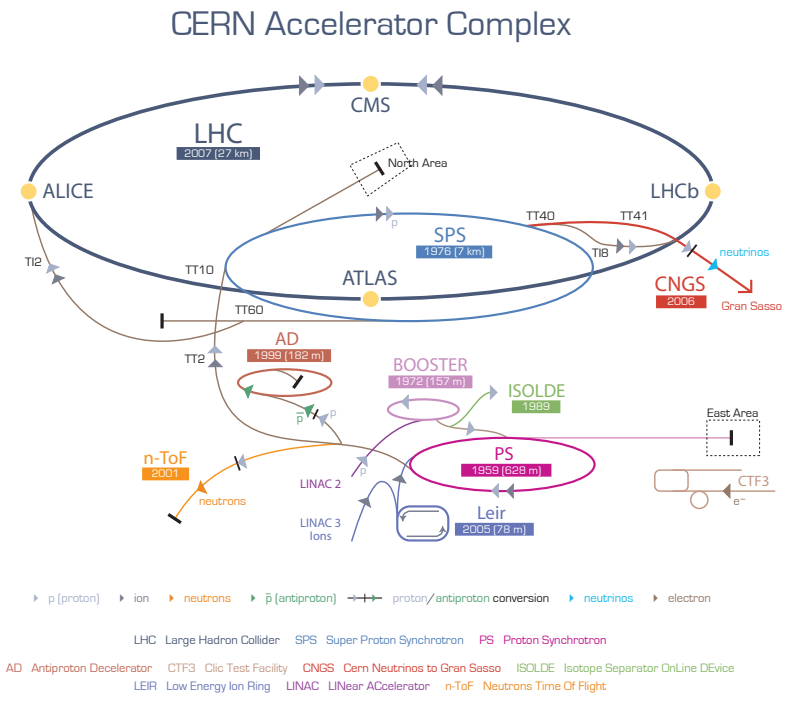

Figure 2.1: LHC accelerator complex with its four main experiments.

beams around the ring while the radiofrequency cavities accelerate each proton beam until a mean energy of 7 TeV. The lead ions will have a collision energy of 575 TeV. Each beam will consist of 2808 bunches of particles. The bunches will be separated by an interval of 25 ns and will contain as many as 10^{10} particles. The beams will be stored at high energy for 10–20 hours. During this time collisions take place inside the four main LHC experiments. When two bunches cross, there will be about 20 collisions at the design luminosity of 10^{34} cm^{-2}s^{-1}, and 1 collision at the LHCb luminosity of $2 \cdot 10^{32}$ cm^{-2}s^{-1}. The luminosity decreases exponentially with a characteristic time of 10 hours. After 10 to 20 hours the beam will be dumped and the LHC will be refilled.

First collisions at high energy are expected mid-2008. The accelerator is expected to run at an initial luminosity of 10^{33} cm^{-2}s^{-1} and it will reach the design luminosity of 10^{34} cm^{-2}s^{-1} the year after. The LHC will provide collisions at the highest energies ever observed in laboratory conditions. Four detectors – ALICE, ATLAS, CMS and LHCb – will observe the collisions.

ATLAS[13] and CMS[14] are two general-purpose detectors. They will inves-

tigate a wide range of physics, including the search for the Higgs boson, extra dimensions, and particles that could make up dark matter like supersymmetry particles. For the ALICE[15] experiment, the LHC will collide lead ions. The data obtained will allow physicists to study a state of matter known as quark−gluon plasma, which is believed to have existed soon after the Big Bang. A fifth experiment, TOTEM[16], installed with CMS, will study collisions in the forward region not accessible to the general-purpose experiments. Among a range of studies, it will measure the total cross section and also monitor accurately the LHC luminosity.

2.2 The LHCb Detector

LHCb[17][18] is the experiment dedicated to b physics at the LHC. It will exploit the large number of b hadrons produced at the LHC to study, with a very high precision, CP violation and rare decays in the b system. The detector is designed and constructed by a collaboration of about 500 physicists, technicians and engineers from 47 universities and laboratories from 15 countries. In the following sections the detector components will be discussed, together with the physics aspects of the experiment.

2.2.1 Physics motivations

The cross section of $b\bar{b}$ production at LHC, even if with large uncertainties, is estimated to be $\sigma_{b\bar{b}} = 500$ μb. This fact, and the high luminosity, will make LHC the most copious source of B mesons compared to the other accelerators operating in the world. At a centre of mass energy of $\sqrt{s} = 14$ TeV all kinds of b hadrons are produced, mesons and baryons.

Since $b\bar{b}$ hadrons at the LHC will be predominately produced at low polar angles in the same forward cone (see Fig. 2.2), the LHCb detector is designed as a single-arm, forward spectrometer with an angular coverage from approximately 10 mrad to 300 (250) mrad in the horizontal (vertical) plane.

Another way to search for physics beyond the Standard Model is to study B-meson decays that are rare or even forbidden in the Standard Model. A possible larger branching ratio (BR) observed for these types of decay represents evidence of physics beyond the Standard Model. To study rare decays a large cross section

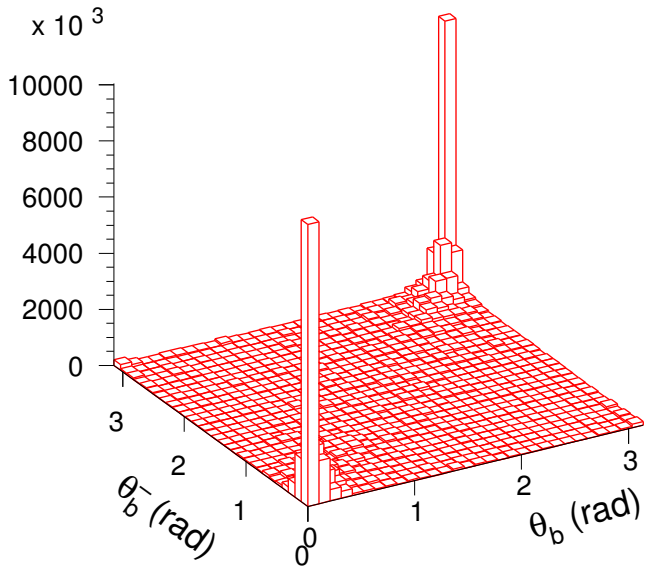

Figure 2.2: Polar angles of the b and $\bar{b}$ hadrons calculated by PYTHIA event generator.

and luminosity is not enough, it's also necessary to have a detector which is able to reconstruct and identify particles with a very high efficiency, in order to separate the rare events from background. To achieve this it is important to have a good vertex resolution in order to separate the B decay vertex from the production vertex. The good vertex resolution is needed also to achieve a good proper time resolutions for the study of the rapidly oscillating B_s mesons and their CP asymmetries. Multiple pp collisions within the same bunch crossing (so called pile–up) can significantly complicate the discrimination between B decay vertex and its corresponding production vertex, besides increasing the background. Therefore, the LHCb experiment will operate at an average luminosity of only 2×10^{32} cm^{-2}s^{-1}, which is less than the designed luminosity of the LHC. This is achieved by slightly changing the focusing of the proton beams at the interaction point. It reduces the mean number of inelastic pp collisions per filled bunch crossing from 27 to 0.53. Another benefit of running at a moderate luminosity is the reduction of radiation damage on the detectors and readout electronics. In LHCb, all subdetectors are

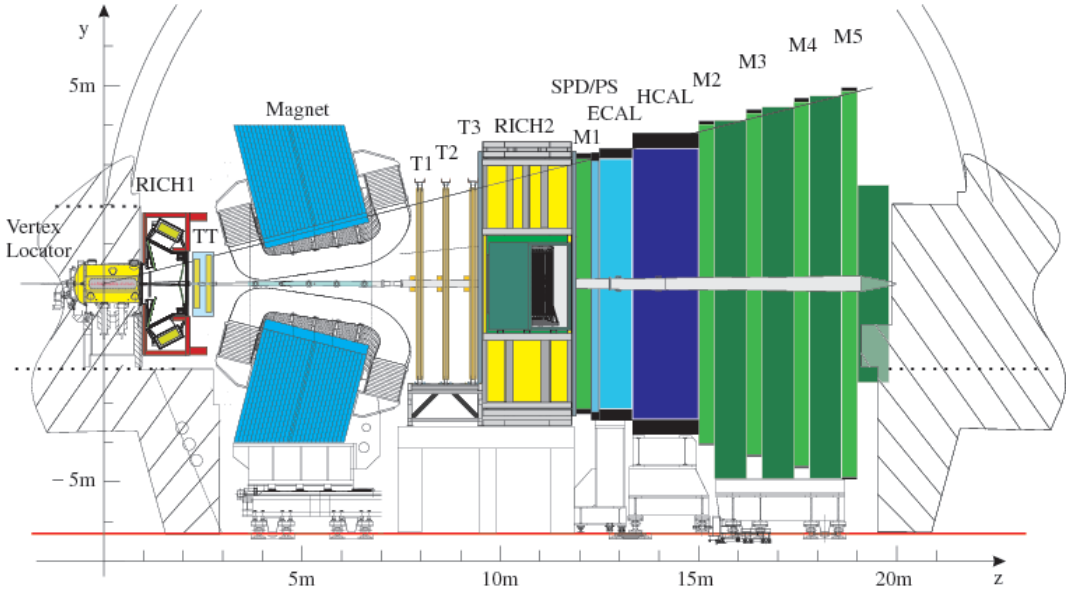

Figure 2.3: Side view of the detector in the non-bending plane

designed to cope with a luminosity of 5×10^{32} cm^{-2}s^{-1}. This leaves room for running at slightly higher luminosities. The integrated luminosity $L_{int} = \int_{year} Ldt$, in one nominal year of 10^7 s is 2 fb^{-1}. Hence, the number of $b\bar{b}$ pairs produced in the LHCb experiment during a nominal year is expected to be $\sigma_{b\bar{b}} \times L_{int} = 10^{12}$.

The high production rate of charmed particles ($\sigma_{c\bar{c}} = 3.600\ \mu$b) allow also the study of CP violation in the c system. For example at LHCb will be investigated the $D^0 - \overline{D}^0$ oscillations. Other physics topics (e.g., Higgs) will also be investigated.

2.2.2 Detector layout

The layout of the LHCb spectrometer is shown in figure 2.3. The coordinate system is: the z axis along the beam, the y axis along the vertical, and the x axis completes the right-handed coordinate system.

LHCb is being installed at IP8 cavern at CERN (see figure 2.1), the site occupied by Delphi during LEP times. A modification to the LHC optics, displacing the interaction point by 11.25 m from the centre, has permitted maximum use to be made of the existing cavern by freeing 19.7 m for the LHCb detector components. All detector subsystems, except the vertex detector, are assembled in two halves, which can be separated horizontally for assembly and maintenance, as well as to provide access to the beam pipe.

A great attention was put into keeping small the material budget. A large amount of material increases secondary interactions. The emission of high energy γ by bremsstrahlung and the multiple scattering of the charged hadrons, degrade the momentum resolution. The reduction of material budget implied a big effort in the optimization of the detector[18]. At the end of the tracking system the particles that come from the interaction point have seen a mean radiation length of $X_0 = 60\%$, and a mean absorption length of $\lambda_I = 20\%$[19].

2.2.3 Beam Pipe

The beam pipe[17][18] is a cylindrical structure sealed to the forward exit window of the vertex locator and composed of four conical sections. From the VELO exit window, a 2 mm thick foil of aluminium, starts the first section of the beam pipe. This is 1840 mm long and is composed by two cones of 1 mm thick beryllium. The cone nearest to the interaction point has a 25 mrad angle, the following has a 10 mrad angle, as the other sections of the beam pipe. The connection between the two cones is made by a window, followed by a short cylindrical section (250 mm in length), with an inner diameter of 50 mm. The first section passes through the Ring Imaging Cherenkov Detector 1 and the Trigger Tracker. The section 2 (which traverses the dipole magnet) and section 3 (which passes through the tracker stations, Ring Imaging Cherenkov Detector 2, the first Muon Station and part of Electromagnetic Calorimeter) are long 3876 mm and 6000 mm respectively. They are 10 mrad beryllium cones formed by several pieces of increasing thickness (1.0 mm to 2.4 mm) welded together. A stainless steel 10 mrad cone 5300 mm long and 4 mm thick, connected with the previous section by a stainless steel bellow 1300 mm long, is the last section of the beam pipe.

The choice of beryllium for the part of the beam pipe that traverse the tracking system and the Ring Imaging Cherenkov Detectors guarantees a reduction of material budget in the highest occupancy area of the detector. This is due to the beryllium higher radiation length with respect to the stainless steel. Beryllium has also an higher modulus of elasticity which permits to built a thin beam pipe mechanically strong enough to stand the difference in pressure between the ultra high vacuum and the surrounding ambient. Because of its toxicity, fragility and cost, it was used only in the areas where an high transparency was needed. In the upstream region, where the particle flux is lower, the classical solution of stainless steel was chosen, which has good mechanical and vacuum properties.

2.2.4 Vertex Locator

The VErtex LOcator (VELO)[17][18][20] provides precise measurements of track coordinates close to the interaction region, and allows to reconstruct the position of the primary vertex as well as of secondary verteces. The b and c–hadrons have a quite long lifetime and, due to their velocity near the speed of light at their production, they exhibit decay verteces displaced from the production one. The presence of a decay vertex far from the primary vertex is an important signature of events in which a pair of $b\bar{b}$ or $c\bar{c}$ is produced. This signature is used by the Trigger system to enrich the amount of $b\bar{b}$ or $c\bar{c}$ events in the saved data, as well as in the LHCb off–line analysis. An accurate measurement of the verteces positions is needed to provide an accurate measurement of the B and D mesons time of flight, and to measure the impact parameter of particles used to tag the b flavour. The VELO is also the first module of the tracking system, therefore it provides information on the momentum of the particles.

The VELO consists of a series of 21 silicon modules positioned along and perpendicular to the beam axis for 1 m (see figure 2.4). To reach the high precision needed in the measurement of the verteces positions, the silicon detectors must be as near a possible to the beam line. Each module, that provides a measurement of the r and ϕ coordinates, starts at 8 mm from the beam axes, which is less than the aperture of 3 cm required by LHC during the injection. Therefore the modules must be retractable. It is adopted a Roman pot system, which is able to move the two detector halves away from the beam during beam injection, avoiding radiation damage. The modules don't surround the beam pipe but are mounted in a vessel that maintains vacuum around the sensors and is separated from the machine vacuum by a thin corrugated aluminium sheet. With this trick the material traversed by the particles before reaching the detector is minimized, thus reducing multiple scattering, and is guaranteed a certain superimposition between the two halves.

Each module is a disc with internal (external) radius of 8 mm (42 mm). It is composed of two single-sided, n–on–n silicon strips detector with a thickness of 220 μm: one side measures the r coordinate, the other the azimuthal ϕ coordinate (see figure 2.5). The information about the impact parameters of the particles is used in the Trigger. The r and ϕ sensors are bonded back-to-back. The r sensors are composed of concentric semi-circles strips subdivided into four 45° regions, with the centre on the beam axes. The radius of the strips change gradually from 40 μm near the centre to 101.6 μm near the outer radius. This guarantees a uniform

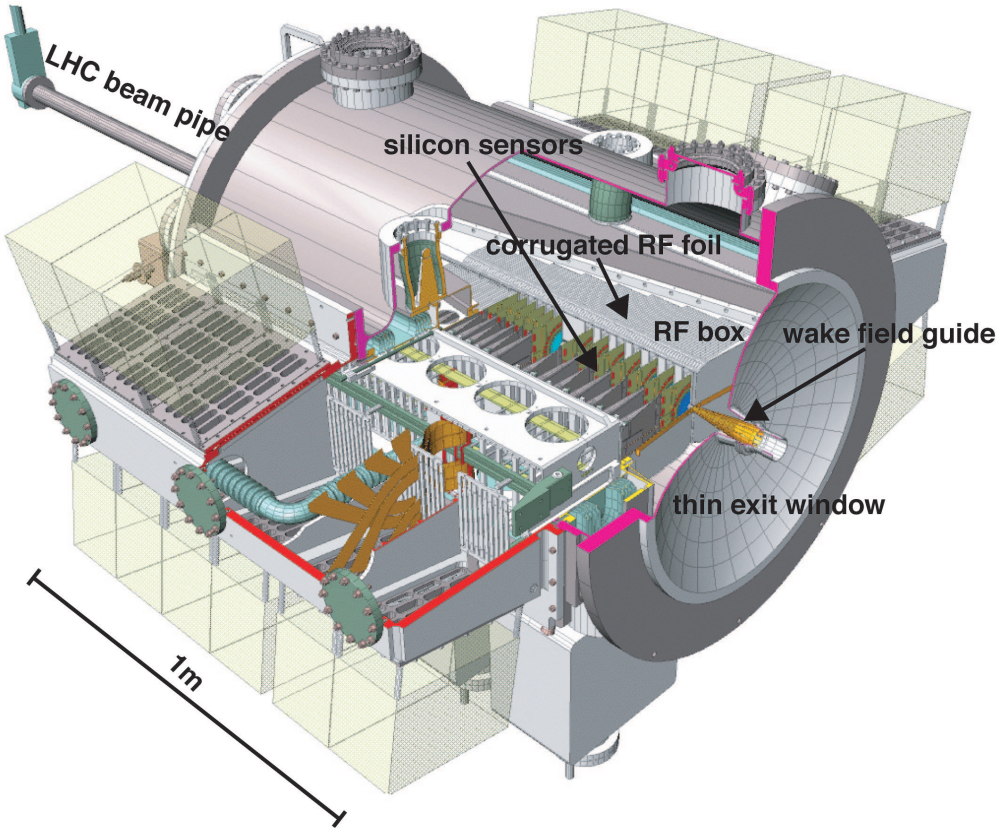

Figure 2.4: The VELO vacuum vessel with the silicon sensors, silicon sensor stations, the corrugated RF foils, the RF box, the wakefield guides and the thin exit window.

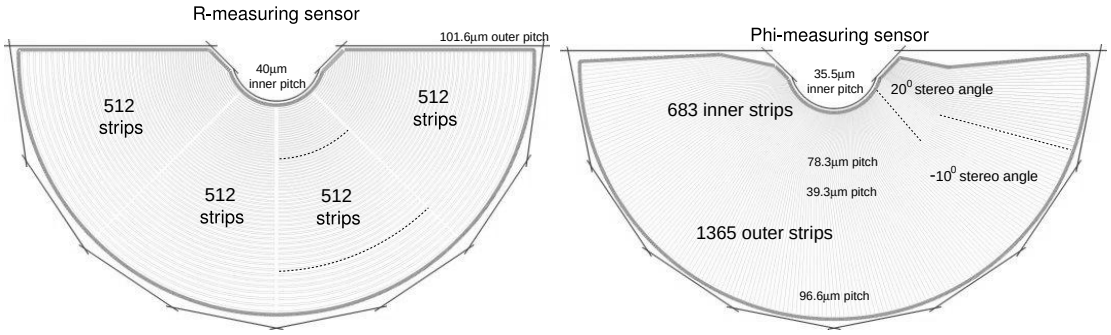

Figure 2.5: Sensor lay–out. Some strips are plotted with dotted lines for illustration.

resolution on the measurement of r. The ϕ sensors are composed of radial vertical strips. An inner set of strips, inclined of 10° with respect to the x (or y) axes

runs from $r = 8$ mm (with a pitch of 35.5 μm) to 17.25 mm (with a pitch of 78.3 μm). Another set of strips, inclined of 20° in the opposite versus, run from $r =$ 17.25 mm (with a pitch of 39.3 μm) to the end of the module (with a pitch of 78.3 μm). This guarantees a uniform resolution on the measurement of ϕ. The z direction is measured knowing in which module was found the hit, this require a great alignment precision.

The DAQ electronics of each sensor are placed on the outer most position of the wafers in order to avoid radiation damage of the acquisition chips and multiple scattering. The radiation damage is a big issue for the VELO, which is the detector traversed by the highest flow rate. To guarantee a minimal operative time of the VELO greater than 3 years it is necessary to maintain the temperature of the silicon detectors constantly lower than -5°C. This is done using a dedicated cooling system based on CO_2 and freon. The VELO main vacuum tank is bonded directly to the beam pipe.

The performances of the VELO are: reconstruction of tracks with angle from 15 up to 390 mrad, a spatial cluster resolution of about 4 μm for 100 mrad tracks and a primary vertex resolution of 42 μm in the z direction and 10 μm in the perpendicular plane. The displacement of production and decay verteces at LHCb is of $\mathcal{O}$(1 cm). The VELO precision on the decay length depends on the decay channel: it ranges from 220 μm to 370 μm.

The system is completed by a pile–up veto counter, composed by two VELO modules located in the downstream part, which are used by the Trigger to reject events with multiple primary verteces.

2.2.5 Magnet

To measure the momentum of charged particles a warm dipole magnet is used at LHCb[21][17]. Its aperture defines the acceptance of the experiment (±250 mrad vertically and ±300 mrad horizontally). The magnet is formed by two identical coils of conical saddle shape, placed mirror–symmetrically to each other in the magnet steel yoke. Each coil is made of aluminium, and consists of fifteen pancakes arranged in five triplets (see figure 2.6). The magnet design was dictated by the necessity of having a field as high as possible in the tracking stations and less than 2 mT in the RICH region. It gives a field vertically oriented and an integrated magnetic field of $\int Bdl = 4$ Tm for tracks of 10 m. This guarantees the desirable momentum resolution of 0.4% for particles with momenta up to

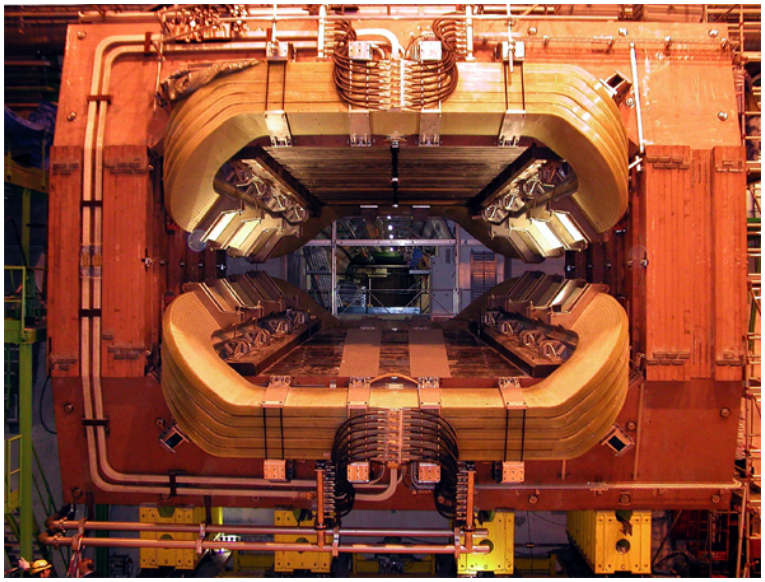

Figure 2.6: The LHCb magnet and surrounding steel yoke mounted in the point 8 cavern.

200 GeV/c. The magnet polarity can be easily and quickly changed, thanks to the non superconductive nature of the magnet, to reduce systematic errors in the CP violation measurements that could result from a left–right asymmetry of the detector. As the main issues with a warm magnet is of thermal origin, the conductor in coils has a central cooling channel of 25 mm diameter. Furthermore the pancakes of the coils are left free to slide upon their supports; only one coil is kept fixed where electrical and hydraulic terminations are located.

2.2.6 Tracking system

The aim of the tracking system is to reconstruct the tracks of the charged particles. From the curvature of a track in the magnetic field the momentum of the particle can be inferred. A $\delta p/p = 0.4\%$ or better is needed to achieve a good mass resolution at LHCb. One of the biggest contribution to the momentum resolution are the random change of track direction due to multiple scattering along the particle trajectory.

The tracking system is composed by the VELO detector (see section 2.2.4), the Trigger Tracker (TT) and the Tracking Stations T1–T3 which are divided into two components known as the Inner Tracker (IT) and the Outer Tracker (OT). Each

subsystem not only contributes to the measurement of the particle momentum but has also other functions. The VELO measures the positions of primary and secondary verteces (see section 2.2.4). The TT gives information like the transverse momentum (p_T) of charged particles to the trigger (hence the name). The T1–T3 stations match the tracks with the calorimeters depositions and the RICH rings, this links the energy and mass information to each track, so to each measured momentum.

Trigger Tracker (TT)

The Trigger Tracker (TT)[17][23] is a planar tracking station 150 cm wide and 130 cm high which covers the full LHCb acceptance. It is located downstream of the magnet, so it feels the same particles flux of the VELO. For this reason a silicon based technology is used. It is composed of silicon microstrip sensors with a strip pitch of about 200 μm which cover four detection layers. These are coupled in two pairs (TTa with x-u and TTb with v-x strip orientation) with a gap of 27 cm along the beam axis. The first and the last layers are covered with vertical strips. Strips rotated by a stereo angle of -5° and +5° compose the second and the third layer respectively (x-u-v-x orientation) (see figure 2.7). The 200 μm pitch of the silicon microstrips guarantees a single-hit resolution of about 50 μm. With this spatial resolution the momentum resolution of the spectrometer is dominated by multiple scattering, therefore the front-end readout electronics and mechanical support are located outside the LHCb acceptance. A read-out electronics with a fast shaping time (order of magnitude of 25 ns, the bunch crossing interval) is used, in order to avoid pile-up with a signal that comes from the previous event. The choice of silicon microstrips is also motivated by the ability to cope with high density of tracks and to be radiation-hard. In fact up to 5×10^{-2} particles per cm^2 per second are expected on the TT. It is maintained at 5°C or less with a dedicated cooling circuit to stand for 10 years at the expected radiation fluence.

The TT gives to the trigger a rough estimation of the particles transverse momentum. It is also fundamental in reconstructing the trajectories of the long–lived neutral particles that decay outside of the VELO fiducial volume and in reconstructing the tracks of the low momentum particles. These particles are deviated by the magnetic field outside the acceptance before reaching the T1–T3, so their momentum is reconstructed only from the VELO and TT information.

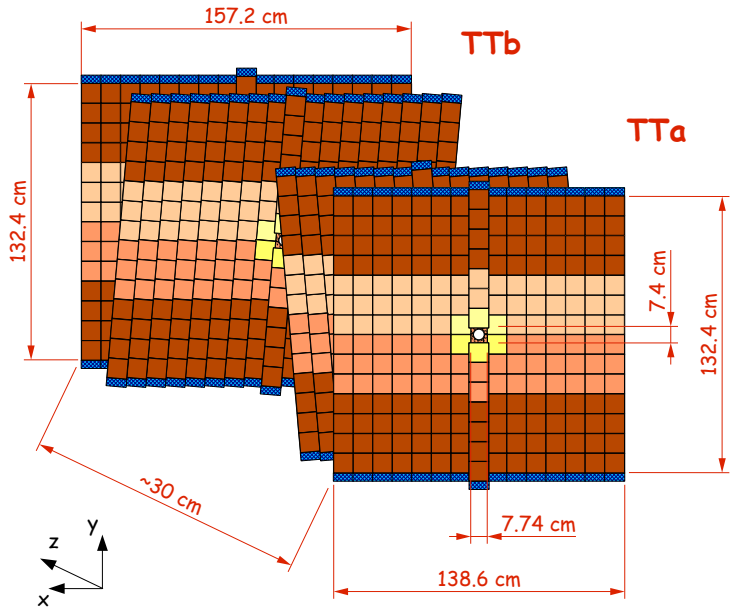

Figure 2.7: Lay–out of vertical read–out strips in TTa and TTb.

Tracking stations (T1–T3)

The three tracking stations[18] are located downstream the magnet at equal distances one to the other in the z direction. They measure the last part of the track of the charged particles, determining their momentum. They also give the directions of the particles which get into the RICH2. This directions are used in the algorithm which calculates the Cherenkov angle for each photon detected in the RICH2 HPDs, and therefore contribute to the particles identification. To cope with the variation of the particles flux with the polar angle the T1–T3 stations are divided into two parts: an Inner Tracker (IT)[23] covering the innermost region around the beam pipe, and an Outer Tracker (OT)[22] covering the outer part of the tracking stations.

The Inner Tracker region is expected to be crossed by a charged-particle density up to 1.5×10^{-2} per cm^2 per second, therefore their design must be constrained by the same consideration done for the TT, like the occupancy and the radiation damage. Each of the three IT stations consists of four individual detectors arranged above, below, left and right of the beam pipe (see figure 2.8); they cover a cross shaped area approximately 120 cm wide and 40 cm high. This geometry covers

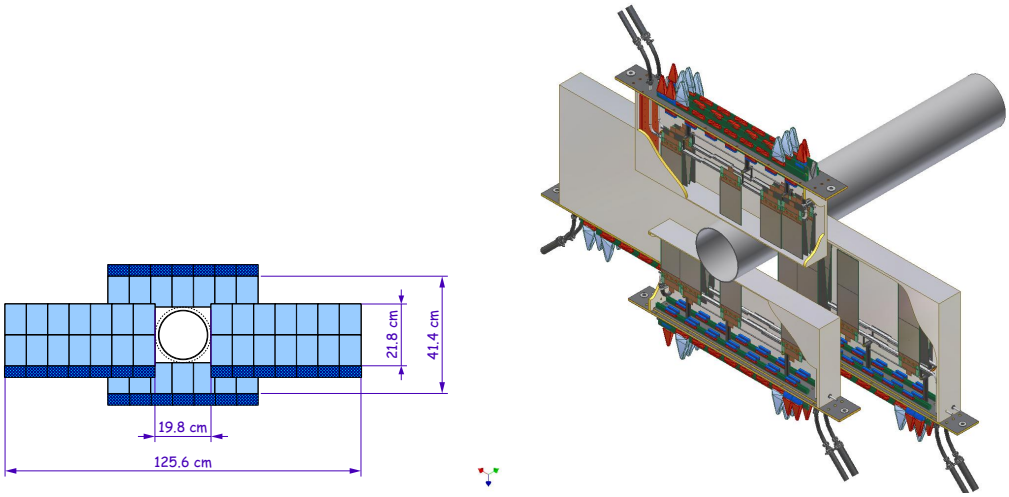

Figure 2.8: Left: layout of an IT x layer with the silicon sensors. Right: the four IT detector boxes arranged around the LHC beam pipe.

the highest particles density area of the T1–T3 plane. In fact, even though it represents the 1.3% of the tracking stations sensitive area, the IT is traversed by the 25% of the particles which reach the tracking stations. Detector modules in the boxes above and below the beampipe consist of a single silicon sensor 320 μm thick and a readout hybrid. Detector modules in the boxes to the left and right of the beampipe consist of two silicon sensors 410 μm thick and a readout hybrid (see figure 2.8). The strips thickness are chosen to maximize the signal to noise ratio and to reduce the material budget. Each IT station is composed of four planes of microstrips oriented in a x-u-v-x topology like the TT. The external planes have vertical strips, the internal are covered with strips rotated by an angle of +5° and -5° with respect to the x direction. The front end read out electronic is in the LHCb acceptance, so an extreme care was put to make it as radiation restistent and trasparent as possible. The silicon sensors and readout are maintained at a temperature of 5°C or lower to remove the heat due to the leakage current and ensure a ten year life for the T1–T3 stations.

The Outer Tracker is composed of four layer of modules for each station; each module contains two layers of drift tubes staggered in such a way that spatial resolution is less than the straw radius (see figure 2.9). The modules are arranged in an x-u-v-x geometry: the modules in the x layers are positioned vertically, while those in the u and v layers are tilted by $\pm 5°$ with respect to the x. Each straw tube

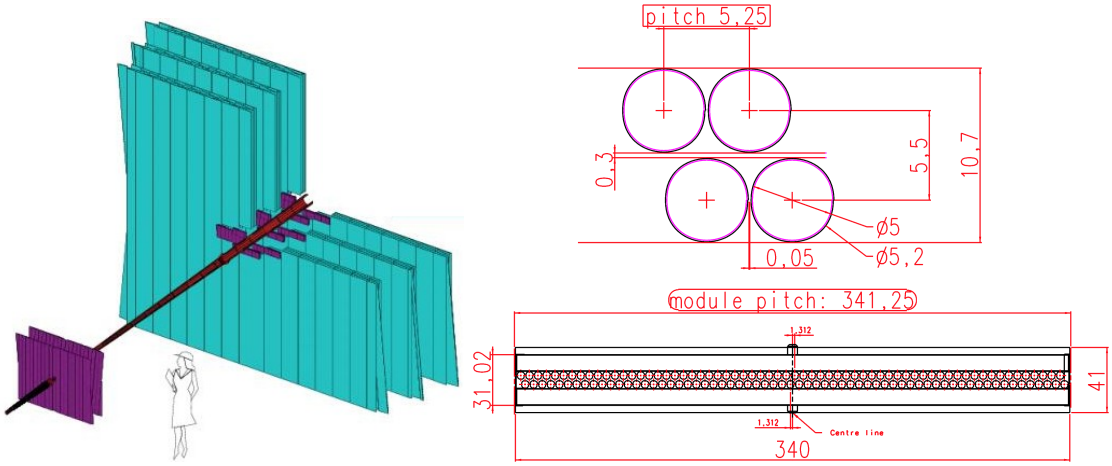

Figure 2.9: Left: 3D view of the tracking stations. Right: cross section of a straw-tubes module.

has an internal diameter of 4.9 mm and is filled with a gas mixture of Argon (70%) and CO_2 (30%). This guarantees a fast drift time (below 50 ns) and a good drift-coordinate resolution (200 μm). The straw tubes gives good spatial resolution, a fast response and they cover a wide area with a relatively low cost. The use of straw tubes is allowed by the fact that the OT covers an area of moderate particles density: it will see an occupancy up to 10%. The signal from each straw tube is collected by the front end read out electronics, which is positioned outside the LHCb acceptance. A special care was put in electrically shielding each straw tube from the front–end electronics and the ground. Problems of aging due to outgassing of a glue used in the assembly of the drift tubes have been seen, this should be cured by flushing CO_2 and heating the tubes, or increasing the high voltage for short periods.

2.2.7 RICH detectors

In order to reconstruct hadronic B mesons decays and to infer the B flavour at the production it is fundamental to distinguish between kaons and pions. To do that it is needed a particles identification detector, which works over a large momentum range. At LHCb this task is performed with two Ring Imaging CHerenkov (RICH)[17][18] [24] detectors which use different radiators to covers the full momentum range. One of them (the RICH1) is located between the VELO and the

TT, and cover the low momentum range. The other one (the RICH2) is located downstream the T3 station, and covers the high momentum range.

RICH1

The RICH1[25] is located upstream the magnet, and identifies the particles with low momentum: these will not reach the RICH2, because they are lost by the magnetic bending. The RICH1 covers the low to mid momentum charged particle range ∼1-60 GeV/c and covers the full LHCb acceptance (from ±25 to ± 300 mrad in the horizontal direction and from ±25 to ± 250 mrad in the vertical direction). It has two radiators: silica aerogel (n = 1.03) to cover the low momentum range and C_4F_{10} (n = 1.0014) to cover the mid momentum range. They both provide 3σ $\pi - K$ separation for the whole dynamic range. The Cherenkov light is focused using spherical mirrors. Flat mirrors guide the image towards the Hybrid Photon Detectors (HPDs), which are positioned out of the acceptance (see figure 2.10). The HPDs detect the Cherenkov photons in the wavelength range 200-600 nm and measure their spatial position. They are surrounded by external iron shields and are placed in MuMetal cylinders to permit operation in magnetic fields up to 50 mT. Once the position of the photon is detected and the direction of the track is known, the Cherenkov angle can be calculated using the algorithm discussed in [27].

RICH2

The RICH2[26] is located downstream the magnet, so it can only identify the particles with mid to high momentum that aren't lost by the magnetic bending. This kind of particles are mainly at low polar angles. For this reason the RICH2 acceptance is from ±15 to ± 120 mrad in the horizontal direction and from ±15 to ± 100 mrad in the vertical direction. It covers the high to mid momentum charged particle range, ∼15-100 GeV/c and beyond, using CF_4 radiator (n = 1.0005). RICH2 has a horizontal optical layout symmetry which is the replica of the RICH1 optic and photon detecting system (see figure 2.10).

2.2.8 Calorimeters

The Calorimeters have the task of measuring the energy and position of electrons, photons and hadrons, by their complete absorption. They also provide particles

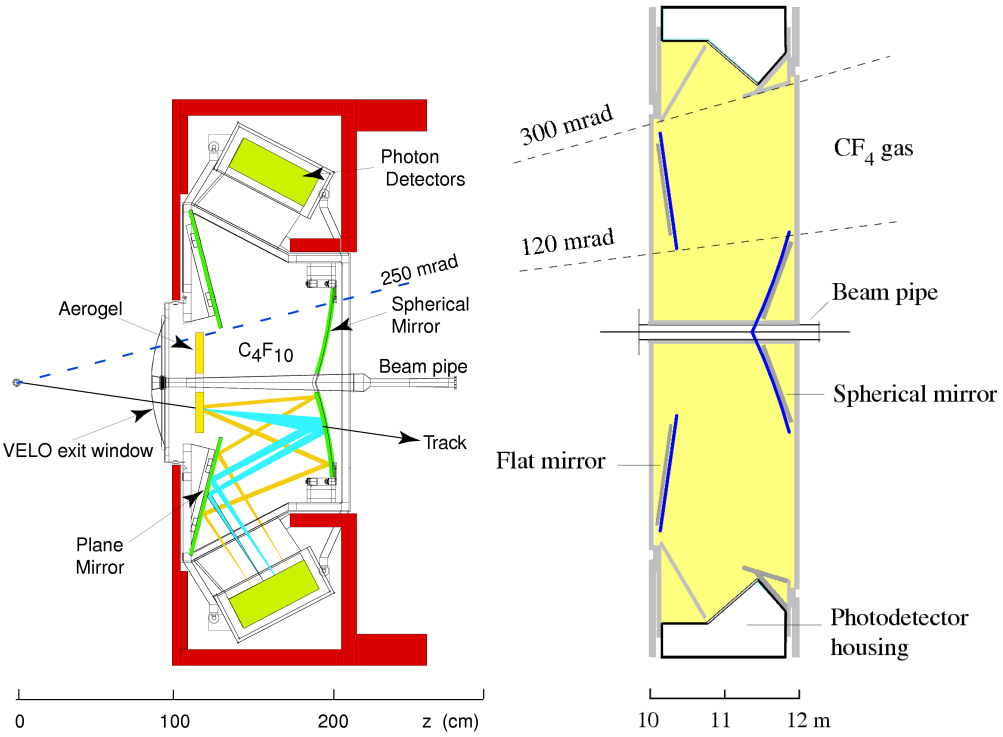

Figure 2.10: Left: side view schematic of the RICH 1 detector. Right: top view schematic layout of the RICH 2 detector.

identification. At LHCb the calorimeters provide a very fast estimation of the transverse energy of electrons, photons and hadrons to the Level-0 (L0) trigger. The LHCb calorimeter system consists of a Scintillating Pad Detector (SPD), a PreShower detector (PS), an Electromagnetic CALorimeter (ECAL) and Hadronic CALorimeter (HCAL).

Scintillating Pad Detector (SPD) and Preshower detector (PS)

The Scintillating Pad Detector (SPD) and the PreShower detector (PS)[17][28] are positioned just in front of the Electromagnetic CALorimeter (ECAL), with a thin 12 mm (2.5 X_0) thick lead converter placed in between. The SPD is used to select charged particles and to distinguish between electrons and photons. Charged particles deposit energy in the scintillator material, while neutrals particles practically do not interact in it. Thus the ECAL showers associated to hits in the SPD come from electrons, the ones without hits in the SPD come from photons. This is used

to separate photons and electrons at the Level 0 of the ECAL trigger.

The Preshower (PS) performs the discrimination between electrons and hadrons, using the difference in interaction lengths in lead: electrons produce a shower that starts in the lead absorber whereas pions do not. Electrons deposit in the scintillator more energy than the pions.

Both SPD and SP are formed of rectangular scintillator pads 15 mm thick. They are subdivided into three lateral concentric sections (see figure 2.11) with different granularity to obtain the same occupancy along the detector. This segmentation is the projection of the ECAL segmentation. In each pad the light is collected by a wavelength shifting (WLS) fiber and collected in 64 channel multianode photomultiplier tubes (MAPMT) located outside the acceptance.

Electromagnetic Calorimeter (ECAL)

The ECAL[17][28] is a sampling calorimeter built with a "shashlik" technology. It's composed of 66 alternating layers of 2 mm thick lead sheets and 4 mm thick scintillator plates, with a total radiation thickness of 25 X_0. Each scintillator plate is read out with a plastic WLS fiber, and the fibers corresponding to an individual ECAL module are grouped together in bunches and read by a single PhotoMultiplier Tube (PMT). The ECAL energy resolution is $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, where the energy is measured in GeV and $\oplus$ means the addition in quadrature. The calorimeter is subdivided into inner, middle and outer sections which differ each other by the density of cells: higher in the inner section, lower in the outer one (see figure 2.11). The density of cells was chosen in such a way that the granularity is high enough to separate two individual showers. This subdetector is used to perform the identification of electrons and photons, and to measure their energy.

Hadronic Calorimeter (HCAL)

The Hadronic CALorimeter (HCAL)[17][28] is used to identify the hadrons and to measure their energy. It is a sampling calorimeter made of 4 mm thick scintillator and 16 mm thick iron plates arranged parallel to the beam axis. This choice was made with the aim of improving the sampling of the more lateral hadronic shower. The overall detector thickness is 1.2 m representing an hadron interaction length of 5.6 λ_I (the upstream ECAL adds a further 1.2 λ_I). The light is collected

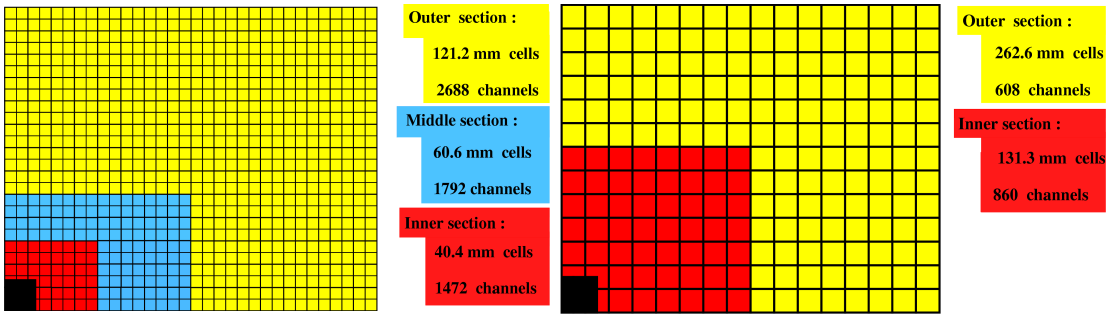

Figure 2.11: Left: lateral segmentation of the SPD, the PS and the ECAL. Right: lateral segmentation of the HCAL. In both figures, one quarter of the detector front face is shown.

by WLS fibres running along the edges of the tiles toward the back side, where photomultiplier tubes (PMTs) are housed. The fibers are grouped together in different sets and than read by a PMT. The HCAL is segmented in two sections of different tile size in order to cope with the different fluences (see figure 2.11). The HCAL energy resolution is $\sigma_E/E = 80\%/\sqrt{E} \oplus 10\%$, where the energy is measured in GeV.

2.2.9 Muon System

The muon identification and detection is fundamental at LHCb. Muons are produced in more than 10% of B decays; like the golden channel $B_d^0 \rightarrow J/\psi\ K_s$, in which the J/ψ goes in two muons. The muons are also used in the Level-0 (L0) trigger and to distinguish the flavour of B at the production (flavour tagging with opposite side muons). The muon identification and reconstruction at LHCb is provided by the Muon System[17][29][30][31]. It is composed by 5 stations (M1–M5) of rectangular shape, orthogonal to the beam axis (see figure 2.12). The M1 station is placed in front of the calorimeters. The stations M2–M5 are placed downstream the calorimeters and are spaced with iron shielding plates 80 cm thick. The M1 station improves the resolution of p_T: it measures the p_T of muons before their interaction with the calorimeters material that increase the multiple scattering. The iron absorbers between the M2–M5 stations select the muons, that are the only particles which can traverse them without being absorbed.

Each Muon Station is divided in four regions with increasing distance from the beam axis; this was done to cope with the variation of flux along the z direction

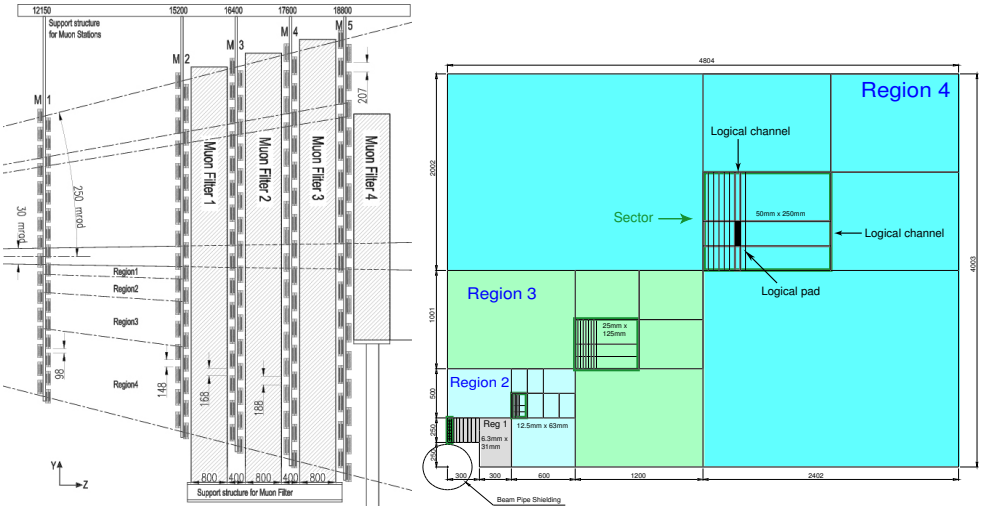

Figure 2.12: Left: side view of the muon system. Right: front view of one quadrant of a muon station showing the partitioning into sectors. In one sector of each region a horizontal and a vertical strip are shown.

(see figure 2.12). The geometry of the five stations is projective: all their transverse dimensions scale with the distance from the interaction point. In all the regions of the muon system Multi Wire Proportional Chambers (MWPC) detectors are used, except in the innermost region of the M1 station, where the fluence is too high and the MWPC have ageing problems. In this internal region Gas Electron Multiplier (GEM) detectors are used. The pad granularity differs in the 4 different regions: it increases towards the inner regions to cope with the variation of particles density. To achieve the required 95% efficiency in stations M2–M5 the MWPCs are composed of four 5 mm gas gaps arranged in two sensitive layers with independent readout. In station M1 the chambers have only two gas gaps to minimize the material in front of the electromagnetic calorimeter. The gas gap thickness and the gas mixture ($Ar/CO_2/CF_4(40:55:5)$) was chosen to obtain the fastest collection of charge possible. This is due to the L0 trigger request of a 5 ns time resolution for each gas gap. Stations M1–M3 have an good spatial resolution and they are used to calculate the muons p_T with a resolution of $\sim$20% and to reconstruct the muons track. The stations M2 and M3 have a poor spatial resolution and are used to identify the most penetrating particles.

2.2.10 Trigger

The bunch crossing frequency expected is 40 MHz. It's not possible to record events at this frequency, because such an enormous amount of data cannot be stored. Furthermore most of the bunch crossing produce events useless for the physical purposes of the LHCb experiment. At a luminosity of 2×10^{32} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ only about 10 MHz is the crossing frequency with interactions that produces at least two charged particles with hits in VELO and T1–T3 stations (these are called "visible interactions"). The visible interactions are expected to contain a rate of about 100 kHz in which a $b\bar{b}$–pair is produced. Only 15% of these events will contain a B meson with all its decay products inside the LHCb acceptance. Furthermore the Brancing Ratios (BRs) of the CP violating decays and rare decays are small (tipically less than 10^{-4}). For these reasons the 40 MHz bunch crossing frequency must be reduced by the trigger to a 2 kHz storage frequency. This reduction is achieved in two trigger levels: Level–0 (L0) and the High Level Trigger (HLT)[32][33]. First the signals are processed by the L0, then the events that pass the L0 are processed by the HLT. Finally the events that pass the HLT are stored.

Level–0 trigger

The L0 is implemented in hardware and works synchronously with the 40 MHz bunch crossing frequency. It has the purpose to reduce the acquisition rate from 40 to 1 MHz. 1 MHz is the frequency at which the whole spectrometer can be read out, therefore all the subdetectors can contribute to the High Level Trigger (HLT) decision. The L0 combines the output from the pile-up system, the calorimeters system and the muon system to decide if the event has a physics interest and must be sent to the HLT or not. The B mesons are the heaviest mesons produced at LHC, and heavy mesons decays predominantly in particles with high p_T and E_T. The trigger L0 looks for the highest E_T hadron, electron and photon clusters in the calorimeters and the two highest p_T muons in the muon chambers. In addition it analyses the pile-up system, estimating the number of primary pp interactions in the bunch crossing in order to reject multiple pp inelastic interactions. Furthermore the number of tracks in the event is estimated by the total energy detected in the calorimeters and the number of hits in the SPD, in such a way the event with too many particles (more than around 150) are rejected. For this type of events, which have an high probability to be selected by the trigger algorithm for their large

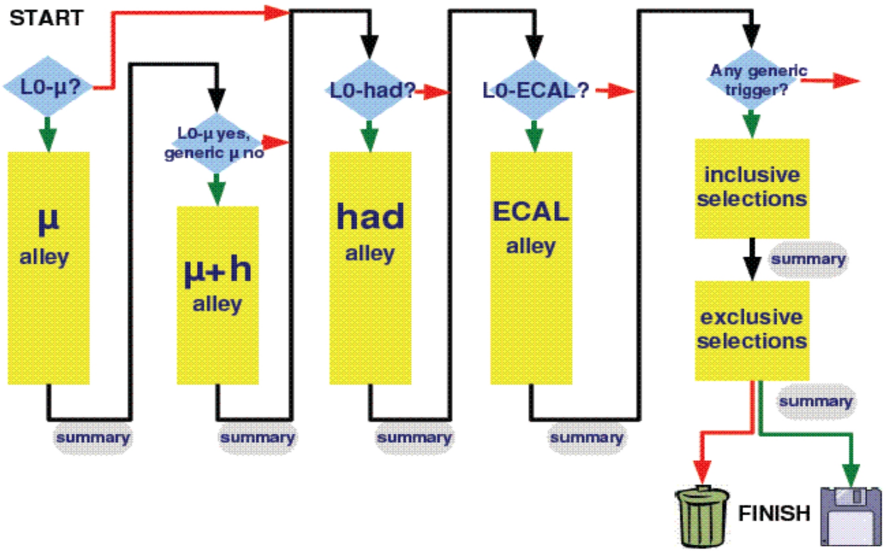

Figure 2.13: The High Level Trigger flow-chart.

combinatorics, the reconstruction would take too much time at the HLT level.

The L0 trigger is performed by dedicated electronics boards on the participating sub-detectors. Each sub-detector has a pipeline buffer 168 bunch–crossings long that has a latency time of 4.2 μs. This is the time between the pp interaction and the L0 trigger decision. If the time of flight and the delays due to the cables are considered, only 2 μs are left to the L0 Decision Unit (L0DU) to make the decision.

High Level Trigger

The High Level Trigger (HLT) is implemented software and works asynchronously on a processor farm. It has the purpose to reduce the acquisition rate from 1 MHz to 2kHz, making use of the full event data from all subdetectors. It consists of a C++ application which runs on over 2000 CPU located in the Event Filter Farm (EFF). It is separated in four independent trigger streams, the so called "alleys". It exists one alley for each trigger type of L0: the muon alley, the muon and hadron alley, the hadron alley and the electromagnetic alley (see figure 2.13). The events selected by an L0 trigger type are processed by the respective alley. In the alley the event information is refined and new information, like the impact parameters, are added. First the L0 selection is validated adding the tracking system information. If the event is selected, a selection specific for each alley is applied using more

information and requesting tracks with high p_T and high impact parameter. If an event pass at least one alley selection, it is processed by the inclusive trigger algorithms where specific resonances are selected. The events are then processed by the exclusive trigger algorithms which looks for specific final states. The final trigger decision is a logical OR of the inclusive and exclusive selections. The advantage of this scheme is that most of the events (∼85%) are selected due only to one L0 candidate, and hence they will only be processed by one alley. The events selected are mainly of physics interest, but a large fraction of the output bandwidth is devoted to calibration and monitoring purposes.

Chapter 3

$B_d^0 \to J/\psi K_S^0$ decay mode and control channel

In this chapter the B_d^0 decay mode that will be used at LHCb to measure the $\sin 2\beta$ parameter is described. In the first paragraph the main features of this channel are shown. In the second paragraph it will be explained why this channel was chosen to measure $\sin 2\beta$. In the third paragraph it will be explained why a measurement of this parameter is so important at LHCb. In the last two paragraphs the strategy of the measurement will be described.

The tree diagram for the decay $B_d^0 \to J/\psi K_{S,L}^0$ is showed in figure 3.1(a). We can notice that the $B_d^0 = \bar{b}d$ can decay only in $K^0 = \bar{s}d$ and the $\overline{B_d^0} = b\bar{d}$ can decay only in $\overline{K^0} = s\bar{d}$. The decay amplitude of the tree level diagram is proportional to $V_{cb}^* V_{cs} \backsimeq A\lambda^2$. Beside tree diagrams there are also gluonic penguin diagrams. These diagrams usually complicate the analysis, since the hadronic matrix elements are non–perturbative at low energy and the QCD constants are not experimentally known at the desired precisions. They not only contribute to the uncertainty of the parameter, but also produce errors in the interpretation of the CP asymmetries and the extraction of the parameters from them. The most important penguin diagram for this decay is showed in figure 3.1(b). The b quark decays in a gluon and a s quark; then the gluon goes in a $c\bar{c}$ pair. Its contribution is proportional to:

$$P \sim V_{us}V_{ub}^* f(m_u) + V_{cs}V_{cb}^* f(m_c) + V_{ts}V_{tb}^* f(m_t) \quad , \tag{3.1}$$

where $f(m_i)$ describes the contribution of the quark loop, and m_i is the mass of the i quark in the loop. Since in the Standard Model (SM) the CKM matrix is

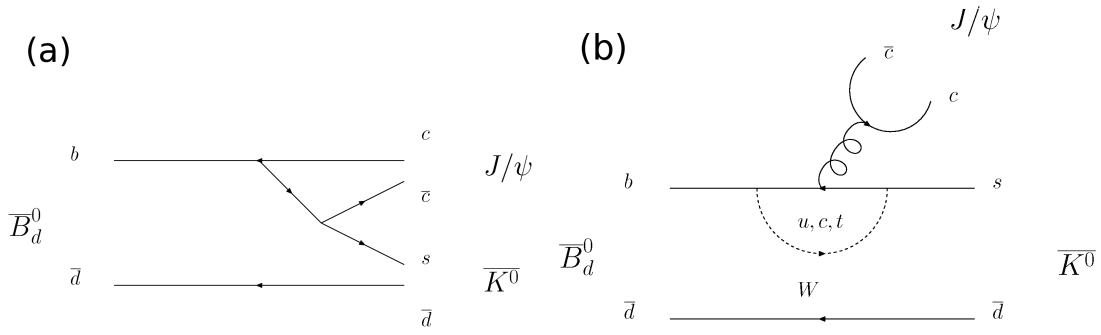

Figure 3.1: Feynman diagrams for $B_d^0 \to J/\psi K_{S,L}^0$.

unitary, we have $\sum_i V_{is}V_{ib}^* = 0$. Thus $f(m_t)\sum_i V_{is}V_{ib}^* = 0$ can be subtracted in expression 3.1, and we have:

$$P \sim V_{us}V_{ub}^*(f(m_u) - f(m_t)) + V_{cs}V_{cb}^*(f(m_c) - f(m_t)) \quad . \tag{3.2}$$

The last term in 3.2 has the same CKM coefficients as the tree amplitude, while the first term has a CKM factor $V_{us}V_{ub}^* \sim \lambda^4(\rho - i\eta)A$, which is two orders of magnitude smaller.

3.1 Physical description of the Decay

Since the penguin amplitude is smaller than the tree one, it can be neglected. Considering only the tree diagram, the weak amplitude of this channel can be calculated with a precision of 1%[34]. This is the reason why this mode is called the *gold–plated mode* or also the *golden channel*; I'll refer to this as the golden channel in the text.

3.2 CP violation in $B_d^0 \to J/\psi K_{S,L}^0$

The golden channel presents several characteristics that simplify the theoretical description of the CP violation in it. The final state is a CP eigenstate, thus both B^0 and $\overline{B^0}$ decay into it. In the following argument we will neglect the rather small CP violation in the mixing of the K^0 which gives that the $K_{S(L)}$ is not a perfect CP eigenstate. In the B_d system the $\Delta\Gamma = \Gamma_1 - \Gamma_2$ is small ($\Delta\Gamma_{B_d}/\Gamma_{B_d} \approx 0.003$) and in first approximation, it can be put equal to zero; thus: $\Gamma_1 = \Gamma_2 = \Gamma$. Substituting

1.34, 1.35 and $\Gamma_1 = \Gamma_2 = \Gamma$ in 1.36 and 1.37 we get the decay rate as a function of time:

$$\Gamma_{B^0\to f}(t) = e^{-\Gamma t}\frac{|\ A_f\ |^2}{2}\left[\left(1+|\ \lambda\ |^2\right) + \left(1-|\ \lambda\ |^2\right)\cos\left(\Delta m t\right) - 2Im\left\{\lambda\right\}\sin(\Delta m t)\right]$$

$$\Gamma_{\bar{B}^0\to f}(t) = e^{-\Gamma t}\frac{|\ \bar{A}_f\ |^2}{2\ |\ \lambda\ |^2}\left[\left(1+|\ \lambda\ |^2\right) - \left(1-|\ \lambda\ |^2\right)\cos\left(\Delta m t\right) + 2Im\left\{\lambda\right\}\sin(\Delta m t)\right] \tag{3.3}$$

where

$$\lambda \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}\ . \tag{3.4}$$

The CP violation in the mixing for the B_d^0 is expected to be very small: $|q/p| - 1 \sim 10^{-4}$[34]. Taking $|q/p| = 1$, $\frac{|\bar{A}_f|^2}{|\lambda|^2} = |\ A_f\ |^2$ and we can define the time dependent *asymmetry* as:

$$a(t) \equiv \frac{\Gamma_{\bar{B}^0\to f} - \Gamma_{B^0\to f}}{\Gamma_{\bar{B}^0\to f} + \Gamma_{B^0\to f}} = -\frac{(1-|\ \lambda\ |^2)}{(1+|\ \lambda\ |^2)}\cos(\Delta m t) + \frac{2Im\lambda}{(1+|\ \lambda\ |^2)}\sin(\Delta m t)\ . \tag{3.5}$$

The term $\frac{(1-|\ \lambda\ |^2)}{(1+|\ \lambda\ |^2)}\cos(\Delta m t)$ in 3.5 is non–zero only when there is CP violation in the decay (see section1.4.2 equation 1.38), which occurs only when more than one amplitude contribute to the decay. In the specific case of the golden channel there is only one relevant decay amplitude, then $|\ \lambda\ | = 1$, but $Im\lambda$ could still be non–zero if CP is violated in the interference. Substituting $|\ \lambda\ | = 1$ in 3.5 we obtain:

$$a(t) \equiv Im\lambda\sin(\Delta m t)\ . \tag{3.6}$$

A_f is the amplitude of B^0 into a given final state f. Since a $K_{S(L)}$ is in the final state, the overall amplitude will involve the amplitudes into the flavour eigenstates and the transformation from the flavour to the mass eigenstates. This is due to the fact that it was chosen the mass eigenstates to develop the calculation. Making the reciprocal of equation 1.31 we have:

$$< K_S\ | = \frac{1}{p_K} < K^0\ | - \frac{1}{q_K} < \overline{K^0}\ |\ ,$$

then, remembering that a $B_d^0(\overline{B_d^0})$ goes only in a $K^0(\overline{K^0})$ we have:

$$A_f = < J/\psi K_S\ |\ T\ |\ B_d^0 > = \frac{1}{p_K} < J/\psi K_0\ |\ T\ |\ B_d^0 > \tag{3.7}$$

and

$$\overline{A}_f =< J/\psi K_S \mid T \mid \overline{B_d^0} >= -\frac{1}{q_K} < J/\psi \overline{K_0} \mid T \mid \overline{B_d^0} > \ , \tag{3.8}$$

then

$$\frac{\overline{A}_{J/\psi K_S}}{A_{J/\psi K_S}} = -\left(\frac{p}{q}\right)_K \frac{\overline{A}_{J/\psi \overline{K_0}}}{A_{J/\psi K_0}} \ ; \tag{3.9}$$

The same equation holds for the K_L, with a plus sign on the right. The CP parity eigenvalues are calculated multiplying the C and P eigenvalues. The J/ψ, which is a vector meson ($l = 0, s = 1$), has C=$(-1)^{l+s} = -1$, P=$(-1)(+1) = -1$ and CP=+1. On the contrary B_d^0 is a pseudoscalar meson ($l = 0, s = 0$), thus has CP=−1. Since K_S has CP=+1 for construction, then the J/ψ and K_S in the final state must have relative angular momentum l=1 (alternatively we can say that they are produced in p–wave). This means that there is an additional minus sign when the CP operator is applied to the final state. This sign cancels the minus in 3.9. We can now write λ as:

$$\lambda = \eta_{S(L)} \left(\frac{p}{q}\right)_{B_d} \left(\frac{p}{q}\right)_K \frac{\overline{A}_{J/\psi \overline{K_0}}}{A_{J/\psi K_0}} \ . \tag{3.10}$$

where $\eta_{S(L)}$ is $+(-)1$. Rewriting the 3.10 in terms of the CKM matrix elements we have:

$$\lambda(J/\Psi K_{S(L)}) = -(+)\frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}} \frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}} \frac{V_{cd}^*V_{cs}}{V_{cd}V_{cs}^*} = -(+)\left(\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right)^* \frac{V_{tb}^*V_{td}}{V_{cd}V_{cb}^*} \ . \tag{3.11}$$

From 1.29 we obtain:

$$\lambda(J/\Psi K_{S(L)}) = -(+)\left(|r|e^{i\beta}\right)^* \left(|r|e^{i\beta}\right)^{-1} = -(+)e^{-i2\beta} \ , \tag{3.12}$$

where $|r| = \left|\frac{V_{tb}^*V_{td}}{V_{cd}V_{cb}^*}\right|$. Substituting 3.12 in 3.6 we finally have:

$$a(t)\Big|_{J/\Psi K_{S(L)}} = (-)\sin(2\beta)\sin(\Delta m t) \ . \tag{3.13}$$

So, for the Standard Model, the measurement of the asymmetry as a function of time in the golden channel gives directly the parameter $\sin(2\beta)$, which is nonzero if and only if CP is violated. If we include the CP violation in the K system we obtain $a(t)_{S(L)} = (-)\sin(2\beta - Re(\varepsilon'))\sin(\Delta m t)$. Since 2β is $\mathcal{O}(10^{-1})$ and $Re(\varepsilon')$ is $\mathcal{O}(10^{-6})$, this justifies the choice of neglecting the CP violation in the K system.

3.3 Experimental motivations

In the golden channel, 50% of K_S^0 and 50% of K_L^0 are produced. The mean life of a K_S^0 and a K_L^0 are respectively 0.89×10^{-10} s and 5.1×10^{-8} s. At their average momentum of 36 GeV, they will travel about 2 m and 1000 m respectively. It means that only the decay products of the K_S^0 can be detected in the tracking system; while the K_L^0 is detected only in the hadronic calorimeter. Thus for the K_S^0 can be reconstructed the invariant mass and the decay vertex. Instead the K_L^0 is detected as a neutral hadron, with a rather poor information on its direction, no information about its mass and no secondary vertex. Therefore a bad reconstruction efficiency is expected for the K_L^0. This leads to consider only final states with K_S^0.

The $\sin 2\beta$ was measured for the first time with the two detectors BaBar and Belle. They are two 4π experiments which work respectively at PEP–II (in the US) and KEK–B (in Japan). These are two e^+e^- machines operating at the $\Upsilon(4s)$ resonance. They have an e^+ beam of $\sim$3 GeV and an e^- beam of $\sim$8 GeV. This gives to the centre of mass of the collision a boost, which separates the decay positions of the two B mesons. The B–factories exploit the correlation between the flavour of the two b quarks produced at the $\Upsilon(4s)$ resonance: the oscillation of one B meson starts only once the opposite B is decayed. This allows to measure the flavour of B at the production with a great precision, detecting the flavour of the opposite B (this is called "flavour tagging", see section 7.1). The B–factories have also the advantage of having a high luminosity ($\sim 10^{33}$ cm^{-2}s^{-1}). Due to the fact that they work at a $b\bar{b}$ resonance in an e^+–e^- machine, and that they can measure also the missing energy, they have a very low background. For all these reasons the B–factories performed the most precise measurement of $\sin 2\beta$. The present world average is[34]:

$$\sin 2\beta = 0.673 \pm 0.028 \quad . \tag{3.14}$$

This value is incompatible with zero by 5σ. From equation 3.14 we obtain four possible solutions for β, which are located in the first and the second quadrant of the $(\overline{\rho}, \overline{\eta})$ plane. Both are shown in figure 3.2, together with other measurements of the unitarity triangle parameters. The ambiguity goes from four to two when the $\cos 2\beta$ measurement is considered[36][37][38]. It can be seen in figure 3.2 that all the different measurements intersect in the same area, delimited with a black line. The fact that all the measurements agree means that they all coincide with the Standard Model expectations.

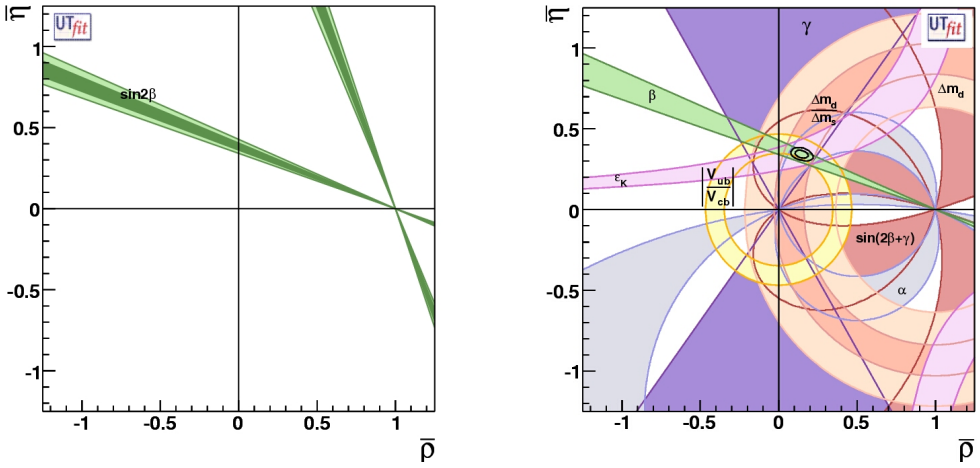

Figure 3.2: Left: $\sin 2\beta$ bound from $B_0 \to J/\psi K_0$ shown on the (ρ, η) plane. Right: Allowed regions for (ρ, η). The closed contours at 68% and 95% probability are shown. The full lines correspond to 95% probability regions for the constraints given by the measurements of $|V_{ub}|/|V_{cb}|$, ε_K, Δm_d, $\Delta m_d/\Delta m_s$, α, β, γ, $\Delta\Gamma_d/\Gamma_d$, $\Delta\Gamma_s/\Gamma_s$, A_{SL}^d, and the dimuon asymmetry.

The statistics is a key issue in the $\sin 2\beta$ measurement. At LHCb, due to the large luminosity and the high $b\bar{b}$ cross section, a huge amount of $b\bar{b}$ is expected. This large statistics can compensate the flavour tagging and background level, worse than at the B–factories. In the first year of data taking at LHCb, which is expected to start in summer 2008, an integrated luminosity of 2fb^{-1} is expected. This luminosity will allow a first significant measurement of $\sin 2\beta$. This will probably be the first measurement of a CP violation parameter at LHCb. Its main purpose will be to verify the power of LHCb to measure the others, more challenging, parameters. It can be estimated the LHCb sensitivity to measure other parameters of the unitarity triangle. Important information can come from a comparison between the $\sin 2\beta$ measured in the first year at LHCb and the current $\sin 2\beta$ result. A discrepancy can be a sign of systematic errors, related for example to a misalignment or a bad calibration of a subdetector. The systematic errors can't be predicted from the detector simulation and must be understood in the first periods of running. A possible way to look for this type of uncertainties is to compare some physical value measured at LHCb (for example the B lifetime) with the one found by other experiments. Once the systematics are understood, any

other differences between the measurements performed by LHCb and the expected values can be indications of new physics. Even if the golden channel isn't one of the most suitable to show new physics, there can be new particles which enter in the loop of the penguin diagram, changing the CP violation behaviour of the channel.

The aim of this study is to define a selection for the golden channel, to estimate the signal and background yield and to perform a fit of $\sin 2\beta$ calculating the error expected after one year of data taking.

3.4 Measurement strategy and control channel

In this section the strategy adopted to measure $\sin 2\beta$ in the channel $B_d^0 \to J/\psi K_S$ is explained. First of all the final state must be exclusively selected. The J/ψ is a narrow resonance which decays immediately after its production, and the K_S decays after many centimeters.

The J/ψ decay modes with the highest Branching Ratio (BR) are $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$; they have both a BR = 5.9%. The other modes with a BR of the order of percent involve a neutral particle, a π^0 or a γ. The leptonic decay shows the highest BR and reconstruction efficiency. The difference between the two leptonic modes is in the reconstruction efficiency: the muons are better identified and reconstructed particles. Their detection is performed by the muon chambers (see section 2.2.9). The electrons are less well reconstructed in the ECAL and can be misidentified with photons; besides the background electrons are more than the muons. Thus it was choosen to reconstruct only the $J/\psi \to \mu^+\mu^-$ decay mode.

The K_S decays in the 99% of cases in two pions: 1/3 of time in $\pi^0\pi^0$, 2/3 in $\pi^+\pi^-$. The reconstruction efficiency of a charged pion is much higher than a neutral one. A charged pion leaves a detectable track, from which the momentum is measured; they are also identified in the RICHes and absorbed in the HCAL. On the contrary the $\pi^0 \to \gamma\gamma$ are only detected in the ECAL, and they are identified with lower efficiency. For these reasons only the $K_S \to \pi^+\pi^-$ will be selected.

To calculate the asymmetry (equation 3.13) and extract $\sin 2\beta$, the lifetime and the flavour of B at the production must be known. Once the final state has been reconstructed we know the momentum and the flight distance of the reconstructed B. With these two quantities the lifetime can be obtained. To know if a B or a $\overline{B}$ is produced, the flavour tagging algorithms are used (see section 7.1). These

algorithms have a significant probability to make the wrong estimation of flavour, this is the so called *mistag probability* or *mistag fraction* (ω_{tag}). The mistag fraction is defined as the fraction of the wrong tagged B over all the tagged B. As it will be demonstrated in section 7.1, the mistag fraction enters in the measured asymmetry in such a way that the measured parameter is no more $\sin 2\beta$ but $(1-2\omega_{tag})\sin 2\beta$. Its value can be obtained only knowing the flavour of B at the moment of the decay. This is not possible in the case of $J/\psi K_S^0$, because both the B^0 and the $\overline{B^0}$ have the same final state. The solution is to measure the mistag fraction in an alternative decay channel, usually called the *control channel*. This channel must be as similar as possible to the golden channel, but mast have a final state that is not a CP eigenstate. The best candidates are: $B_d^0 \to J/\psi K^*$ and $B_u^0 \to J/\psi K^+$ with the respective charge conjugates. The mistag of the signal and control channels should be as similar as possible. In fact, as $\sin 2\beta$ depends linearly on the value of the mistag, a small difference introduces a systematic error.

3.5 The unified and unbiased selection

The control channel chosen is kinematically very similar to the golden channel. This guarantees that the overall value of the mistag is the same. If the mistag depends on a kinematic variable, we should proof that the distribution of this variable is the same for both channels. Otherwise the mistag measured in the control channel must be corrected to be used for the golden channel. The estimation of this correction from real data is extremely difficult. Indeed, to calculate the mistag with an acceptable error, a large amount of data is needed. This is due to the low tagging efficiency and to the difficulty of fitting the mistag from the oscillation pattern in case of B^0 channels. For example if we calculate the mistag as a function of a kinematic variable, we must divide the phase space in intervals of this variable and calculate the mistag in each interval. Of course, the more are the intervals and the less is the number of events for each interval. Thus a determination of the mistag as a function of a kinematic variable leads to a large error in the determination of the mistag itself. This is the reason why we tried to perform a selection which preserve the kinematical similarity of the channels. This selection will be called in the text the *unified selection*. It does not mean that the selection cuts are exactly the same for all the channels, it means only that particular care is reserved to the dependence of the mistag with respect to some kinematic variables.

If this dependence exists, the selection cuts are chosen in such a way that the final distributions for these variables are the same.

For both the $\sin 2\beta$ and the mistag determination, a time dependent asymmetry must be measured. In the selection there can be some cuts on kinematic variables which eliminate preferentially signal events with low B lifetime, or vice versa. This situation introduces a bias in the lifetime distribution of signal events; this bias must be taken into account in the $\sin 2\beta$ and mistag fits. To cure the bias one must know the so called acceptance function, this is just the selection efficiency as a function of the decay time. For example, if a cut is applied, which eliminates predominantly events with short B lifetime, the result will be a lower efficiency at low lifetime. It is extremely difficult to know the acceptance function of the signal events from real data. The problem with real data is that we know only the time distribution of the events after the selection, so we know only the biased distribution. To know the acceptance function also the unbiased distribution is needed. Moreover, with real data, we can't distinguish between signal and background. There are two possible solutions: one is to use an independent sample of data selected with and without the bias cuts, and then to calculate the acceptance function. This method suffer of large uncertainties due to the impossibility of distinguishing between signal and background. The other method consists in calculating the acceptance function from simulated data. Of course the simulated data could not correspond to the real one. So, both the methods can introduce systematic errors in the mistag and the $\sin 2\beta$ estimation. To avoid this problem it was chosen to tune a selection which do not bias the lifetime distribution of B for both the golden channel and the control channel. This kind of selection will be called *the unbiased selection* in the text. To perform an unbiased selection, for each cut applied, it was checked that the efficiency does not depend on the lifetime of B.

My work was to develop the selection of the golden channel and the $\sin 2\beta$ fit. I also took part in the selection of the $B_d^0 \to J/\psi K^*$, which will be presented in chapter 6. In this thesis I won't speak about the $B_u^0 \to J/\psi K^+$ control channel, I'll only briefly show its use in chapter 7.

Chapter 4

Simulation and the LHCb software

The work of selection and fitting was performed on simulated data produced in 2006. In the first paragraph of this chapter the simulation is described.

The simulated data, stored in a particular database, had to be manipulated and saved locally, to perform the physical studies. To complete this task it is necessary not only to know how the simulation worked, but also to learn the LHCb software. As part of my thesis work I run on the stored data, I selected them and saved the most important information locally, in a format which permits successive manipulations. These works have taken long time, not only for the huge amount of information that had to be processed, but also for the complexity of the software that must be managed. This software involves several different programs that have to interface together. All these programs are in a developing stage, therefore change continuously, and bugs are often pointed out by users. For these reasons I dedicate part of the current chapter to describe the LHCb software.

4.1 LHCb simulation

The real data from LHCb will not be available until the summer of 2008, start up of the LHC machine. Until then the physicists need simulated data to do things like: studying the detector behaviour, optimizing the selections for the physics channels, setting up all kinds of software that will be used for real data. To satisfy these requests it was performed a complete simulation of both the LHCb detector and the pp inelastic interactions.

The simulation is performed by a software package[35] called Gauss (see figure

4.1). It first generates events of pp collisions as if they come from an LHC bunch crossing, producing $b\bar{b}$, then simulates the decays of the B-mesons in channels of interest for the LHCb physics program. It uses the software Pythia[39] for the event production, and the package EvtGen[40] which is specialized in accurately simulating B decays.

After the generation phase, Gauss traces in the LHCb detector the generated particles. The simulation of the physics processes which the particles undergo when travelling through the experimental setup, is delegated to the Geant4 toolkit. The end of this phase is the output of Gauss, that consists in hits produced in the subdetectors as well as the Monte Carlo (MC) truth history.

The final stage of the LHCb detectors simulation is performed by the "Boole" digitization program. It simulates the response of each subdetector to the hits left by the particles. It takes into account the electronic readout of subdetectors as well as the L0 trigger. The output of Boole is the response of the LHCb subdetectors as if it comes out directly from LHCb while it is running. Information from the Monte Carlo truth are kept.

The subdetector response is taken as input by the reconstruction application called "Brunel". Since the format of simulated input data is identical to the real data that come from the DAQ system, Brunel can run without distinction both on real and on simulated data. It starts with a clustering in the tracking detectors, then the clusters are used as input in the tracking pattern recognition. The resulting tracks are passed to the Calorimeter, RICH, and Muon detectors for Particle IDentification reconstruction (PID). The reconstruction phase in the case of simulated data is followed by a Monte Carlo association phase. This phase can be switched off when Brunel runs on real data. It consists of an association between the clusters and the MC particles. If a track has nearly all its clusters associated with the same MC particle, it is itself associated to the MC particle. Otherwise it is defined as a *ghost*. If more than one track is associated to the same MC Particle, the tracks are classified as *clones*. Finally the clusters, the tracks, the particle ID information and the association tables between clusters, tracks and MC particles are stored on the DST. The DST is the output of Brunel reconstruction and it is used in the offline analysis performed with the DaVinci application.

The simulation chain described is used to generate very large samples of simulated data. These data are stored in DST files, that are used for the analysis. Particular care is put in simulating the material budget, which is important for

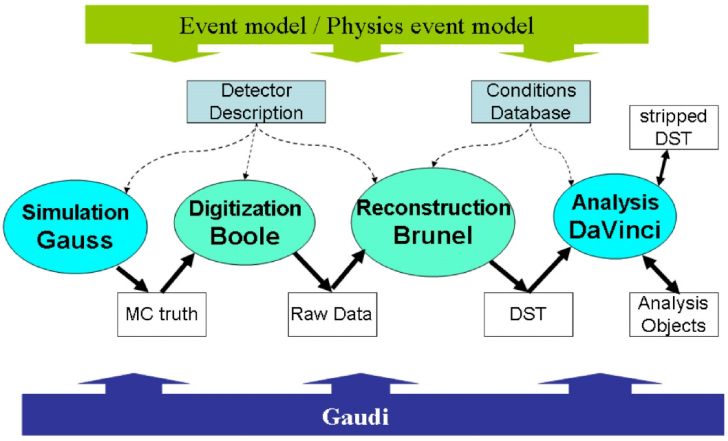

Figure 4.1: The LHCb data processing applications and data flow. Underlying all of the applications is the Gaudi framework and the event model describes the data expected. The arrows represent input/output data.

secondary interactions. A realistic detector simulation leads to an overall deterioration of the reconstruction performances, which is expected each time one pass from an ideal situation to a more realistic one.

4.2 Data manipulation

The size of all the data files stored after a simulation is huge (about 300 kB per event), but it is nothing compared to the amount of data stored after an year of data taking at LHCb. The solution adopted is to replicate the data stored at CERN in other sites, called Tier, provided with computing facilities. In this way the analysis use all the available facilities in a distributed computing model called GRID. The facility at CERN is a Tier 0 centre. Tier 1 centres serve large regions or countries; Tier 2 centres do the same on a smaller scale. This system is part of the LHC Computing GRID (LCG). The LHCb simulated data are available on GRID as if they were real data.

The selection of events and analysis is performed with the LHCb analysis framework called DaVinci. All the LHCb software, from the simulation to the DAQ, is embedded in an architecture called Gaudi. This provides a common framework for all the packages and is widely used by DaVinci. DaVinci takes as input the DST files, and it can give outputs in several formats. The output format used by me is the so called "NTuple file", which contains the physical information of the

selected events. The NTuple files represent the last step of data manipulation, they are read by an object-oriented framework, independent on the LHCb framework, which is called ROOT[41]. ROOT is a generic data analysis software, written to provide histograms, fits and other kinds of useful manipulations of data stored in an NTuple file. The most part of the golden and control channel analysis showed in this text was performed with ROOT. Once a DaVinci program was written, it must be submitted to Grid. The program used by the LHCb users to work on Grid is called Dirac. An application called Ganga was created to submit the DaVinci programs on Grid using Dirac. It menages the submission of the analysis job and the retrieval of the output.

To perform the analysis I first wrote a program in the DaVinci framework, which selects the events stored in DST files. This program had to save all the necessary physical information of the particles in the events, without including unnecessary information, to avoid a too big output file. Then I wrote a script file to submit this program on Grid by Ganga. The whole process of the NTuple creation takes a time of the order of a week. Once the output was retrieved, I analysed it locally with a ROOT executable.

Chapter 5

$B_d^0 \to J/\psi(\mu^+\mu^-)K_S^0(\pi^+\pi^-)$ selection

In the following paragraphs the selection of $B_d^0 \to J/\psi K_S^0$ will be discussed in detail. First the selection strategy and the preselection will be described, then each selection cut will be justified. Finally the selection results in terms of expected annual signal yield will be presented. The final paragraph is dedicated to the rate of selected events and its impact to the HLT and data stripping.

5.1 Selection strategy

In the text I'll call *signal* the set of particles which compose exactly the decay under study. Every set of particles which are not signal will be called *background*. If this particular decay is produced in an event, the event will be called a signal event; otherwise it will be called a background event. If a reconstructed particle corresponds to a signal particle in the Monte Carlo truth, it will be called a signal particle; and so on. The selection of a channel is a sequence of requirements on kinematics variables which are preferentially satisfied by the signal particles, and not satisfied by the background particles. These conditions are usually called selection cuts. To quantify how good a selection performs its task, a quantity is defined: the selection efficiency ($\varepsilon_{sel.}$). It is defined as the ratio between the number of selected and reconstructed signal events and the number of signal events before the selection. The larger $\varepsilon_{sel.}$ is, and the more are the signal events that can be used to measure a physical quantity. Since the error on a physical quantity depends on the statistics used, the maximization of $\varepsilon_{sel.}$ must be taken into account during the tuning of the selection.

The maximization of $\varepsilon_{sel.}$ is not the only parameter a selection has to take into account. The other important quantity is the signal to background ratio (B/S). This is the number of background events after the selection, divided by the number of signal events after the selection. The less is the B/S ratio, and the more the signal is discriminated from background. Of course the minimization of B/S and maximization of $\varepsilon_{sel.}$ are two conflicting tasks.

The minimization of B/S and maximization of $\varepsilon_{sel.}$ are common to all the selections. As this is an "unbiased and unified" selection (see paragraph 3.5) there are two additional requests to be satisfied. First of all the selection must be unbiased in the lifetime of B; which means that each selection cut can't change the selection efficiency as a function of the B lifetime. At the end of the selection $\varepsilon_{sel.}$ calculated for signal events must not be dependent on the B lifetime. An identical check is the B mean lifetime: the B mean lifetime before the selection must be the same as the one after the selection. This prevents the use of some cuts very effective in reducing B/S without reducing too much the $\varepsilon_{sel.}$. The other constraint is the unification of the selections of both the golden channel and the control channel. The strategy chosen was to perform selections as similar as possible and then compare the results, and eventually correct the discrepancies.

A previous work on the golden channel and the $B_d^0 \to J/\psi K^*$ control channel exists [42]. This work was taken into account, then a completely independent study was performed. The selection strategy was the following:

- Characterization of the kinematics of the golden channel.
- Comparison between the distributions of the signal and the background, for each of the considered kinematic variables.
- Choice of the value of the cut. This value must select more signal than background, and is chosen by comparing distributions.
- Calculation of the efficiency of the single cut. If it's too low (less then 80%) or the reduction of B/S is minimal, the process restarts.

5.2 Preselection

The objects saved in the DST files are called protoparticles. They represent the end of the reconstruction stage for the final state particles and they are the starting

point of the physics analysis. They contain the information on how they were reconstructed and all the kinematic variables. The first part of the selection is performed in the Gaudi framework and will be called *preselection*. It consists in two steps: create the final state particles and then create the resonances. To create a final state particle from a protoparticle, the mass must be identified (PID) and the four–momentum known. One can then reconstruct resonances with them. There is a set of "standard particles" predefined in DaVinci. These are a series of weak conditions on the PID probability and on kinematic variables, which decide if a protoparticle is a $K^{\pm}$ a $\mu^{\pm}$, a $\pi^{\pm}$, but also a K_S^0 or a J/ψ. These cuts are very loose and do not prevent the possibility that more than one particle is created from the same protoparticle; in these case the particles are treated separately and never mixed.

When a resonance decays the particles produced come from the same space point, which is called secondary vertex. So, retracing back the particles, we can find the secondary vertex as the intersection point of the tracks. Once a vertex is found, we can reconstruct the resonance (or mother) particle, simply combining the daughter particles. At this point a wide cut on the mass of the resonance is made, together with a wide cut on the $\chi^2/ndof$ of the secondary vertex. The above operations are applied to build K_S^0 form $\pi^+\pi^-$, J/ψ from $\mu^+\mu^-$ and B_d^0 from $J/\psi K_S^0$.

To measure the lifetime of a B hadron the pp interaction point is needed. This is called primary vertex and is built combining all the tracks together and looking for the intersection point. If more than one primary vertex is created, the Impact Parameter Significance (IPS) of the B candidate with respect to each primary vertex is calculated. The primary vertex which corresponds to the minimum IPS is used. For each event there may be several combinations of reconstructed particles that pass the preselection, both because the preselection is done with loose cuts and because the number of particles produced in each event is huge and often more than one particle is associated to one protoparticle. These different combinations are called candidates and are all saved for the successive offline selections.

Reconstructed tracks are assigned to different categories: in my thesis the relevant categories are

1. "Long tracks" (L)

2. "Downstream tracks" (D)

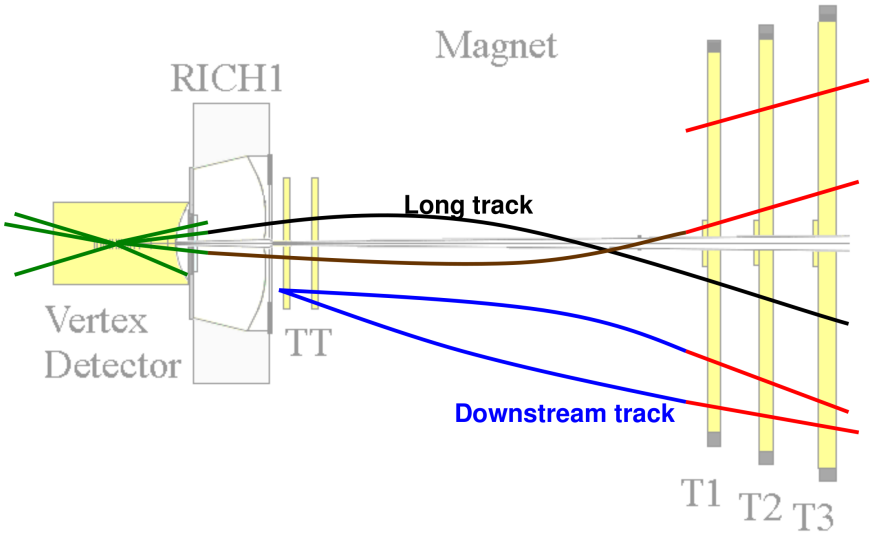

Figure 5.1: LHCb track types used in this text.

(see figure 5.1). Long tracks are composed by clusters in all the tracking stations. Since they are detected by the whole tracking system and they traverse the whole LHCb, they have the best impact parameter and p resolution. The downstream tracks have clusters in the TT and T1–T3 stations but not in the VELO. More than one half of K_S^0, which has a relatively long life, decays preferentially outside the VELO. The D tracks show a good p resolution but a worse Impact Parameter (IP) resolution. The muons are preselected as the particles which passed through the muons chambers and have long tracks. Pion preselection does not apply any cut on the PID of the protoparticles, but only a condition on the track type (L or D). The K_S^0 resonance is made either from a pair of "long" pions or from a pair of "downstream" pions. In the first case the K_S^0 and the B candidate constructed with it are called LL (long long); in the latter case the K_S^0 and B are called DD (downstream downstream). A long pion is never combined with a downstream one.

5.3 Samples used

The generic pp inelastic interactions simulated and saved in DSTs correspond to a very short period of data taking at LHCb. This is due to the computing time necessary to simulate the detector response. This sample of data (corresponding to minimum bias events) contains a negligible amount of signal, due to the fact that

the exclusive channel under study has a very small BR. A sample is therefore generated that contains only events in which the golden channel was generated. This sample was used to characterize the signal in detail and to accurately determine the efficiencies.

From the minimum bias simulated events a detailed study of the background is not possible. The main background are events similar to the signal events, with high probability to be selected. In the minimum bias the fraction of these events is low. A class of events similar to the signal are the events containing $B \to J/\psi X$, where X means all possible kind of final state. At LHCb three such kind of data samples are created: $B_d^0 \to J/\psi X$, $B_u^0 \to J/\psi X$ and $B_s^0 \to J/\psi X$. They were also created events of type $\Lambda_b \to J/\psi X$. In the analysis the events from the sample $B_{d,u,s}^0 \to J/\psi X$ are mixed together following the ratios 39.8 : 39.8 : 10.3[34].

Another source of background analysed in this thesis is the so called J/ψ inclusive events. These are defined as the events in which at least one J/ψ was produced. Since the production of J/ψ is very large in pp interactions, it is likely that these J/ψ combine randomly with two other particles to give an object with a B mass. Beside a true J/ψ plus a randomly chosen K_S, there can be the opposite situation in which a true K_S and two randomly chosen particles form a B. This type of background is contained into the $b\bar{b}$ inclusive events. These are simulated events in which at least a $b\bar{b}$ couple was produced. These three samples are to represent the most probable sources of background, and they were studied in detail. Finally, even the minimum bias data sample was taken into account, to estimate how many events are selected per second. The minimum bias events analysed have passed the L0 trigger requirements.

In all the data samples analysed, to reduce the number of simulated and saved events, when the event is outside the LHCb acceptance, the simulation is stopped. For a sample with an exclusive decay, only the events with the decay products inside the acceptance are totally simulated. In the $B \to J/\psi X$ sample an event is inside the acceptance when the muons from J/ψ are. In the inclusive data samples, not all the known decay modes are simulated: the most rare decay modes are omitted, since the probability to produce a super rare decay in these data samples is too low. Finally, among the used samples, the CP violation effect is simulated only in the golden channel sample.

5.4 Selection cuts

In the following section the selection cuts are described. For each cut it will be explained why it was chosen, and the effect for both the signal and background will be shown. To calculate the efficiency, the cuts are applied sequentially: first the pion cuts are studied, then the K_S^0 ones. After selecting the K_S^0, the muons are analysed and then the J/ψ. Finally, cuts on B will be applied. This choice helps to understand how each cut is effective. Indeed, applying a cut after another, helps to eliminate those that are just a replica of a previous one. If a cut is not effective in reducing B/S could reduce unnecessarily the efficiency.

5.4.1 Pions track quality and momentum

We define the Track Quality (TQ) as the ratio between the χ^2 and the number of degrees of freedom ($ndof$) of a track. It was chosen to select the events that have long pions with TQ < 10 and downstream pions with TQ < 20. This cut is particularly powerful in rejecting the ghost tracks (see section 4.1), which have, on average, a TQ worse than the signal tracks. The TQ distributions for both the L and the D pions are shown in figure 5.2. The efficiency for this cut is 89179/90814 = 98.2%.

The pions that come from a B are energetic since they are originated from an heavy hadron, which is also boosted. As they are light particles, their momentum is on average high, around 10 GeV/c. It exists also an enormous amount of prompt pions (about 70 per event). They come out from the fragments of pp collisions both directly and as decay products of light resonances (like kaons, or ρ). Therefore they have a small momentum, meanly around 7 GeV/c. Thus a cut on the pions momentum is applied: the pions with $p_\pi < 2$ GeV/c are rejected. This cut have an efficiency of 88542/89179 = 99.3%. As explained above the cut efficiencies are calculated with respect to the events which have passed the previous cuts. It was verified that the cuts on pions track quality and momentum do not bias the lifetime.

5.4.2 Pions IPS

The Impact Parameter (IP) is defined as the perpendicular distance between the momentum vector of a particle and the primary vertex. The Impact Parameter

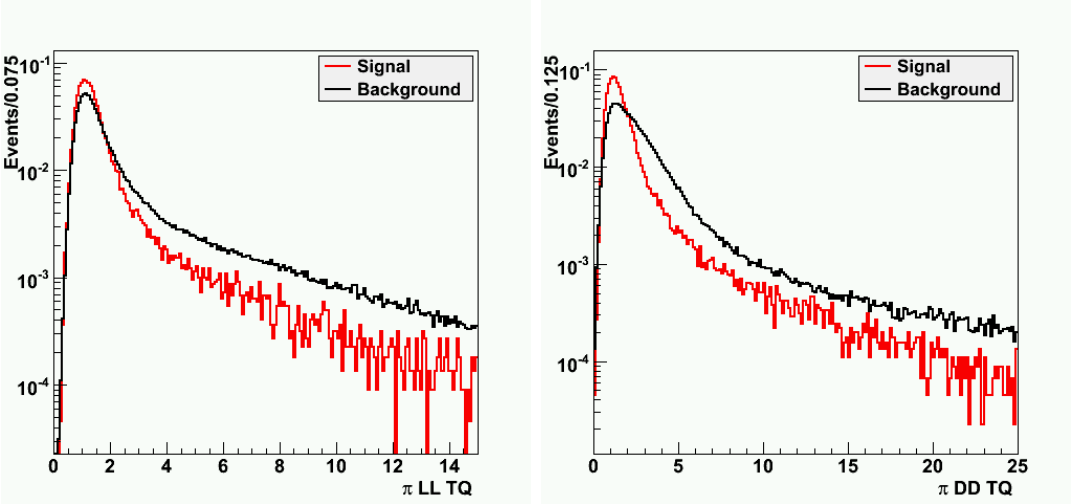

Figure 5.2: Left: TQ distribution for long pions. Right: TQ distribution for downstream pions. In red the signal, in black the background. The long pions with TQ $<$ 10 and downstream pions with TQ $<$ 20 are selected. The background shown was taken from the $B^0_{d,u,s} \to J/\psi X$. The distributions are normalized to their total area.

Significance (IPS) is the IP divided by its error. This last quantity is preferred with respect to the simple IP, since the IP errors can be large and can vary with the track. The heavier a particle is, and the more the transverse momentum (p_T) of its decay products is high. We know also that B mesons are preferentially produced in the z direction (see figure 2.2), they are boosted and travel a lot (1 cm on average) before decaying. The decay products of a B meson have therefore a high p_T and a secondary vertex far from the primary vertex. The K^0_S decay products have an IP even higher because of the long lifetime. The prompt pions come directly from the primary vertex or from a light resonance, which usually decays immediately after its production. Thus they have a low IP. Only events that have long pions with IPS $>$ 3, and downstream pions with IPS $>$ 2 are selected. This choice was made since the IP of the D tracks is worse measured, hence the distribution of the downstream pions is wider.

The efficiency for this cut is 86344/88542 = 97.5%. In figure 5.4 it is shown the cut efficiency as a function of the B lifetime. This plot was made in the following way: the content of each bin in the lifetime distribution of B *after* the pions IPS cut was divided by the content of the corresponding bin in the lifetime distribution

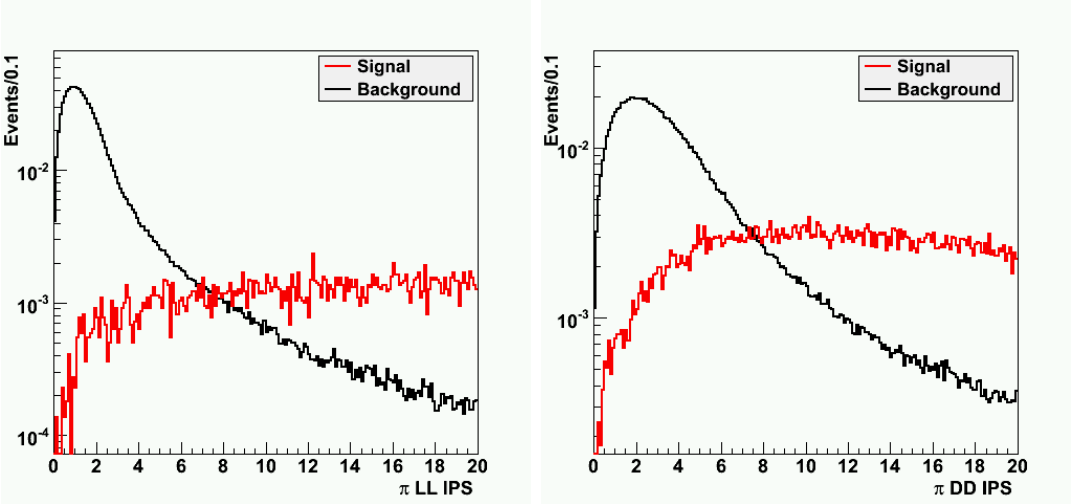

Figure 5.3: Left: IPS distribution for long pions. Right: IPS distribution for downstream pions. In red the signal, in black the background. The background shown is from the $B^0_{d,u,s} \to J/\psi X$. Long pions with IPS > 3, and downstream pions with IPS > 2 are selected.

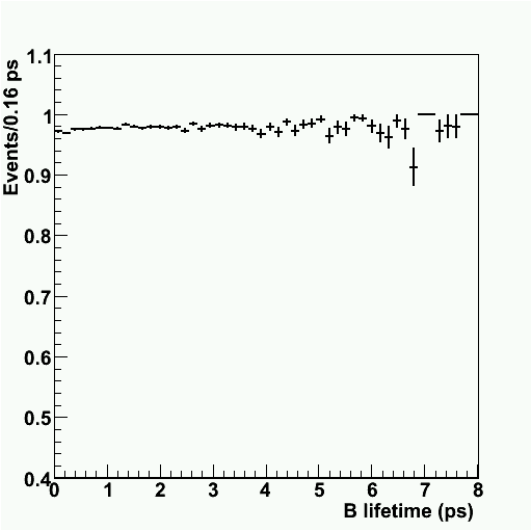

Figure 5.4: IPS cut efficiency for the pions as a function of the B lifetime.

of B *without* the pions IPS cut. The fact that the distribution is flat guarantees no need to introduce an acceptance function for the B lifetime. This cut is very powerful in rejecting background from prompt pions, and hence prompt K^0_S.

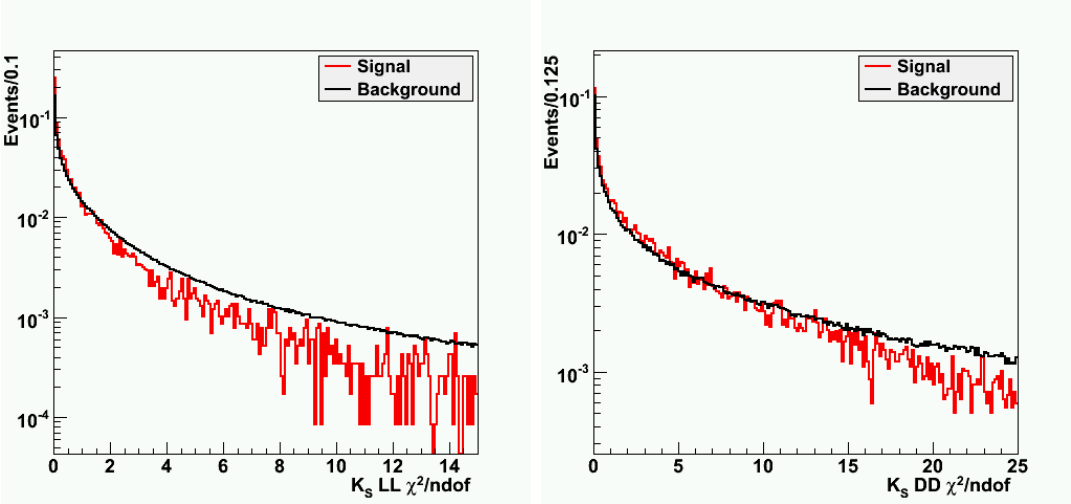

Figure 5.5: Left: $\chi^2/ndof$ distribution for LL K_S^0. Right: $\chi^2/ndof$ distribution for DD K_S^0. In red the signal, in black the background. The background shown is taken from the $B_{d,u,s}^0 \to J/\psi X$. The K_S^0 LL with $\chi^2/ndof < 10$ and the K_S^0 DD with $\chi^2/ndof < 20$ are selected.

5.4.3 K_S^0 selection

Once the two pions are reconstructed and identified, a vertex fit is made to find if they make a common secondary vertex. Each vertex fit has a χ^2, and a number of degree of freedom ($ndof$) which depends on the number of tracks used in the fit. For particles which come from the same vertex, even if they are not coming from a true K_S^0, the $\chi^2/ndof$ should be near one. For tracks which do not come from the same secondary vertex, because they are randomly matched together, the $\chi^2/ndof$ should be $\gg 1$. The determination of secondary vertex for the K_S^0 made from two downstream pions is worse, because the D tracks are determined from a smaller number of points, and these points are far from the production point. Therefore it was chosen to select as K_S^0 candidates combinations having $\chi^2/ndof < 10$ for the K_S^0 LL, and $\chi^2/ndof < 20$ for DD. The efficiency for this cut is 82273/86344= 95.3%.

Since the K_S^0 comes from an heavy particle like the B meson, a large p_T is expected. The same is not true for the prompt background. Thus only the events in which the p_T of K_S^0 is greater than 1 GeV/c are kept. The efficiency for this cut is $75759/82273 = 92.1\%$. It was checked that the cuts on the K_S^0 vertex $\chi^2/ndof$ and p_T do not bias the lifetime.

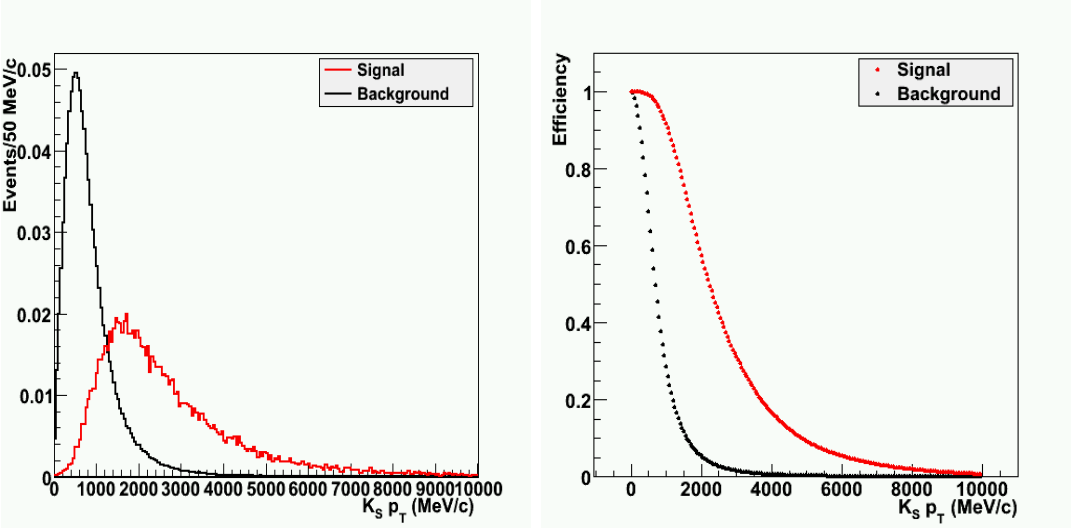

Figure 5.6: Left: p_T distribution for K^0_S. Right: cut efficiency of K^0_S p_T cut. In red the signal, in black the background. The background shown is taken from the $B^0_{d,u,s} \to J/\psi X$. K^0_S with $p_T > 1$ GeV/c are chosen.

5.4.4 Muons selection

For each protoparticle a likelihood function is constructed from the available PID information of the relevant subdetector. It tells us the combined probability to lie under a certain hypothesis. The initial hypothesis formulated for each particle is that the particle is a pion. A total likelihood function is built as the product of each subdetector likelihood. The PID information is extracted from the comparison between a given PID hypothesis and the probability of being a pion. The use of pion as initial hypothesis is dictated by the fact that the pion is the particle most abundantly produced, and it is easy to misidentify a pion with another particle. From the two hypothesis a Delta Log Likelihood (DLL) function is constructed. In the case of muons we have:

$$\Delta \ln L_{\mu\pi} = \ln \left[\frac{L(\mu)}{L(\pi)} \right] = \ln L(\mu) - \ln L(\pi) \tag{5.1}$$

For correctly identified muons $L(\mu) > L(\pi)$, thus the DLL function should give positive results. The muons are the best identified particles, because they are the only ones that traverse the whole LHCb and are detected by all the subdetectors. Moreover the muon chambers provide a precise μ identification. For these reasons the DLL variable is particularly effective in eliminating the fake muons, that have

predominantly a negative DLL value. Only the events with DLL($\mu - \pi$) greater than -5 are selected. The efficiency for this cut is 73881/75759 = 97.5%.

There is a part of background that comes from decays in flight. These are particles, mainly kaons, which travel for a certain length, and then decay into muons or pions before the muons chambers. If in their flight they passed through the RICHes while they were still kaons, the DLL($K - \pi$) from the RICHes is positive.

That is in general true also for the prompt kaons misidentified as muons, even if in this case the combined DLL($\mu - \pi$) should also work. Therefore it was chosen to select events in which the DLL($K - \pi$) for both the muons is less than 2. The efficiency for this cut is 70593/73881 = 95.5%.

In order to separate muons from the ghosts a TQ < 5 cut is applied (see paragraph 5.4.1). The efficiency for this cut is 69278/70593 = 98.1%.

Unfortunately an IPS cut can't be used to eliminate prompt muons or J/ψ. The reason is simple: a cut on the muons IPS with respect to the primary vertex corresponds to an indirect cut on the J/ψ IPS, and it bias the B lifetime. In fact the J/ψ decay immediately, thus the contribution to its IP is given by its p_T and the B flight distance. As a consequence a cut on the IPS of J/ψ corresponds to a cut on the B flight distance; and hence to a cut on the B lifetime. This effect is hidden in the cut on the pions IPS: here the cut is applied indirectly on the B flight distance plus the K_S^0 flight distance, which is almost 58 time longer than the B one. Thus the impact on the B lifetime is negligible.

The muons coming from the J/ψ ($m_{J/\psi} = 3.1$ GeV/c^2) have a large p_T. So part of prompt background is eliminated requesting that the p_T of muons is larger than 500 MeV/c. The efficiency for this cut is 67120/69278 = 96.9%. It was verified that the cuts on muons described above do not bias the lifetime.

5.4.5 J/ψ selection

The same consideration made for K_S^0 in paragraph 5.4.3 are true for the J/ψ. A cut on $\chi^2/ndof$ is useful to reject combinatorial background and a cut on p_T is useful to reject prompt background. The IPS cut, which is very powerful to eliminate the prompt background, as shown in paragraph 5.4.2, is not applied in the J/ψ case. Thus the p_T cuts, applied to both muons and to J/ψ, had to be hardened to better reject the prompt background; this leads to a significant loss of efficiency. It was chosen to select events in which the J/ψ has $\chi^2/ndof < 6$ and $p_T > 1$ GeV/c.

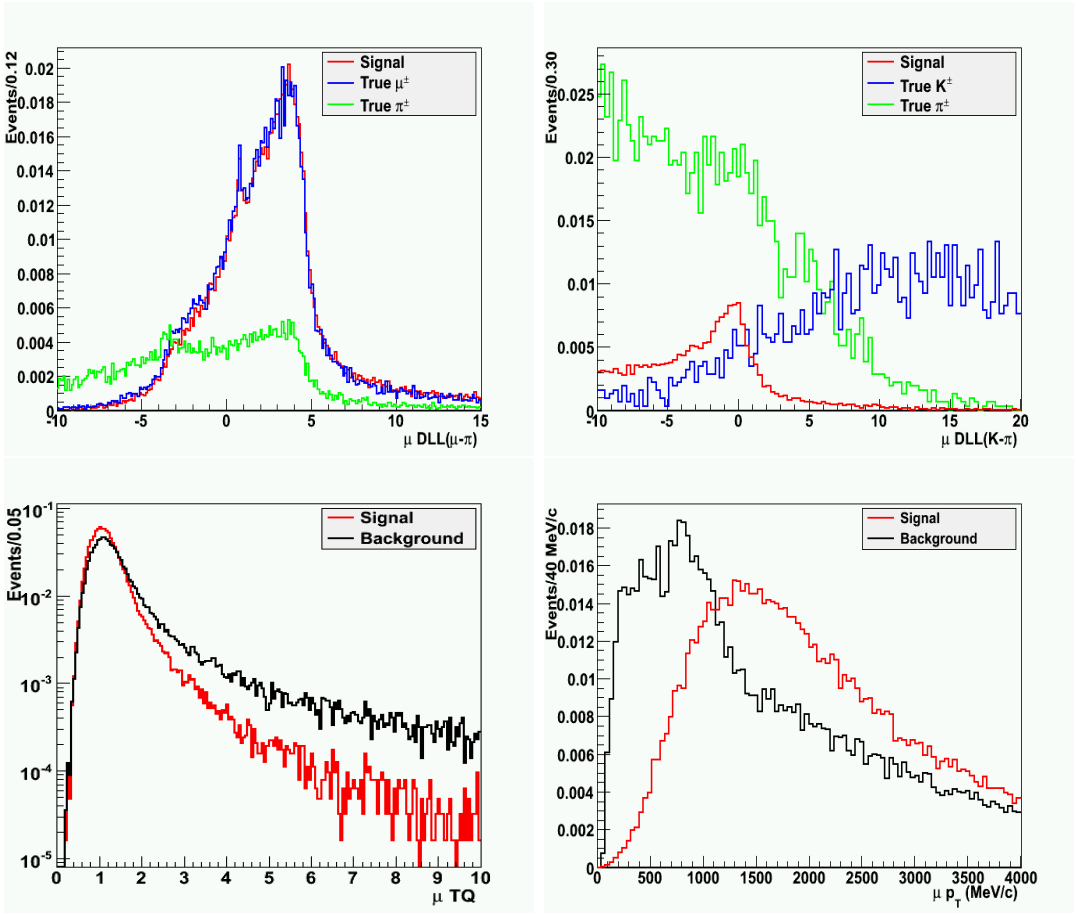

Figure 5.7: Top left: DLL($\mu-\pi$) distribution for muons. In red the signal, in blue the background of true muons, in green the background of true pions. Top right: DLL($K-\pi$) distribution for muons. In red the signal, in blue the background with true kaons, in green the background with true pions. Bottom left: TQ distribution for muons. Bottom right: p_T distribution for muons. In red the signal, in black the background. The background shown is taken from the $b\bar{b}$ inclusive. The events in which both muons have DLL($\mu-\pi$) > -5, DLL($K-\pi$)$_\mu$ <2, TQ < 5 and p_T > 500 MeV/c are selected.

The efficiencies of these cuts are 64673/67120 = 96.3% and 59306/64673 = 91.7% respectively. It was verified that the cuts on J/ψ showed above do not bias the B lifetime.

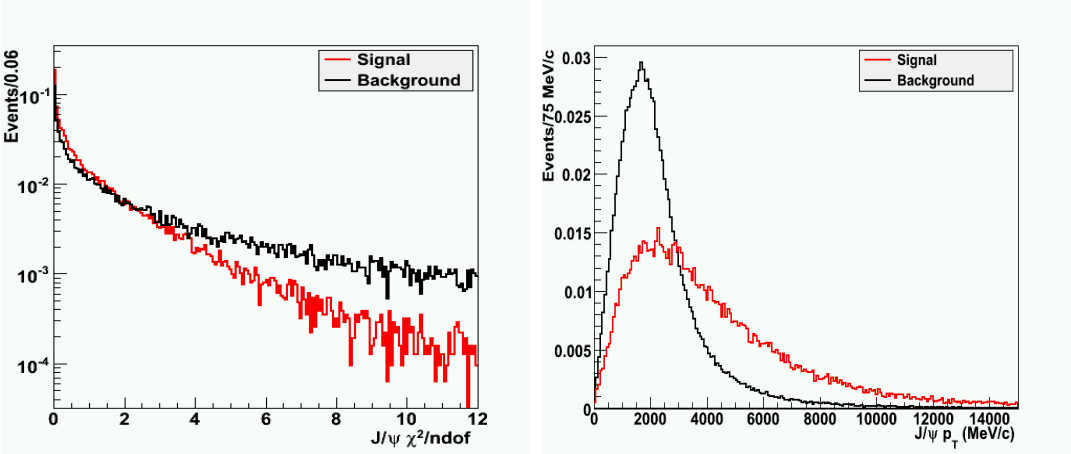

Figure 5.8: Left: $\chi^2/ndof$ distribution for J/ψ. The background shown is taken from the $b\bar{b}$ inclusive. Right: p_T distribution for J/ψ. The background shown is taken from the J/ψ inclusive. In red the signal, in black the background. The J/ψ with $\chi^2/ndof < 6$ and $p_T > 1$ GeV/c are selected.

5.4.6 J/ψ and K_S^0 masses and Δz

The most powerful selection that can be applied once a resonance is reconstructed is a cut on its mass. The masses of both the J/ψ and the K_S^0 are selected in a 3σ mass window around 497 MeV/c^2 and 3.096 GeV/c^2 respectively. The background has an uniform distribution under the mass peak, hence the better the mass resolution is, the more the B/S is reduced without loosing efficiency. Because the downstream tracks have a momentum resolution worse than the long tracks, the K_S^0 mass resolution is worse if it is reconstructed by them. Thus the LL and DD K_S^0 are handled separately. In figure 5.9 the invariant mass distribution for both the K_S^0 and the J/ψ is shown. The mass resolutions taken from a Gaussian fit on signal events are the following:

$$\sigma_{K_S^0 LL} = 4.1 \text{ MeV/c}^2 \ , \ \ \sigma_{K_S^0 DD} = 9.9 \text{ MeV/c}^2 \ \text{ and } \ \sigma_{J/\psi} = 13.7 \text{ MeV/c}^2 \ .$$

The efficiencies of the selection cuts on masses are $57712/59306 = 97.3\%$ and $54800/57712 = 94.9\%$. From figure 5.9 bottom right is also evident a tail on the left of the J/ψ mass distribution. This is due to the radiative energy loss of muons, which emit photons by bremsstrahlung and thus are reconstructed with lower momentum.

The K_S^0 has a long lifetime, and its secondary vertex is far from its production

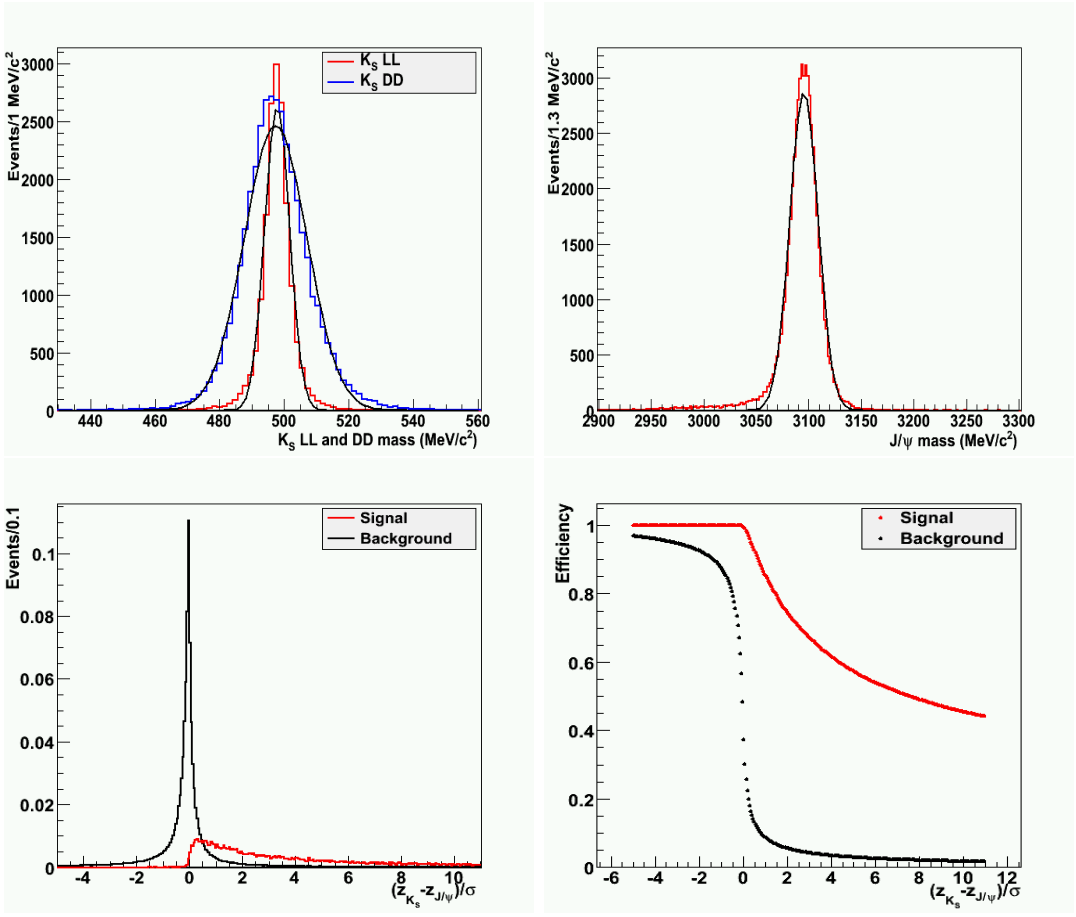

Figure 5.9: Top left: mass distribution for LL K^0_S, in red, and DD K^0_S, in blue. Top right: mass distribution for J/ψ. The selected signal events are showed. The distributions are fitted with a Gaussian. Bottom left: $(z_{K^0_S} - z_{J/\psi})/\sigma$ distribution. Bottom right: cut efficiency of $(z_{K^0_S} - z_{J/\psi})/\sigma$. In red the signal, in black the background. The background shown is taken from the $B^0_{d,u,s} \to J/\psi X$. The events with $(z_{K^0_S} - z_{J/\psi}) < 0$, and K^0_S and J/ψ mass withing 3σ mass window are selected.

vertex (i.e. the B decay vertex). On the contrary the J/ψ decays immediately, and its secondary vertex coincides with the production vertex (the B decay vertex). The J/ψ secondary vertex is the K^0_S production vertex. As the K^0_S flies in the z direction, the difference between the z coordinate of its secondary vertex and the z coordinate of its production vertex must be larger than zero. This is not always the case for the prompt and the combinatorial background. The efficiency obtained in eliminating the events with $(z_{K^0_S} - z_{J/\psi}) < 0$ is $54714/54800 = 99.8\%$. Moreover

this cut does not show bias on the B lifetime.

5.4.7 B selection

The B decay vertex is reconstructed by combining the two resonances to a common vertex. Thus a vertex $\chi^2/ndof$ selection can be applied. This selection is more effective of the one applied to each K_S^0 and J/ψ, since the probability that four randomly chosen tracks (π^+, π^-, μ^+ and μ^-) pass through a same point is lower. Because of the different vertex resolution between long and downstream tracks the events are divided between B LL (composed by LL K_S^0), and B DD (composed by DD K_S^0). Only the events in which the B LL mesons have $\chi^2/ndof < 5$, and the B DD mesons have $\chi^2/ndof < 15$ are selected. The efficiency for this selection cut is: $52936/54714 = 96.7\%$.

A cut on the IPS of B is also applied. The B is generated directly at the primary vertex, and preferentially at low angles with respect to the z axis (see figure 2.2). Therefore a low IP is expected for the B mesons. Only the B mesons with IPS less than 5 are selected. It was demonstrated that this cut, because it puts an upper limit to the IPS of B, does not bias the lifetime. The efficiency for this selection cut is: $52482/52936 = 99.1\%$.

At the end of the selection, in a restricted number of events (about 3%), there can be more than one candidate. This is mainly due to the presence of the so called "clones". The clones (see section 4.1) are tracks almost identical, which are associated to the same Monte Carlo truth. Since these tracks have practically the same kinematic variables, it doesn't matter what is the track chosen. Two criteria were used to chose the best candidate. If the 3σ mass window is considered, the candidate with the best B vertex is chosen. If the large 500 MeV/c^2 mass window is considered, the candidate with a B mass closest to the expected one and outside a 60 MeV/c^2 mass window is chosen. If both candidates are inside the 60 MeV/c^2 mass window the minimum $\chi^2/ndof$, criterion is applied.

The efficiency of this criterion is $52326/52482 = 99.7\%$. The very low number of multi–candidate events and the presence of clones reduce the impact of this choice, which results negligible.

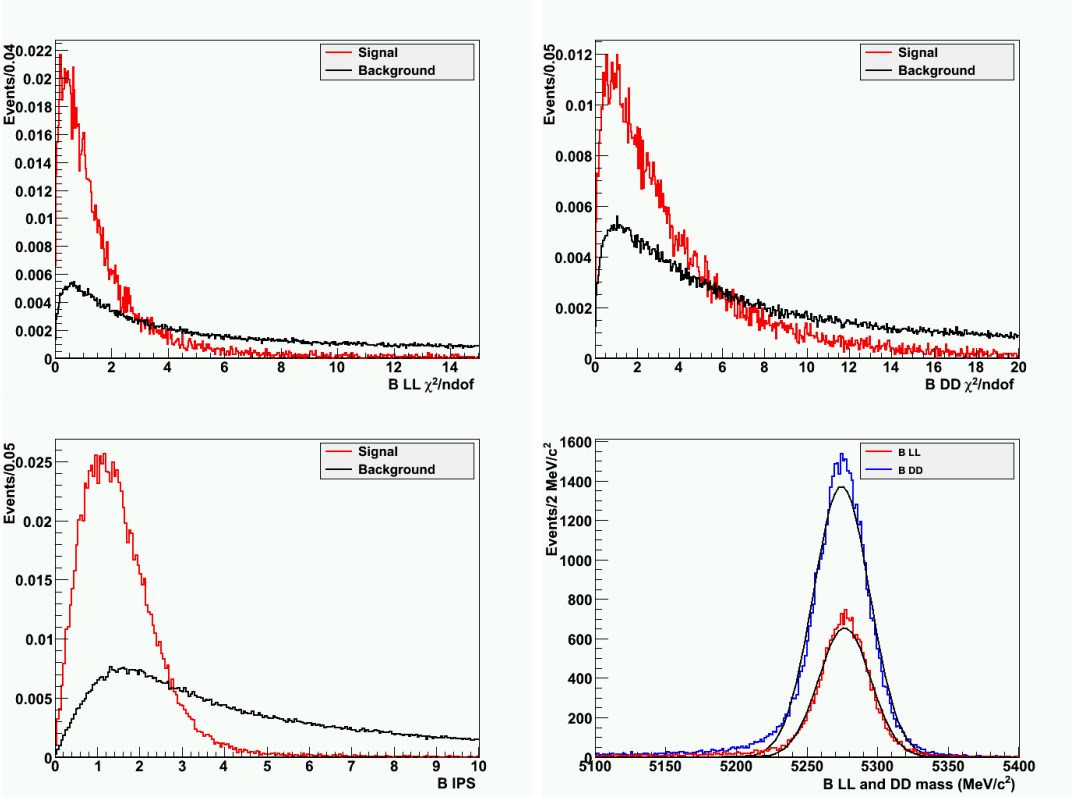

Figure 5.10: Top left: vertex $\chi^2/ndof$ distribution for LL B. Top right: vertex $\chi^2/ndof$ distribution for DD B. Bottom left: IPS distribution for B. In red the signal, in black the background. The background shown is taken from the $B^0_{d,u,s} \to J/\psi X$. Bottom right: mass distribution for LL B (in red) and DD B (in blue). Only the selected signal events are shown. The distributions are fitted with a Gaussian.

5.5 Selection results

At the end of the selection the percentage of B LL is 30%, that one of B DD is 70%. The signal mass distributions of the selected and reconstructed B are shown in figure 5.10. The uncertainties on the B mass are: $\sigma_{m(LL)} = 18$ MeV/c^2 for the B LL, and $\sigma_{m(DD)} = 20.4$ MeV/c^2 for the B DD.

The lifetime distribution of the selected and reconstructed B is shown in figure 5.11; it is also shown the lifetime from the Monte Carlo truth. The lifetime distribution of signal events taken before the selection is plotted in figure 5.11 in blue. Each histogram of figure 5.11 is fitted with an exponential, from which the mean

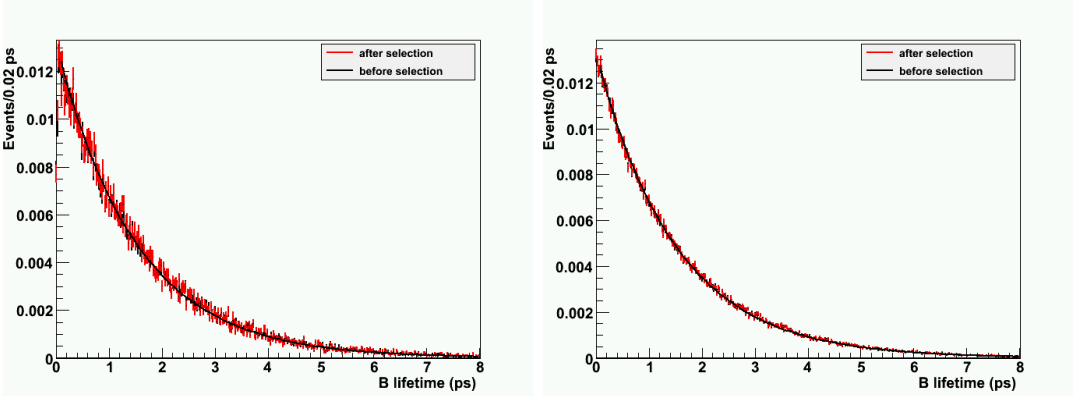

Figure 5.11: Left: lifetime distribution for reconstructed B before the selection (blue) and after the selection (red). Right: lifetime distribution for Monte Carlo truth B before the selection (blue) and after the selection (red). Only the signal events are shown. The distributions of selected events are fitted with an exponential.

B signal	τ (ps)
reconstructed after the selection	1.514 ± 0.008
Monte Carlo truth after the selection	1.512 ± 0.008
reconstructed before the selection	1.516 ± 0.006
Monte Carlo truth before the selection	1.512 ± 0.006
Value put in the simulation	1.536

Table 5.1: Results of the exponential fits performed on the distributions of figure 5.11.

lifetime is obtained. The fitted values are reported in table 5.1.

The reconstructed B lifetime agrees with the Monte Carlo truth, and the results before and after the selection agree. This is the important result achieved with the unbiased selection. There is a discrepancy with the value of the B lifetime used in the simulation. The reconstructed B lifetime is 2.75 σ far from the expected one. This underestimation was observed also for other simulated channel. This indicates a possible bias not related to the selection, but to the simulation or the reconstruction, and it must be understood. It was observed that the tracking efficiency decreases with the flight distance. This is due to requirements in the pattern recognition on the presence of VELO hits.

Cuts	efficiency
TQ ($\pi^\pm$ LL) < 10 TQ ($\pi^\pm$ DD) < 20	98.2 %
p ($\pi^\pm$) > 2 GeV/c	99.3 %
IPS ($\pi^\pm$ LL) > 3 IPS ($\pi^\pm$ DD) > 2	97.5 %
vertex $\chi^2/ndof$ (K_S^0 LL) < 10 vertex $\chi^2/ndof$ (K_S^0 DD) < 20	95.3 %
p_T (K_S^0) > 1 GeV/c	92.1 %
DLL $(\mu - \pi)_\mu >$ -5	97.5 %
DLL $(K - \pi)_\mu < 2$	95.5 %
TQ ($\mu^\pm$) < 5	98.1 %
p_T ($\mu^\pm$) > 500 MeV/c	96.9 %
vertex $\chi^2/ndof$ (J/ψ) < 6	96.3 %
p_T (J/ψ) > 1 GeV/c	91.7 %
mass window (K_S^0 LL) $= 3 \times 4.1$ MeV/c^2 mass window (K_S^0 DD) $= 3 \times 9.9$ MeV/c^2	97.3 %
mass window (J/ψ) $= 3 \times 13.7$ MeV/c^2	94.9 %
$(z_{K_S^0} - z_{J/\psi}) > 0$	99.8 %
vertex $\chi^2/ndof$ (B LL) < 5 vertex $\chi^2/ndof$ (B DD) < 15	96.7 %
IPS (B) < 5	99.1 %
B with minimum vertex $\chi^2/ndof$	99.7 %
$\varepsilon_{sel.} =$	57.4 %

Table 5.2: Selection cuts and the relative efficiencies.

In table 5.2 all selection cuts are reported. In a 3σ mass window around the B, the reconstruction and selection efficiency of signal events is $\varepsilon_{rec.\&sel.} = 0.0522 \pm 0.0002$. The L0 trigger efficiency (defined as the number of reconstructed events selected by L0 trigger, divided by the total number of reconstructed events) is $\varepsilon_{L0} = 0.9390 \pm 0.0009$. This is a very high value of trigger efficiency, due mainly to the presence of a $J/\psi \to \mu^+\mu^-$ in the decay. In fact the two muons from the J/ψ have an high p_T and are the perfect candidates to be triggered by the muon alleys of the L0 trigger. It implies another considerable advantage: for both the golden channel and the control channel almost all the events are triggered by a signal particle. It was observed a difference in the tagging performances between events triggered by a signal particle (Trigger On Signal, TOS) and events triggered by a non signal particle (Trigger Independent of Signal, TIS). That is due to the fact that in the TIS events the tagging particles are more likely in the LHCb acceptance. Therefore if the ratio between TIS and TOS events of the golden channel was different from the one of the control channel, the tagging performances could be different. At the time in which the study was performed the HLT algorithms were still in a development status, thus the HLT efficiency is not considered in this thesis.

As explained in paragraph 5.3, only the events in the LHCb acceptance are completely simulated. Thus the generation efficiency ($\varepsilon_{gen.}$) must be taken into account. It is defined as the number of events completely simulated, divided by the number of events generated in 4π. For the golden channel data sample used $\varepsilon_{gen.} = 0.207 \pm 0.004$.

The calculation of the B/S is more difficult, because the B/S obtained with different kind of data sample must be combined. There are some kind of background events that are present in more than one data sample. Thus there is a risk of double counting such background events and then overestimating the total B/S. To avoid this problem I divided the background events in three independent categories and I calculated the B/S for each of them:

1. $B_{(d,u,s)} \to J/\psi X$ background: $(\mathrm{B/S})_X = 258 \ / \ 1397 = 0.185 \pm 0.012$
2. J/ψ inclusive background, without $B_{(d,u,s)} \to J/\psi X$ events: $(\mathrm{B/S})_{J/\psi\ incl.} = 197 \ / \ 73 = 2.7 \pm 0.4$
3. $b\bar{b}$ inclusive background, without events with a J/ψ: $(\mathrm{B/S})_{b\bar{b}\ incl.} = 1 \ / \ 7 = 0.14 \pm 0.15$

These three categories are completely independent, thus their B/S can be summed to have the total expected B/S, $(\mathrm{B/S})_{tot} = 3.02 \pm 0.43$. From the three values of B/S, it is clear that the dominant of background is due to the events with a prompt J/ψ (category 2). This was expected, because the IPS cuts, which are very powerful in rejecting prompt background, can't be used in the J/ψ selection. This type of background has a B lifetime distribution which is concentrated at very low lifetimes. This characteristic is taken into account in the $\sin 2\beta$ fit, and reduces drastically the impact of this background to the $\sin 2\beta$ sensitivity. This is not true for the $B_{(d,u,s)} \to J/\psi X$ background (category 1), which is composed by events with a true B. These events have the B lifetime and they can undergo a $B\overline{B}$ oscillation due to the presence of B_d^0 and B_s^0. This is superimposed to the signal oscillation and spoils the $\sin 2\beta$ fit. For these reasons the amount of this type of background was reduced down to a reasonably low value.

A study of the $b\bar{b}$ inclusive background (category 3) shows that the few number events selected are all combinatorial background.

5.6 Signal and background statistics

To calculate the yield of events in one year (10^7 s), the branching ratios of the observed decays must be known. This is called the "visible" branching ratio (BR_{vis}) and it is the product of the BR's of all the decay modes considered. From [34] we have:

$$\begin{aligned} \mathrm{BR}_{vis} &= BR_{B_d^0 \to J/\psi K_S^0} \times BR_{J/\psi \to \mu^+\mu^-} \times BR_{K_S^0 \to \pi^+\pi^-} \\ &= 4.36 \cdot 10^{-4} \times 5.93 \cdot 10^{-2} \times 69.2 \cdot 10^{-2} = 1.79 \cdot 10^{-5} \ . \end{aligned} \tag{5.2}$$

Thus the annual signal yield of reconstructed, selected and L0 triggered events expected at LHCb is:

$$\begin{aligned} Y_S &= L_{int} \times \sigma_{b\bar{b}} \times 2 \cdot f_B \times \mathrm{BR}_{vis} \times \varepsilon_{rec.\&sel.} \times \varepsilon_{gen.} \times \varepsilon_{L0} \\ &= 2\ \mathrm{fb}^{-1} \times 500\ \mu b \times 2 \cdot 0.398 \times 1.79 \cdot 10^{-5}\ \mathrm{cm}^{-2}\mathrm{s}^{-1} \times \\ &\quad \times 0.0522 \times 0.9390 \times 0.207 \\ &= 144568 \text{ events}, \end{aligned} \tag{5.3}$$

where L_{int} is the luminosity collected in one year by LHCb; for convention $L_{int} = 2 \cdot 10^{32}\ \mathrm{cm}^{-2}\mathrm{s}^{-1} \times 10^7\ \mathrm{s} = 2\ \mathrm{fb}^{-1}$. $\sigma_{b\bar{b}} = 500\ \mu b$ is the $b\bar{b}$ cross section production at 14 TeV. Its value is an extrapolation from the value measured at Tevatron ($\sqrt{s} = 2$

TeV) , and is not well known. f_B is the fraction of b quarks which hadronize into a B_d^0 meson, and the factor 2 stands for the two quarks. From [34] we take: $f_B = 0.398$.

The expected annual background yield is:

$$Y_B = Y_S \times \left(\frac{B}{S}\right)_{tot} = 436596 \text{ events.} \quad (5.4)$$

5.7 Minimum bias and stripping selection

As explained in paragraph 2.2.10 HLT last selection is the OR between a set of inclusive selections and a set of exclusive ones. Therefore it is enough that the signal pass one of these selections to be kept. The final acquisition rate of 2 kHz is divided between these selections: to the inclusive selections is assigned an high rate, instead the exclusive selections have to work with a lower rate. For instance the rate of the J/ψ inclusive selection is 600 Hz, while the rate dedicated to an exclusive selection is about 1.4 Hz. This decay mode can be selected by the HLT algorithms in two different ways: either in the J/ψ inclusive selection, or in a dedicated exclusive selection. The selection described above, applied to the 2526620 minimum bias L0 triggered events, selects only one event in the 500 MeV/c^2 B mass window. It corresponds to a rate approximately around 0.4 Hz. If all these events are kept, they would be 4 million per year. This is only a very approximate estimate, due to the very low statistics of minimum bias events. The number of events per year could be considerably larger, and in this case it will be very difficult to accurately reconstruct them.

Once the events will be saved selection criteria must be applied to the events, and only the events which will pass them will be saved in DST format ("stripping"). The stripping selection has the same structure of the HLT selection, but have a lower writing rate.

A study was made on the minimum bias events, to understand how the rate of selected events can be lowered. This study revealed that the minimum bias background is almost exclusively of prompt type. As previously explained, the weak point in rejecting prompt background is the selection of the J/ψ. To reject prompt background in the J/ψ system without biasing the B lifetime, only the minimum p_T of either muons or J/ψ can be changed. In particular my study confirmed that the cut on the p_T of muons is the most effective in rejecting minimum bias back-

ground. If p_T $(\mu^{\pm}) < 1$ GeV/c, all the minimum bias background is eliminated, even lowering the other selection cuts. Such a hard cut dramatically reduces the selection efficiency to $\varepsilon_{sel.\&rec.} = 0.040 \pm 0.00018$. It should be accurately tuned to reduce the amount of minimum bias background without loosing too much efficiency. This tuning can't be performed with the low amount of minimum bias events simulated. From the study performed, the minimum value of $p_T(\mu)$ should be taken between 0.5 and 1 GeV/c. The choice of this value depends on the HLT and stripping constraints.

Chapter 6

$B_d^0 \to J/\psi(\mu^+\mu^-)K^{*0}(K^+\pi^-)$ selection

The $B_d^0 \to J/\psi K^{*0}$ decay mode is the ideal control channel for the $B_d^0 \to J/\psi K_S^0$. They are both decays of $B_d^0 \to J/\psi X$ type, and they only differ for the mass of one of the decay products. Moreover the final state of $J/\psi K^{*0}$ is not a CP eigenstate, thus the mistag fraction can be determined. We chose to reconstruct the $K^{*0} \to K^+\pi^-$ and $\overline{K}^{*0} \to K^-\pi^+$ (corresponding to the 66%) final states: the charge of the kaon gives the flavour of the s quark in K^{*0}, and thus the flavour of B at the decay. For this reason the $K^{*0} \to K^0\pi^0$ decay mode is ignored. My work on this channel concerned mainly the selection and the comparison of the tagging results with the golden channel.

In this paragraph the selection of the control channel is shown using the same structure and formalism of the previous chapter. Many considerations are identical to the ones given in the previous chapter, so only a short description of the selection and its results will be given.

6.1 Selection strategy

The selection adopted must be as similar as possible to that for the golden channel. The data sets analysed to study the background were the same as for the golden channel. As far the J/ψ is concerned, the same preselection was applied. The K^{*0} considered is the $K^*(892)^0$. It is a resonance with an extremely short lifetime and a width $\Gamma = 50$ MeV. The selection strategy was to minimize B/S and maximize

$\varepsilon_{sel.}$ maintaining the lifetime unbiased and replicating, where it was possible, the golden channel selection cuts.

6.2 Selection cuts

The main difference with respect to the golden channel is the difficulty to reject the background in the K^{*0} selection. In the golden channel the main issue was to reduce the prompt background in the J/ψ selection; the K_S^0 didn't give particular problems. The rejection of prompt K^{*0} presents the same limitations as of the prompt J/ψ, and moreover the K^{*0} has a mass width four time larger than the K_S^0. For these reasons there must be differences between the selection of the two channels. In general the control channel selection can be seen as a tighter version of the golden channel one. The cuts were tightened until an acceptable background amount was reached.

The selected pions and kaons must have a track quality smaller than 10 and a momentum larger than 2 GeV/c. The first request reduces the amount of ghosts, the second reduces the prompt background. The cut on momentum is identical to the one applied to the K_S^0 decay products.

In many events in which prompt background is selected a proton reconstructed as a kaon is observed. This misidentification has a low probability, but the amount of protons produced is very large. For this reason more strict kaon identification is applied to the $K^{\pm}$: DLL($K^{\pm} - p$) > -10. This exploits the ability of the RICHes to identify the charged hadrons (see figure 6.1 left).

As explained in paragraph 5.4.2, it is not possible to apply an IPS cut on the pion and kaon from the K^{*0} because this should bias the B lifetime. Moreover the K^{*0} mass is only 896 MeV/c^2, too low to apply p_T cuts to its daughter particles. These facts represent two serious limits to the prompt background rejection in this channel.

The K^{*0}, as the K_S^0, is selected by requesting a small $\chi^2/ndof$ of the vertex and a high p_T, namely: $\chi^2/ndof < 20$ and $p_T > 1.5$ GeV/c. The tightening of the p_T is the only way to reduce the prompt background from K^{*0}. As a consequence, a difference in the p_T spectrum of the B with respect to the golden channel is introduced. This can give a difference in tagging performances of the two channels (as will be discussed in the following chapter) and must be kept under control. Unfortunately it is not possible to apply the same strong cut to the K_S^0, since it

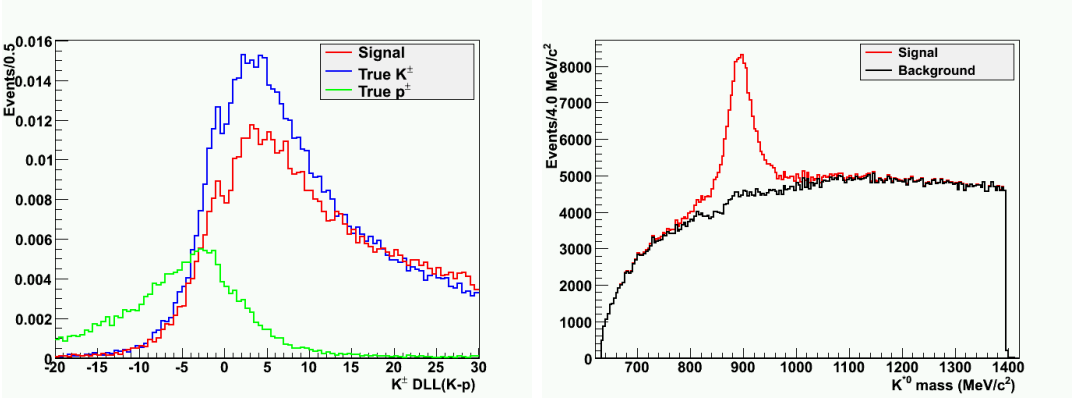

Figure 6.1: Left: DLL($K^\pm - p$) distribution for the $K^\pm$. In red the signal, in black the background of true $K^\pm$, in green the background of true protons. The events with DLL($K^\pm - p$) $>$ -10 are selected. Right: mass distribution for the $K^*(892)^0$. In red the signal, in blue the background. The events within the 3σ mass window are selected. The background shown is taken from the $B^0_{d,u,s} \to J/\psi X$.

implies a too strong reduction in $\varepsilon_{sel.}$.

A 3σ mass window cut is finally applied. The $K^*(892)^0$ width is $\Gamma = 50$ MeV. This means that the cut can't be reduced because the main contribute is given by the intrinsic width of the K^{*0}. Obviously with a larger mass window a larger amount of background is collected.

The J/ψ and B selections applied are the same as for the golden channel (see paragraphs from 5.4.4 to 5.4.7). The only relevant difference is the request of $\chi^2/ndof < 5$ on the B vertex.

6.3 Selection results

The mass distributions of the selected and reconstructed B is shown in figure 6.2. The width of the B mass distribution is: $\sigma_m = 18.7$ MeV/c^2 .

The mean lifetime is reported in table 6.1. From table 6.1 we can see that the selection doesn't bias the lifetime.

In table 6.2 the selection cuts are reported. In a 3σ mass window of B candidates, the reconstruction and selection efficiency of signal events is $\varepsilon_{rec.\&sel.} = 0.0923 \pm 0.0002$. If we compare it with the $\varepsilon_{rec.\&sel.}$ of the golden channel, we see that it is almost the double, despite of the harder selection. One reason is the

B signal	τ (ps)
reconstructed after the selection	1.526 ± 0.004
Monte Carlo truth after the selection	1.518 ± 0.004
reconstructed before the selection	1.536 ± 0.003
Monte Carlo truth before the selection	1.526 ± 0.003

Table 6.1: Results of the exponential fits performed on the lifetime distribution of the B mesons.

Cuts
TQ ($K^\pm$ and $\pi^\pm$) < 10
p ($K^\pm$ and $\pi^\pm$) > 2 GeV/c
DLL($K - p$) > -10
vertex $\chi^2/ndof$ (K^*) < 20
p_T (K^*) > 1.5 GeV/c
TQ ($\mu^\pm$) < 5
p_T ($\mu^\pm$) > 500 MeV/c
DLL $(\mu - \pi)_\mu$ > -5
DLL $(K - \pi)_\mu < 2$
vertex $\chi^2/ndof$ (J/ψ) < 6
p_T (J/ψ) > 1 GeV/c
mass window (K^*) $= 3 \times 35$ MeV/c^2
mass window (J/ψ) $= 3 \times 13.7$ MeV/c^2
vertex $\chi^2/ndof$ (B) < 5
IPS (B) < 5

Table 6.2: Selection cuts.

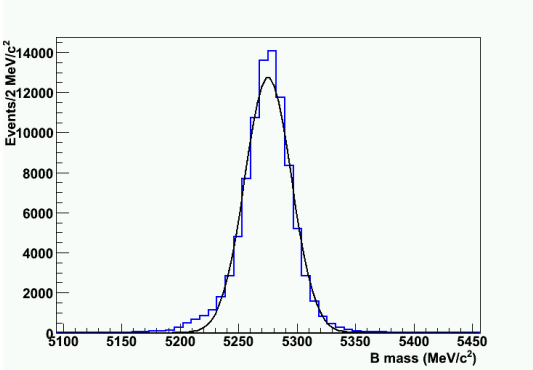

Figure 6.2: Mass distribution for B candidates. Only the signal events are shown. The distribution is fitted with a Gaussian.

better reconstruction efficiency of the K^{0*}'s decay products. They have mainly long tracks and thus are better reconstructed. Another reason is the fact that the K^{0*} decay always inside the acceptance, which is not true for the K_S^0's. The L0 trigger efficiency is $\varepsilon_{L0} = 0.94 \pm 0.0004$. The generation efficiency for this channel is $\varepsilon_{gen.} = 0.17 \pm 0.0011$.

In order to calculate the B/S ratio, the same categories shown in paragraph 5.5 were used:

1. $(\mathrm{B/S})_X = 3357/6048 = 0.555 \pm 0.012$ from $B_{(d,u,s)} \to J/\psi X$
2. $(\mathrm{B/S})_{J/\psi\ incl.} = 1456/239 = 6.1 \pm 0.4$ from J/ψ inclusive background, without $B_{(d,u,s)} \to J/\psi X$ events
3. $(\mathrm{B/S})_{b\bar{b}\ incl.} = 1/9 = 0.11 \pm 0.12$ from $b\bar{b}$ inclusive background, without any events with a J/ψ

The total B/S expected is $(\mathrm{B/S})_{tot} = 6.8 \pm 0.4$. Despite the harder selection, the B/S are almost two time larger than those of the golden channel.

In this selection study the minimum bias background was not considered. A run on minimum bias data samples with this selection shows a huge amount of events selected (about 100 million per year). Clearly this amount is too large and this selection must be tightened. The selection efficiency is large enough to apply harder cuts. The same is not true for the golden channel. Therefore a tightener selection of the control channel means an even larger difference between the two selections, which could imply differences in the tagging performances.

6.4 Signal and background yield

From [34] we have:

$$\begin{aligned} \mathrm{BR}_{vis} &= BR_{B_d^0 \to J/\psi K^*(892)^0} \times BR_{J/\psi \to \mu^+\mu^-} \times BR_{K^{0*} \to K^+\pi^-} \\ &= 1.33 \cdot 10^{-3} \times 5.93 \cdot 10^{-2} \times 66.6 \cdot 10^{-2} = 5.25 \cdot 10^{-5} \ . \end{aligned} \tag{6.1}$$

Thus the annual yield expected at LHCb is:

$$\begin{aligned} Y_S &= L_{int} \times \sigma_{b\bar{b}} \times 2 \cdot f_B \times \mathrm{BR}_{vis} \times \varepsilon_{rec.\&sel.} \times \varepsilon_{gen.} \times \varepsilon_{L0} \\ &= 2 \text{ fb}^{-1} \times 500 \ \mu b \times 2 \cdot 0.398 \times 5.25 \cdot 10^{-5} \text{ cm}^{-2}\text{s}^{-1} \times \\ &\quad \times 0.092318 \times 0.1726 \times 0.9415 \\ &= 626931 \text{ events}, \end{aligned} \tag{6.2}$$

to be compared to the signal $B_d^0 \to J/\psi K_S^0$ yield calculated in 5.2 and 5.3. The expected annual background yield is:

$$Y_B = Y_S \times \left(\frac{B}{S}\right)_{tot} = 4263131 \text{ events}. \tag{6.3}$$

Chapter 7

Comparison between the golden channel and the control channel mistags

To measure $\sin 2\beta$ from a time dependent asymmetry of $B_d^0 \to J/\psi K_S^0$ events, the flavour of B must be known. In this chapter the algorithms which give the B flavour are described. The uncertainty on $\sin 2\beta$ depends on the probability of determining incorrectly the B flavour. This probability, called mistag, can be measured in the control channel, and applied to the signal channel providing that it is the same in the two channels. In this chapter a comparison between the mistag in both channels is made, and it will be demonstrated that the signal events after the selection have the same mistags than the control channel events. In this chapter only selected and L0 triggered events are considered.

7.1 Flavour tagging

The flavour of a quark is a quantum number conserved in the strong and electromagnetic interactions and violated in the weak interactions. There are 6 flavour numbers which correspond to the 6 quarks: u, d, s, c, b and t. Each quark has the corresponding flavour number f = -1 if it is a quark of the down–type, and $f = +1$ if it is up–type. All the other flavour numbers are $f = 0$. The antiquark has the corresponding flavour number $f = +1$ if it is down–type, f = -1 if it is up–type. For instance a b quark has flavour number b = -1, and $\overline{b}$ has $b = +1$. Therefore, the

type of quark and the flavour identify a quark or an antiquark. In the experimental physics the quark flavour is often used as a synonymous for knowing if a quark is a particle or an antiparticle. The quarks that compose a hadron are the C conjugate of those ones which compose the antihadron. Thus, if the neutral hadron is not a C eigenstate, the knowledge of the quark flavour corresponds to the knowledge of the neutral meson flavour. For instance the B_d^0 is composed by $\overline{b}d$, while the $\overline{B}_d^0$ is composed by $b\overline{d}$. Hence the flavour of b gives the flavour of the meson.

The flavour tagging[43] is the algorithm which infers the flavour of a neutral meson at the production. There are two types of flavour tagging algorithms: Opposite Side (OS) and Same Side (SS).

The OS tagging looks for the information on the flavour of the opposite side b. In fact the b quarks are always produced in pairs, and the probability to produce more than one pair per bunch crossing is low. Thus each time a b is produced, there is another b with opposite flavour. If there is no oscillation, when one flavour is known, automatically also the other is known.

Conventionally the B under study is called "the B", the other is called "opposite side B". The opposite b will hadronize in all kinds of b hadrons. These hadrons have so many decay modes that their exclusive reconstruction has a very low efficiency. So an inclusive selection is done, with the aim of finding the particles which are the decay products of the opposite side b. Since a b goes into a positive lepton ($b \to l$), and an anti b into a negative lepton ($\overline{b} \to \overline{l}$), the charge of the lepton that comes from a semileptonic decay mode gives directly the flavour of b. The most frequent hadronic decay chain is $b \to c \to s$ or $\overline{b} \to \overline{c} \to \overline{s}$. The OS tagging algorithm consists in a selection of all the electrons, muons and charged kaons which most probably come from the opposite b hadron. These particles are called OS taggers. They are selected with tight requirements on the PID likelihood. The usual selection cuts for b decay products are applied: high p and p_T. An additional cut on the IPS is applied to the OS kaons, because they come from a B meson that decay in a D meson. More information on the OS flavour tagging algorithms are given in reference [43]. Another OS tagger is the so called vertex–charge. It consists in reconstructing inclusively a b vertex and summing the charge of the tracks from that vertex. The vertex–charge is used to determine the charge of the b quark: if the vertex–charge is positive the quark is a $\overline{b}$ and vice versa.

The SS tagging looks for the information on the flavour of the signal b. When a $B_d^0(\overline{b}d)$ meson is formed in the hadronization process, a d quark is pulled out from

the sea. Thus a $\overline{d}$ remains. This $\overline{d}$ hadronizes with high probability in a $\pi^+(u\overline{d})$. If the signal meson was an $\overline{B^0_d}$, a π^- would be produced. These tagger particles are called "same side pions". The identification of their charge gives directly the flavour of the signal b at the production. They are detected as the most energetic pions coming from the primary vertex; thus cuts on p, p_T and IPS are applied. More information on the SS flavour tagging algorithms are given in reference [43].

Even if a signal B meson is completely reconstructed, it is not guaranteed that a tagger is found. The tagger particle must be inside the LHCb acceptance, it must be reconstructed and it must satisfy the tagging selection. The tagging efficiency (ε_{tag}) is defined as the ratio between the selected and tagged signal events divided by all the selected signal events. It is usually less than 60%. The tagging efficiencies are defined for each tagger.

The information given by a tagger can be wrong for many reasons. For instance the selected particle is not a b decay product, or it has been misidentified, or it is a ghost. If this wrong tagger is selected randomly among all the particles in the event, the probability to give the wrong flavour is 50%. However there are particular cases in which this probability is higher. For istance if the OS b decays following the chain: $b \to c \to \overline{l}$, and the $\overline{l}$ is selected as the opposite side lepton, it gives always the wrong flavour of signal b.

For the OS taggers, there is an intrinsic reason that brings to a wrong tagging. They detect the flavour of the OS B at the moment of decay, but if it has oscillated, the flavour assigned to the signal b is the wrong one. To take these cases into account, the *wrong tag fraction*, or *mistag probability* (ω_{tag}) is defined. If N_R is the number of events in which the flavour is correctly tagged, and N_W is the number of events in which the flavour is wrongly tagged, we define:

$$\omega_{tag} = \frac{N_W}{N_R + N_W} \tag{7.1}$$

it is usually less than 40%.

The mistag affects all the measured asymmetries. An asymmetry is defined as:

$$a_{th}(t) = \frac{\Gamma_{\overline{B^0}}(t) - \Gamma_{B^0}(t)}{\Gamma_{\overline{B^0}}(t) + \Gamma_{B^0}(t)} \quad , \tag{7.2}$$

where $a_{th}(t)$ is the precise asymmetry, and $\Gamma_{B^0}(t)(\Gamma_{\overline{B^0}}(t))$ are the decay rates as a function of time of an initial $B^0(\overline{B^0})$ that decays into a given final state. Experimentally if we tag a particle as a B^0, we can only say that it has a decay rate

$\Gamma_{B^0}(t)$ with the probability $(1-\omega_{tag})$, and a decay rate $\Gamma_{\overline{B^0}}(t)$ with a probability ω_{tag}. Therefore equation 7.2 becomes:

$$a_m(t) = \frac{(1-\omega_{tag})\Gamma_{B^0}(t) + \omega_{tag}\Gamma_{\overline{B^0}}(t) - (1-\omega_{tag})\Gamma_{\overline{B^0}}(t) - \omega_{tag}\Gamma_{B^0}(t)}{(1-\omega_{tag})\Gamma_{B^0}(t) + \omega_{tag}\Gamma_{\overline{B^0}}(t) + (1-\omega_{tag})\Gamma_{\overline{B^0}}(t) + \omega_{tag}\Gamma_{B^0}(t)} =$$

$$= \frac{\Gamma_{B^0}(t) - \Gamma_{\overline{B^0}}(t) - 2\omega_{tag}(\Gamma_{B^0}(t) - \Gamma_{\overline{B^0}}(t))}{\Gamma_{B^0}(t) + \Gamma_{\overline{B^0}}(t)} = (1-2\omega_{tag})a_{th}(t) \quad . \tag{7.3}$$

Thus the measured asymmetry $(a_m(t))$ is the precise asymmetry reduced by a factor $(1-2\omega_{tag})$. It is clear now why the $\sin 2\beta$ measurement depends on ω_{tag}. The equation 3.13 becomes:

$$a(t)_m \Big|_{J/\Psi K_S} = (1-2\omega_{tag})\sin(2\beta)\sin(\Delta m t) \quad . \tag{7.4}$$

From a fit to 7.4 with real data, we can only extract $(1-2\omega_{tag})\sin(2\beta)$, thus an independent measurement of ω_{tag} is needed. Since ω_{tag} appears in each asymmetry, it can be extracted from the asymmetry of a different channel. This channel must not be a CP violating channel. This is the case of the control channel, for which the asymmetry is:

$$a(t)_m \Big|_{J/\Psi K^*} = \frac{\Gamma_{mix}(t) - \Gamma_{unmix}(t)}{\Gamma_{mix}(t) + \Gamma_{unmix}(t)} = (1-2\omega_{tag})\sin(\Delta m t) \quad . \tag{7.5}$$

Where Γ_{mix} is the number of the B mesons which oscillated before the decay, and Γ_{unmix} is the number of B which decayed without oscillating. To know if a B has oscillated, the flavor at the production, given by the tagging, is compared with the flavour at the decay, given by the sign of the kaon in the K^{0*} final state.

A quantity which takes into account both the ε_{tag} and ω_{tag}, is the effective efficiency, defined as:

$$\varepsilon_{eff} = \varepsilon_{tag}(1-2\omega_{tag})^2 \quad . \tag{7.6}$$

To minimize the error on an asymmetry measurement, it is necessary to maximize ε_{eff}; so ε_{eff} is considered an estimator of the tagging performances.

The mistag can be inferred only at the end of a selection, from the full sample of data. To have an estimation of the mistag event by event a Neural Network (NN) is used. It will be trained on real data using an independent control channel, and it gives an estimation of the mistag probability as a function of the kinematic variables of the taggers. This allows to split the events in categories with similar mistag. It is observed that dividing the events in categories and combining the

Category	ε_{eff}	ε_{tag}	ω_{tag}	N_R	N_W
OS μ	0.79 ± 0.07 %	5.76 ± 0.10 %	31.5 ± 0.8 %	2312	1062
OS e	0.37 ± 0.05 %	4.13 ± 0.08 %	35.0 ± 1.0 %	1571	847
OS K	1.43 ± 0.10 %	23.42 ± 0.18 %	37.7 ± 0.4 %	8548	5162
SS π	0.77 ± 0.07 %	17.43 ± 0.16 %	39.5 ± 0.5 %	6176	4030
Vertex–charge	0.16 ± 0.03 %	22.35 ± 0.17 %	45.8 ± 0.4 %	7097	5988

Table 7.1: Tagging performances for $B_d^0 \to J/\psi K_S^0$. The results for each tagger are shown.

Category	ε_{eff}	ε_{tag}	ω_{tag}	N_R	N_W
OS μ	0.65 ± 0.03 %	5.46 ± 0.05 %	32.7 ± 0.4 %	8282	4023
OS e	0.32 ± 0.02 %	4.26 ± 0.04 %	36.3 ± 0.5 %	6110	3483
OS K	1.33 ± 0.05 %	23.46 ± 0.09 %	38.1 ± 0.2 %	32719	20128
SS π	0.87 ± 0.04 %	19.03 ± 0.08 %	39.3 ± 0.2 %	26004	16850
Vertex–charge	0.18 ± 0.02 %	23.30 ± 0.09 %	45.6 ± 0.2 %	28558	23921

Table 7.2: Tagging performances for $B_d^0 \to J/\psi K^{0*}$. The results for each tagger are shown.

results of each category improves the tagging performances by about 25%. The combination is done summing the ε_{eff} and the ε_{tag} for each category; this gives the total ε_{eff} and ε_{tag}. Then the equation 7.6 is inverted to calculate the total ω_{tag}. The optimal number of categories is 5: it increases ε_{eff} without reducing too much the number of events in each category. To train the NN the decay mode $B_u^{\pm} \to J/\psi(\mu^+\mu^-)K^{\pm}$ is used. It is very similar to the control channel, and is partially selected with the same selection (see paragraph 7.2). Moreover it is self tagging, because the charged mesons do not oscillate, therefore its flavour is given by the kaon charge. The tagging decision from the algorithm is compared with the true flavour tag, and this is the way the NN is trained.

7.2 Mistag comparison

In tables 7.1 and 7.2 the tagging results of each tagger for both the golden channel and the control channel are shown. The mistags are obtained comparing the tagger

	$B_d^0 \to J/\psi K_S^0$	$B_d^0 \to J/\psi K^{0*}$
Total ω_{tag}	39.10 ± 0.27 %	39.59 ± 0.13 %
Total ε_{tag}	57.4 ± 0.2 %	58.6 ± 0.1 %
Total ε_{eff}	2.73 ± 0.13 %	2.54 ± 0.06 %

Table 7.3: Total tagging performances for both $B_d^0 \to J/\psi K^{0*}$ and $B_d^0 \to J/\psi K_S^0$.

NN Category	ε_{eff}	ε_{tag}	ω_{tag}	N_R	N_W
1	0.81 ± 0.07 %	39.56 ± 0.20 %	42.8 ± 0.3 %	13239	9924
2	0.84 ± 0.07 %	9.62 ± 0.12 %	35.3 ± 0.6 %	3646	1985
3	0.86 ± 0.07 %	4.47 ± 0.09 %	28.1 ± 0.9 %	1882	734
4	0.55 ± 0.06 %	2.10 ± 0.06 %	24.4 ± 1.2 %	931	300
5	0.58 ± 0.05 %	1.60 ± 0.05 %	19.8 ± 1.3 %	752	186

Table 7.4: Tagging performances for $B_d^0 \to J/\psi K_S^0$ divided in categories.

NN Category	ε_{eff}	ε_{tag}	ω_{tag}	N_R	N_W
1	0.75 ± 0.04 %	41.34 ± 0.10 %	43.3 ± 0.2 %	52809	40293
2	0.93 ± 0.04 %	9.31 ± 0.06 %	34.2 ± 0.3 %	13793	7178
3	0.76 ± 0.03 %	4.42 ± 0.04 %	29.2 ± 0.5 %	7049	2912
4	0.52 ± 0.03 %	1.99 ± 0.03 %	24.5 ± 0.6 %	3380	1099
5	0.48 ± 0.03 %	1.49 ± 0.03 %	21.8 ± 0.7 %	2634	733

Table 7.5: Tagging performances for $B_d^0 \to J/\psi K^{0*}$ divided in categories.

decision with the true B flavour given by the Monte Carlo. We can see that the numbers agree inside the errors. The total mistag is obtained combining the decision of each tagger. The total mistag together with the total ε_{eff} and ε_{tag} are shown in table 7.3. For more information about the combination of the tagger decisions in an event see [43].

To exploit the tagging improvement given by a division of the events in five tagging categories, the ω_{tag} of the golden channel and the control channel must be the same for each category. These are shown in tables 7.4, 7.5 and 7.6, together with the tagging efficiencies. The agreement between the mistags in table 7.4 and 7.5 is the successful result of the unified selection.

	$B_d^0 \to J/\psi K_S^0$	$B_d^0 \to J/\psi K^{0*}$
Total ω_{tag}	37.4 ± 0.3 %	37.9 ± 0.1 %
Total ε_{tag}	57.4 ± 0.2 %	58.6 ± 0.1 %
Total ε_{eff}	3.6 ± 0.1 %	3.4 ± 0.1 %

Table 7.6: Total tagging performances for both $B_d^0 \to J/\psi K^{0*}$ and $B_d^0 \to J/\psi K_S^0$ after the division in 5 categories.

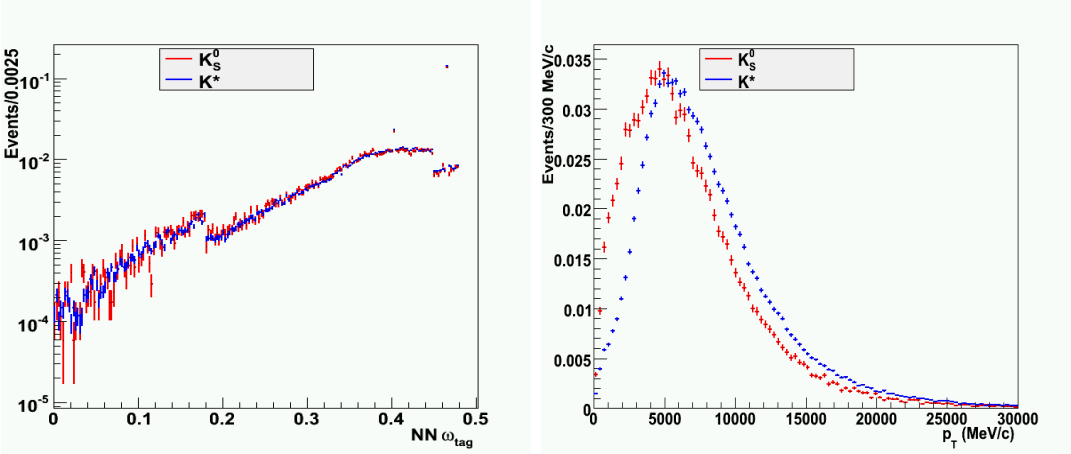

Figure 7.1: Left: ω_{NN}. Right: p_T of the B. In red the golden channel, in blue the control channel. Only selected and L0 triggered signal events are used.

The $\sin 2\beta$ fit can be performed in different ways. The simplest one consists in a fit to all the events, but it uses an higher overall mistag. Another way is to divide the data in the five tagging categories, performing the fit in each category and then combining the results of each category with a weighted mean. If $\sin 2\beta$ sensibly depends on the mistag, an improvement should be seen. The most effective reduction in the $\sin 2\beta$ uncertainty due to the tagging, can be obtained using the NN mistag (ω_{NN}) per event directly in the fit. In fact it can be used in the likelihood, as an estimate of the tagging error. This can work if the NN mistag is a good estimator of the real ω_{tag} (as it seems), and if the ω_{NN} of the golden channel is the same as the ω_{NN} of the control channel for each event. The last is a stronger condition to be satisfied. It is essential that the ω_{NN} distributions and the ω_{NN} dependencies on the kinematic variables are the same for both channels. The first condition is satisfied, see figure 7.1. For other channels it was observed a dependence of the mistag on the p_T of the B meson. This is expected since the

two B mesons emerge from the primary vertex with a strong angular correlation. The B angle is correlated with its momentum: higher the momentum, lower the angle. Thus the momentum distribution of the signal B is correlated to the one of the opposite B. Hence the tagging performances are modified, since there are selection cuts on the p_T of each tagger. The p_T distributions of both the golden channel and the control channel are shown in figure 7.1. These are clearly different since the cuts on the p_T of K^0_S and K^{0*} are different. The p_T of K^0_S is cut at 1 GeV/c, the p_T of K^{0*} is cut at 1.5 GeV/c. If these cuts are taken equal, the differences are strongly reduced, even if they are never cancelled because of the mass difference between K^0_S and K^{0*}. This can potentially give a difference in the tagging performances of the two channels. Figure 7.2 shows the mistags of each tagger as a function of the p_T of B; these plots are fitted with a straight line ($\omega_{tag} = a + b \cdot p_T$). These mistags are calculated from the Monte Carlo truth. The plots in figure 7.2 show different dependences of the ω_{tag} as function of p_T: the most relevant is the one of the same side pions. This dependence is expected since the same side pion is strongly related to the same side B. In figure 7.2 the total mistag as a function of the transverse momentum is also shown. In the total mistag, the different dependencies of ω_{tag} cancel out giving a flat distribution. It means that the total mistags of both the golden channel and the control channel do not depend on the p_T of the B.

The mistag of the background represents a problem in the determination of ω_{tag} of the control channel. It is different from 0.5 for the background events in which a true b was produced. To take it into account, it must be extrapolated from the events in the sidebands of the mass spectrum, provided that the mistag of background in the sidebands is the same as the one under the mass peak. Unfortunately this is not true for the background of $B^0 \to J/\psi K^{0*}$. In figure 7.3, in black, it is shown the mass spectrum of the control channel after the selection performed on the $B^0_{d,u,s} \to J/\psi X$ data, in which the background events always contain a true b. The shoulder structure in the left sideband is called "low mass background". It is due to events in which there is a B with a final state that has μ^+, μ^-, $K^\pm$ and $\pi^\pm$, plus other particles. A B candidate is selected without taking these other particles into account. Therefore the invariant mass is smaller than the true B mass. These are all events with a true B mesons, thus the ω_{tag} is less then 0.5. The events which form the bump in the right sideband come from the $B^\pm_u \to J/\psi(\mu^+\mu^-)K^\pm$ decay mode, a pion has been randomly associated to

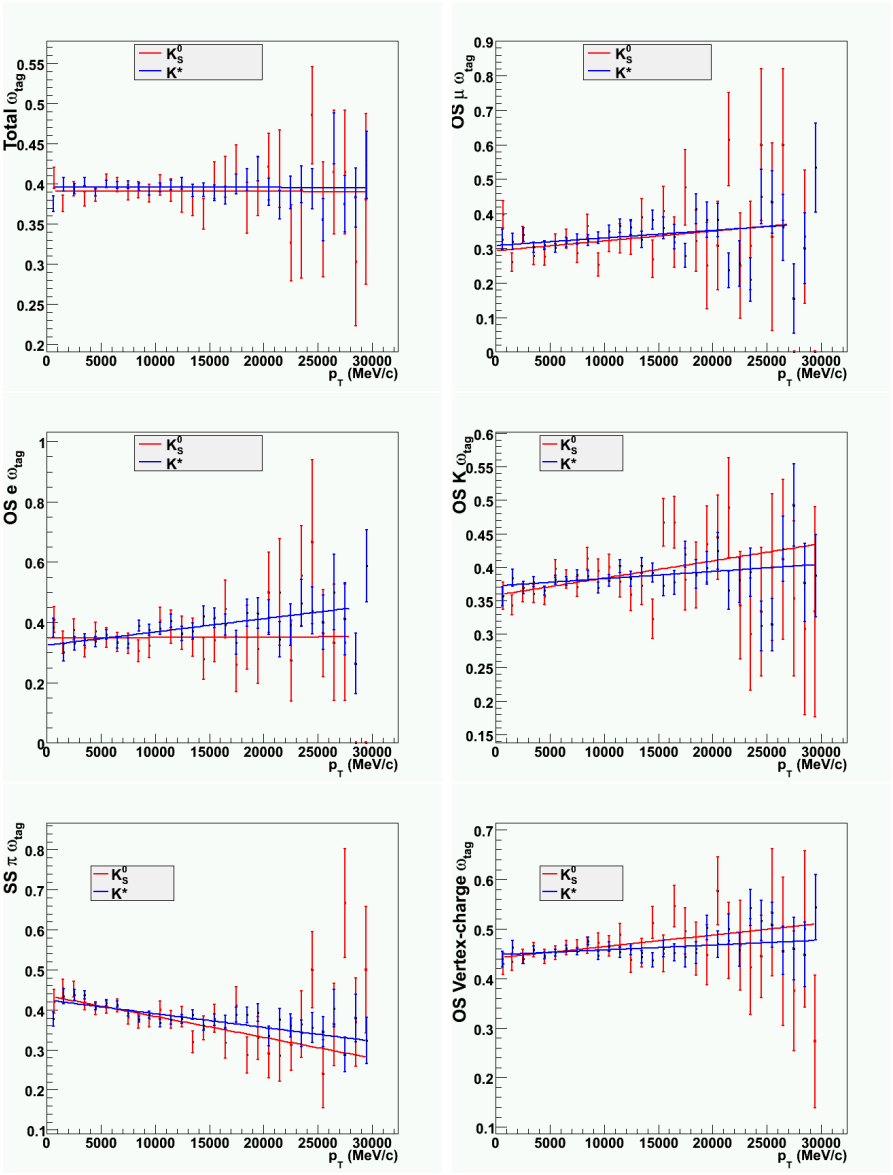

Figure 7.2: ω_{tag} as a function of p_T for the golden channel, in red, and the control channel, in blue. Only selected signal events are used. The plots are made for each tagger separately and for the total ω_{tag}. The distributions are fitted with $\omega_{tag} = a + b \cdot p_T$.

the (μ^+, μ^-, $K^\pm$) of this decay mode, giving a mass very similar to the control channel. Since there is one extra particle, the invariant mass is larger than the B mass. Also this type of background has a mistag lower than 0.5. These two types of background are not present under the mass peak, in which the combinatorial

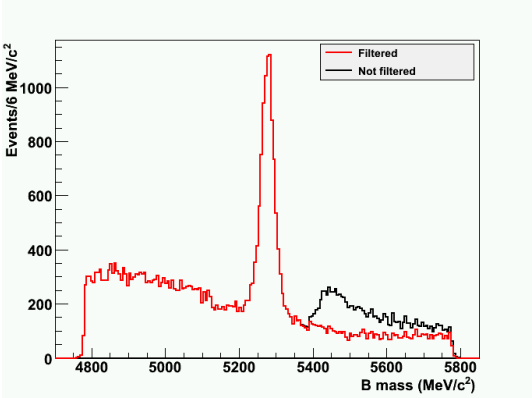

Figure 7.3: B mass spectrum for $B^0_{d,u,s} \to J/\psi X$ events. In black the distribution before the filter. In red the events which pass the filter. Only selected and L0 triggered events are used.

	$m < (m_B - 3\sigma)$	$(m_B - 3\sigma) < m < (m_B + 3\sigma)$	$m < (m_B + 3\sigma)$
$\tau < 0.3$ ps	46.7 ± 0.9 %	45.5 ± 1.7 %	46.0 ± 1.0 %
	$(48.4 \pm 1.2$ %)	$(56.1 \pm 2.2$ %)	$(51.6 \pm 1.2$ %)
$\tau > 0.3$ ps	43.7 ± 0.7 %	46.5 ± 1.6 %	41.2 ± 1.2 %
	$(42.3 \pm 2.9$ %)	$(60.1 \pm 6.5$ %)	$(38.6 \pm 4.5$ %)

Table 7.7: ω_{tag} for $B^0_d \to J/\psi K^{0*}$ selected background events, divided in lifetime and mass intervals. The results for $B^0_{d,u,s} \to J/\psi X$ background are reported out of parenthesis. The results for J/ψ inclusive background are in parenthesis.

and prompt background dominate, and hence ω_{tag} is close to 0.5. This is clear in the comparison of mistags given in table 7.7.

In this table the background events are divided in 6 categories, which take into account not only the mass distribution but also the lifetime distribution. The majority of events at a very low B lifetime are expected to come from prompt background (ω_{tag} nearer 0.5). The other events come more probably from B decays (ω_{tag} less than 0.5). From table 7.7 it is clear that the mistag under the mass peak can't be extracted from the sidebands.

To solve this problem a filter was written which eliminates almost all the $B^\pm_u \to J/\psi(\mu^+\mu^-)K^\pm$ background events from the right sideband. It simply consists in requesting that, for each event, the invariant mass of μ^+, μ^- and $K^\pm$ is outside the 3σ mass window around the B mass. The mass distribution of the events which

	$m < (m_B - 3\sigma)$	$(m_B - 3\sigma) < m < (m_B + 3\sigma)$	$m < (m_B + 3\sigma)$
$\tau < 0.3$ ps	46.7 ± 0.9 %	45.5 ± 1.7 %	47.9 ± 1.1 %
	$(48.4 \pm 1.2$ %)	$(56.1 \pm 2.2$ %)	$(52.0 \pm 1.2$ %)
$\tau > 0.3$ ps	43.7 ± 0.7 %	46.5 ± 1.6 %	44.8 ± 2.1 %
	$(42.3 \pm 2.9$ %)	$(60.1 \pm 6.5$ %)	$(63.1 \pm 6.2$ %)

Table 7.8: ω_{tag} for $B_d^0 \to J/\psi K^{0*}$ selected background events, divided in lifetime and mass categories. The events are filtered requesting $m(\mu^+, \mu^-, K^\pm) < m_B \pm 3\sigma$. The results for $B^0_{d,u,s} \to J/\psi X$ are reported out of parenthesis. The results for J/ψ inclusive background are in parenthesis.

pass the filter is shown in figure 7.3, in read. The mistag results after the filter are shown in table 7.8. The filter eliminates only the events in the right mass window. After the background filter the mistag under the peak and in the right sideband have the same behaviour. Events in the right sideband can be used to estimate the mistag from background events under the peak.

Chapter 8

$\sin 2\beta$ fit

In this chapter the fit strategy of measuring $\sin 2\beta$ from the selected events is described. In the first part of the chapter the model used to fit the data is presented. In the second part it is explained how a simulation was performed to estimate the expected $\sin 2\beta$ sensitivity after one year of data taking at LHCb. The results of this simulation are exposed at the end of the chapter. Only selected and L0 triggered events are considered through the whole chapter.

8.1 Likelihood fit

Given n occurrences $x_1, ..., x_n$ of a sample, with a probability density function (*pdf*) $f(x_i; \theta)$, where θ is a parameter (eventually vector–valued), we call the likelihood function as:

$$L(\theta) = f(x_1, .., x_n; \theta) \ . \tag{8.1}$$

This is the probability density that the sample $x_1, .., x_n$ occurs, given the particular value of the parameter θ. The known outcomes (x_i) are given by the experiment; the only variable is the parameter θ. The value of the parameter which maximizes the likelihood is a good estimator of the parameter itself. This is not always an unbiased estimator, but it asymptotically goes to the unbiased estimator for large n. This method of finding a parameter given a set of data and their *pdf* functions have some advantages in the calculations and it can be used unbinned.

The outcomes of an experiment are usually many measurements of the same variable performed in the same way. They can be considered as independent and obeying the same *pdf*. The joint *pdf* of independent variables is the product of the

pdf of each variable, hence we have:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) \quad . \tag{8.2}$$

As a matter of convenience in calculations this product is transformed in a sum with a logarithm; this is a monotonic function which does not affect the maximum. Finally it's multiplied by a minus, such that it is transformed in a minimization problem, and all the minimization algorithm can be exploited. Therefore we will deal with the minimization of the so called negative log–likelihood (*nll*), defined as:

$$nll(\theta) = -\sum_{i=1}^{n} \log\left(f(x_i;\theta)\right) \quad . \tag{8.3}$$

In the $\sin 2\beta$ case the variables are the measurements of the B lifetime, and the *pdf* of a single variable is the B^0_d decay rate into a CP eigenstate. Since the B^0_d and $\overline{B^0_d}$ decay rates differ, we should divide the sample in two. To distinguish between B^0_d and $\overline{B^0_d}$ the flavour tagging is used. The mistag probability ω_{tag} is part of the *pdf*s of B^0_d and $\overline{B^0_d}$.

The probability of having a signal event or a background event is just the sum of the two individual probabilities. Therefore the combined *pdf* is the sum of the signal *pdf* plus the background *pdf*.

The measured lifetime has a certain error; if we do not include this error, we make an approximation that depends on the size of the error. If we call t the lifetime and σ_t its error, we have that the joined probability density function to have both t and σ_t is:

$$f(t,\sigma_t) = f(t|\sigma_t)g(\sigma_t) = \left[\Gamma(t) \otimes G(t,\sigma_t)\right] g(\sigma_t) \quad , \tag{8.4}$$

where $g(\sigma_t)$ is the *pdf* of σ_t. In general, the observed distribution $f(t|\sigma_t)$ is the conditional *pdf* of having t given a fixed σ_t. It is described by the convolution of the physics model $\Gamma(t)$ and the detector response function $G(t,\sigma_t)$. In this particular case, the physics model is the decay rate. For sake of simplicity it was chosen $g(\sigma_t) = 1$ for the observed σ_t and 0 elsewhere, which means that it was considered a σ_t equal for all the events.

The introduction of an additional variable, the mass, helps in reducing the parameter error due to the presence of background. If we include the mass variable and we make a model of the mass distribution of this background, the background

and the signal in the lifetime distribution will be better separated. Since the mass is independent on the lifetime, we have that the joined *pdf* is the product of the mass *pdf* times the lifetime *pdf*: $f(t,m) = f(t)f(m)$

Therefore the final *pdf* used to build the *nll* is a function of the B lifetime, mass and flavour. To make a model of it, the *pdf* of each variable must be known. They are the sum of the *pdf* for the signal events plus the one for the background events.

The background events are divided in two categories: *prompt background* and *long–life background*. They show two different mass and lifetime distributions. Since these two classes of background are independent, their *pdf*s can be summed to find the total background *pdf*. To define the *pdf* for the signal events the signal data sets were studied. To define the *pdf* for the prompt background the J/ψ inclusive background (without $B_{(d,u,s)} \to J/\psi X$ events) was considered. The $B_{(d,u,s)} \to J/\psi X$ background was considered to build the long–life background *pdf* (both these categories of background are defined in paragraph 5.5). The LHCb simulated data are not enough to be used to estimate the $\sin 2\beta$ value expected after one year of data taking at LHCb.

The ROOT library called RooFit[44] was used to perform the fit. This is a toolkit for modelling the expected distribution of events in a physics analysis. It can manage the convolution and the normalization of the *pdf*s, performing the integrals both analytically (when possible) and numerically. It was originally developed for BaBar, and it contains the decay mode of the B mesons. Finally it provides the instruments to perform the so called "Toy Monte Carlo". This one consists in the generation of a certain amount of data with a given *pdf* that depend on a certain number of parameters. Then the *nll* given by the generated data and the *pdf* are minimized finding the parameters and their errors. The minimization is performed by a specific minimization package called Minuit[45]. In the $\sin 2\beta$ case I generated n times the variables $(t, m, tag\ decision)$ distributed as the total *pdf*: $f(t, m, tag\ decision; \sin 2\beta)$. If n is the number of events selected in one year at LHCb, the data simulated with the Toy Monte Carlo correspond to a simulation of the expected LHCb results. The minimization of *nll* from these data gives $\sin 2\beta$ and its error. This error is considered as the LHCb sensitivity after one year of data taking.

In the following sections the *pdf*s used for signal, long–life background, and prompt background are described. To perform a study as similar as possible to

the analysis made on real data, only the selected events that pass the L0 trigger are considered. A 500 MeV/c^2 B mass window was used. This allows to make a model of the background in the mass sidebands and to extrapolate it under the peak.

8.2 Signal model

Mass

The B mass distribution for the signal events shows a left tail due to events with radiative energy loss. A simple Gaussian cannot describe the tail. The probability density function chosen to describe the mass distribution of signal events is the so called "Crystal Ball":

$$\begin{aligned} S_m(m; m_B, \sigma_{m_B}, a, n) &= \left. \frac{\left(\frac{n}{|a|}\right)^n e^{-\frac{1}{2}a^2}}{\left(\frac{n}{|a|} - |a| - m\right)^n} \right|_{m<-|a|} , \\ &= \left. \exp\left(-\frac{1}{2}\left(\frac{m - m_B}{\sigma_{m_B}}\right)^2\right) \right|_{m>-|a|} \end{aligned} \tag{8.5}$$

This is a Gaussian with a power-law tail. The parameter are m_B and σ_{m_B}: the centre of the Gaussian and its sigma. a represents the threshold at which the tail starts, and n its slope. It is an empirical *pdf* used to describe the effect of radiative energy loss in an invariant mass. The fit of this *pdf* to the simulated signal events is showed in figure 8.1. All the fits shown in this section are performed with the maximum likelihood method. The correspondence between the dots representing the data and the line of $S_m(m; m_B, \sigma_{m_B}, a, n)$ is good. There are some points at the peak that do not match perfectly the line, but they are inside the errors. This small mismatching is explained by the fact that the signal mass distribution is not a single Crystal Ball, but a double one. This is due to the different mass distribution of the LL and DD reconstructed B mesons (see paragraph 5.5). If we fit the mass distribution with a double Crystal Ball, the agreement is very good. The Crystal Ball is a complicated function of four parameters; this imply that a fit of the sum of two Crystal Ball is quite difficult. It does not converge to the expected values when it is performed, in combination with the background *pdf*, on both signal and background data. Since the Crystal Ball functions found by

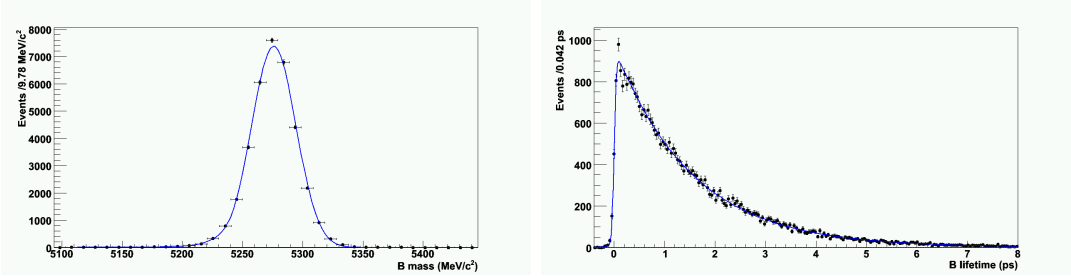

Figure 8.1: Left: B mass from signal events (black dots) fitted with a Crystal Ball (blue line) function. Right: B lifetime from signal events (black dots) fitted with the decay rate function convoluted with a double Gaussian (blue line).

fitting separately the LL B and the DD B mass distributions differ only for the σ_{m_B} value, and since this difference is small, it was decided to use a single Crystal Ball distribution.

Lifetime

From paragraph 3.2, substituting 3.12 in 3.3, we have the B meson decay rates for the golden channel decay mode:

$$\begin{aligned}
\Gamma_{B^0\to f}(t) &= e^{-\frac{t}{\tau}} \mid A_f \mid^2 [1 - \sin 2\beta \sin(\Delta m t)] \quad , \\
\Gamma_{\overline{B^0}\to f}(t) &= e^{-\frac{t}{\tau}} \mid A_f \mid^2 [1 + \sin 2\beta \sin(\Delta m t)] \quad ,
\end{aligned} \tag{8.6}$$

where Γ was substituted with $1/\tau$. Since the mistag probability is non zero, the decay rates become:

$$\begin{aligned}
\Gamma^{obs}_{B^0\to f}(t) &= (1-\omega_{tag})\Gamma_{B^0\to f}(t) + \omega_{tag}\Gamma_{\overline{B^0}\to f}(t) = \\
&= e^{-\frac{t}{\tau}} \mid A_f \mid^2 [1 - (1-2\omega_{tag}) \sin 2\beta \sin(\Delta m t)] \\
\Gamma^{obs}_{\overline{B^0}\to f}(t) &= (1-\omega_{tag})\Gamma_{\overline{B^0}\to f}(t) + \omega_{tag}\Gamma_{B^0\to f}(t) = \\
&= e^{-\frac{t}{\tau}} \mid A_f \mid^2 [1 + (1-2\omega_{tag}) \sin 2\beta \sin(\Delta m t)]
\end{aligned} \tag{8.7}$$

The two rates differ only for a sign: instead of dividing the events in two categories, we can define a joint probability density function of the continuous variable t, and the discrete variable *tag*. *tag* is 1 when the B is tagged as an antiparticle; it is -1 when it is tagged as a particle. Since a *pdf* must be normalized, this eliminates the positive constant factor $\mid A_f \mid^2$. Each positive constant factor entering in the

pdf does not affect the *nll* minimum, and thus can be neglected. The 8.7 becomes:

$$T(t, tag; \tau, \sin 2\beta, \omega_{tag}, \Delta m) = \frac{e^{-\frac{t}{\tau}} \left[1 + tag(1 - 2\omega_{tag}) \sin 2\beta \sin(\Delta m t)\right]}{\int e^{-\frac{t'}{\tau}} \left[1 + tag(1 - 2\omega_{tag}) \sin 2\beta \sin(\Delta m t')\right] dt'} \quad . \tag{8.8}$$

It must be combined with the detector response function. The lifetime resolution can be fitted by a double gaussian distribution. The observed *pdf* for the t and tag variables of the signal B mesons is the convolution between $T(t, tag; \tau, \sin 2\beta, \omega_{tag}, \Delta m)$ and the double Gaussian $G(t; \sigma_{ts1}, \sigma_{ts2}, w_{ts})$:

$$G(t; \sigma_{ts1}, \sigma_{ts2}, w_{ts}) = w_{ts} G(t; \sigma_{ts1}) + (1 - w_{ts}) G(t; \sigma_{ts2}) \quad ,$$

$$\begin{aligned} &S_t(t, tag; \tau, \sin 2\beta, \omega_{tag}, \Delta m, \sigma_{ts1}, \sigma_{ts2}, w_{ts}) = \\ &= T(t, tag; \tau, \sin 2\beta, \omega_{tag}, \Delta m) \otimes G(t; \sigma_{ts1}, \sigma_{ts2}, w_{ts}) \quad , \end{aligned} \tag{8.9}$$

where $\sigma_{ts1}$and σ_{ts2} are the standard deviations of the two Gaussian and w_{ts} is the relative weight. The fit of S_t to the simulated signal events is shown in figure 8.1 (right). There is a good agreement between the *pdf* and the data. The value of τ obtained with this fit is practically the same as the mean lifetime obtained by fitting the simulated data with a single exponential. We fixed ω_{tag} to the value given in table 7.3. The $\sin 2\beta$ obtained is 0.61 ± 0.04: 2.2σ far from the expected value (0.7). The fit repeated using the true B flavour taken from the Monte Carlo truth, and putting in equation 8.8 $\omega_{tag} = 0$, gives $\sin 2\beta = 0.7$. This proves two things: the event selection used do not bias $\sin 2\beta$, and the underestimate is related to the tagging. Several fits were done varying σ_{ts1} and σ_{ts2}, or using a single gaussian model, and no change was observed in the τ and $\sin 2\beta$ values. This is probably due both to the lifetime good resolution (about 0.04 ps) and to the use of a double Gaussian. When the two standard deviations of a double Gaussian are varied the weight w_{ts} can vary to fit the *pdf* to the data.

8.3 Long–life background model

Mass

The mass distribution of the long–life background shows a shoulder in the left sideband. This shoulder, as explained in section 7.2, is the result of the low mass

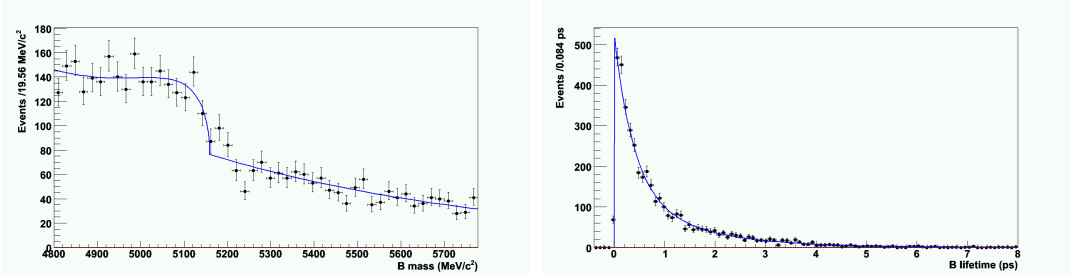

Figure 8.2: Left: B mass from long–life background events (black dots) fitted with an Argus plus an exponential (blue line). Right: B lifetime from long–life background events (black dots) fitted with a double exponential (blue line).

background. The Argus *pdf* is an empirical formula that models the mass distribution of decays near threshold. It is:

$$A(m; m_c, p) = m\sqrt{\left(1 - \left(\frac{m}{m_c}\right)^2\right)} \exp\left(p\left(1 - \left(\frac{m}{m_c}\right)^2\right)\right) \quad , \tag{8.10}$$

where m_c represents a cutoff and p represents the curvature. This *pdf* is very powerful in describing the low mass background distribution. This background shows something like a cutoff, since it is unlikely that the reconstructed mass is larger than the B mass minus the pion mass. In fact the pions are the lighter particles most probably lost in the low mass background. The fit on data confirms this assumption.

The low mass background is superimposed to a combinatorial background. This has an exponential behaviour described by $C(m; s_{ml}) = e^{m\, s_{ml}}$, where s_{ml} is the slope and is about 0. The *pdf* of long–life and combinatorial background must be added. The mass *pdf* for the long–life background events is:

$$L_m(m; m_c, p, s_{ml}, w_{ml}) = w_{ml} C(m; s_{ml}) + (1 - w_{ml}) A(m; m_c, p) \quad , \tag{8.11}$$

where w_{ml} is the fraction of the exponential background over the total long–life background.

Lifetime

The long–life background events are characterized by a mean lifetime significantly larger than 0. This is because they are events in which a B has decayed and

the lifetime presents the typical exponential behaviour of a decay time histogram. Nevertheless it is not well fitted by a single exponential. This is due to the fact that the B produced can be of different type (B_d, B_u, B_s) and they are only partially reconstructed. The particles daughters of B have the production verteces far from the primary vertex, but in the background they are combined with prompt tracks. This shifts the reconstructed secondary vertex towards the primary one, and explains why the distribution has a mean lifetime smaller than the B one. The slope of the exponential is not unique because it is not governed from a unique, distinct physical phenomena as the B_d^0 decay. There are two ways to proceed: either a non parametric *pdf* is used, or a specific model among the known *pdf* is chosen. I followed the second way; in particular I verified that a double exponential is very suited to fit the lifetime distribution of the long–life background (see figure 8.2). It gives a good description of the simulated data and it has not the computing problems of a non parametric *pdf*. This *pdf* is defined as:

$$L_t(t;\tau_{t1},\tau_{t2},w_{tl}) = w_{tl}e^{-\frac{t}{\tau_{t1}}} + (1-w_{tl})e^{-\frac{t}{\tau_{t2}}} \quad , \tag{8.12}$$

where τ_{t1} and τ_{t2} are the slopes of the two exponentials and w_{tl} is the weight. Unlike the signal case, here there is no reason to include the lifetime resolution, since no physical quantity must be extracted from the lifetime of the background.

8.4 Prompt background model

Mass

The prompt background is made of non B events. Thus the mass distribution is a typical exponential:

$$P_m(m;s_{mp}) = e^{s_{mp}m} \quad , \tag{8.13}$$

where s_{mp} is the slope of the exponential and it is about 0. The fit of $P_m(m;s_{mp})$ to simulated data is shown in figure 8.3.

Lifetime

The background events of prompt type have a lifetime equal to zero. They are reconstructed from particles that come from the primary vertex. The fact that the observed distribution is a gaussian centred at zero is due to the (detector or

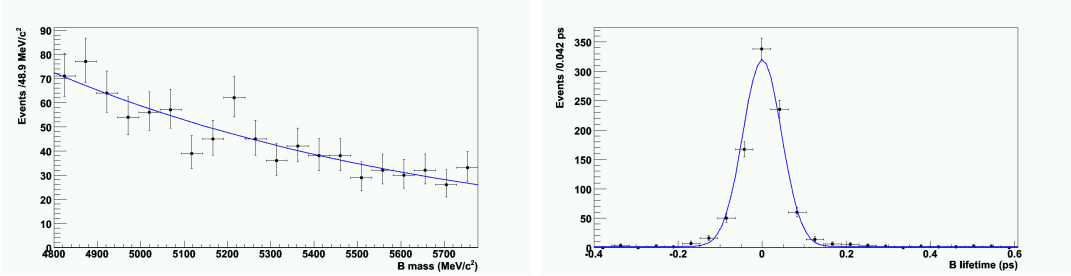

Figure 8.3: Left: B mass of prompt background events (black dots) fitted with an exponential (blue line). Right: B lifetime of prompt background events (black dots) fitted with a double Gaussian (blue line).

experimental) resolution. As in the case of the signal, the lifetime resolution is well fitted with a double Gaussian:

$$P_t(t;\sigma_{tp1},\sigma_{tp2},w_{tp}) = w_{tp}G(t;0,\sigma_{tp1}) + (1-w_{tp})G(t;0,\sigma_{tp2}) \quad , \tag{8.14}$$

where $G(t;0,\sigma_{tp1})$ is a Gaussian with mean 0 and standard deviation σ_{tp1}. The fit of $P_t(t;\sigma_{tp1},\sigma_{tp2},w_{tp})$ to the data is shown in figure 8.3.

8.5 Toy Monte Carlo fit

The total *pdf* is obtained by multiplying the mass *pdf* by the lifetime *pdf* for each category (signal, long–life background and prompt background), and then summing the three *pdf*. The *pdf* of the three categories must be weighted by a factor which takes into account the amount of each category in the sample of selected events. The fraction of signal events is (S/(S+B)) and it will be called the signal fraction. The prompt background fraction f_p is defined as the fraction of prompt events in the background. Thus the total *pdf* becomes:

$$\begin{aligned} &pdf_{tot}(m,t,tag) = \\ &= \left(\tfrac{S}{S+B}\right)\left[S_m(m;m_B,\sigma_{m_B},a,n)S_t(t,tag;\tau,\sin 2\beta,\omega_{tag},\Delta m,\sigma_{ts1},\sigma_{ts2},w_{ts})\right] + \\ &+ \left(1-\tfrac{S}{S+B}\right)\left\{(1-f_p)\left[L_m(m;m_c,p,s_{ml},w_{ml})L_t(t;\tau_{t1},\tau_{t2},w_{tl})\right] + \right. \\ &\left. + f_p\left[P_m(m;s_{mp})P_t(t;\sigma_{tp1},\sigma_{tp2},w_{tp})\right]\right\} \end{aligned} \tag{8.15}$$

The Toy Monte Carlo study consists in generating n times the variables (m,t,tag) according to the pdf_{tot} distribution. If n is the number of selected and triggered events after one year (10^7 s) of data taking at LHCb, this is a simulation of the

first year LHCb results. The simulated variables, together with pdf_{tot}, are used to form a likelihood which is a function of the parameters: (S/(S+B)),f_p,m_B, σ_{m_B}, a, n, τ, $\sin 2\beta$, ω_{tag}, Δm, σ_{ts1}, σ_{ts2}, w_{ts}, m_c, p, s_{ml},w_{ml}, τ_{t1}, τ_{t2}, w_{tl}, s_{mp}, σ_{tp1}, σ_{tp2}, w_{tp}. In principle the minimization of the *nll* made with this likelihood gives the best estimator of each parameter, and hence also of $\sin 2\beta$. Unfortunately it is not easy to minimize a function of 24 parameters, especially because some of these parameters are correlated, and this makes the minimization hard to converge. Some of the parameters will be easily extracted from real data without performing the whole *nll* minimization; these parameters will be fixed in the pdf_{tot}.

The B_d^0 oscillation frequency Δm is very well measured at the B factories, and will be fixed to its world average value[34].

The parameter n of the mass distribution for the signal (S_m) can be extracted from the real data, using the selected and untagged events, which cannot be used in the $\sin 2\beta$ fit. The idea is to consider only untagged events which have a B lifetime larger than 0.3 ps; this eliminates the largest part of the background, which is mainly prompt, and leaves a quite pure sample of signal events. A fit of $(\alpha S_m + (1-\alpha)L_m)$ to these events can be performed, and n can be extracted.

The parameters of the lifetime resolution (σ_{ts1}, σ_{ts2}, w_{ts}) are related to the detector and to the decay mode. It was observed that the value of $\sin 2\beta$ does not change with small variations in the lifetime resolution model. It is possible therefore to use the momentum resolution extracted in control channels.

The slope of the exponential in the mass distribution of the long–life background (s_{ml}) can be easily extracted using the selected events with a B lifetime larger than 0.3 ps and a mass larger than $m_B + 3\sigma_{m_B}$. This is a pure sample of long–life background events; fitting the mass distribution with a simple exponential, gives directly s_{ml}.

A little bit more complicated argument deserves the parameter τ_{t2} of L_t. It defined the largest lifetime in the lifetime distribution of the long–life background. If we integrate the exponentials of equation 8.12 in the interval $[t', +\infty]$ we get:

$$A_1 = \int_{t'}^{+\infty} e^{-\frac{t}{\tau_{t1}}} dt = \tau_{t1} e^{-\frac{t'}{\tau_{t1}}} \quad ,$$

$$A_2 = \int_{t'}^{+\infty} e^{-\frac{t}{\tau_{t2}}} dt = \tau_{t2} e^{-\frac{t'}{\tau_{t2}}} \quad . \tag{8.16}$$

From the fit to simulated data, we have $\tau_{t1} \simeq 1.35$ ps and $\tau_{t2} \simeq 0.38$ ps. Thus substituting $t' = 1$ ps in 8.16 we have: $A_1 = 0.644$ and $A_2 = 0.027$. These are

the areas of the two exponential from 1 ps to infinite. Their ratio is 0.04, which means that the 96% of the background events with B lifetime larger than 1 ps follows the distribution $e^{-\frac{t}{\tau_{t2}}}$. Therefore we can consider all the events that are in the sidebands of the mass distribution, and fit their lifetime distribution with an exponential that goes from 1 ps to infinite. This gives a quite good estimation of the parameter τ_{t2}. The estimation of s_{ml} and τ_{t2} can be performed both on the tagged events and on the untagged ones. The possible bad estimation of these parameters does not introduce any bias in the $\sin 2\beta$ value: they are background parameters, and a small variation of them can be compensated by the weights w_{ml} and w_{tl}.

The values of the parameters used in the Toy Monte Carlo are obtained by fitting the simulated data with the described *pdf*. These values are reported in table 8.1. To know the number of events selected and tagged in a 500 MeV/c^2 mass window, together with (S/(S+B)) and $f_p \equiv (B/S)_{J/\psi\, incl.}/[(B/S)_{tot}]$, calculations identical to that one of paragraph 5.6 are performed. There are two differences with respect to the previous calculations: they were performed using the number of selected events in a 500 MeV/c^2 mass window, and the tagging efficiencies are considered. The second point is crucial, because the tagging efficiency is not the same for the three categories of events. The ε_{tag} for the signal events is 0.574 (from table 7.3), while the ε_{tag} for the long–life background events is 0.651. This is explained by the fact that in the long–life background there are events in which a true B decays into μ^+, μ^-, π^+, π^- plus another particle, for example a $K^\pm$. This particle, since it comes from the B, has a high probability to be selected as an OS kaon by the tagging algorithms. This is not true for the signal events, where the SS B goes only in μ^+, μ^-, π^+, π^-, and the particles used to reconstruct the B can't be reused as taggers. The ε_{tag} for the prompt background events is 0.35 (about 1/2 of the signal ε_{tag}). This is due to the fact that the prompt background is mainly composed by events in which no $b\bar{b}$ pairs are produced. This is extremely helpful in reducing the amount of prompt background in the tagged events. Thus we have the signal yield:

$$\begin{aligned} Y_S &= L_{int} \times \sigma_{b\bar{b}} \times 2 \cdot f_B \times \mathrm{BR}_{vis} \times \varepsilon_{rec.\&sel.} \times \varepsilon_{gen.} \times \varepsilon_{L0} \times \varepsilon_{tag}^{sig.} \\ &= 2\ \mathrm{fb}^{-1} \times 500\ \mu b \times 2 \cdot 0.398 \times 1.79 \cdot 10^{-5}\ \mathrm{cm}^{-2}\mathrm{s}^{-1} \times \\ &\quad \times 0.0527 \times 0.207 \times 0.9390 \times 0.574 \\ &= 83777\ . \end{aligned} \tag{8.17}$$

The background yield and the f_p become:

$$\left(\frac{B}{S}\right)_{tot} = \varepsilon_{tag}^{bkg\ prompt}\left(\frac{B}{S}\right)_{J/\psi\ incl.} + \varepsilon_{tag}^{bkg\ long-life}\left[\left(\frac{B}{S}\right)_{X} + \left(\frac{B}{S}\right)_{b\bar{b}\ incl.}\right] =$$

$$= 0.35 \cdot 24 + 0.651 \cdot [2.20 + 5] = 13.1 \ ,$$

$$f_p = \frac{\varepsilon_{tag}^{bkg\ prompt}\left(\frac{B}{S}\right)_{J/\psi\ incl.}}{\left(\frac{B}{S}\right)_{tot}} = 0.64 \ . \tag{8.18}$$

The total number of events expected as background is:

$$Y_S + \frac{Y_S}{\varepsilon_{tag}^{sig.}} \cdot \left(\frac{B}{S}\right)_{tot} = 1995761 \ . \tag{8.19}$$

To check the correctness of the Toy, 172 Toys were performed. In each Toy the full expected amount of data was generated; each sample differs in the seed of the random number generator. 72% of these Toys converged: this is a good result for such a complicated model with so many parameters. In the unconverged Toys the minimization algorithm does not find the point of absolute minimum. We define the pool of a parameter as the fitted value minus the expected one, all divided by the error on the fitted value. If the Monte Carlo Toys are correct, the distribution of the pool for each variable is expected to be a gaussian centered at zero and with $\sigma = 1$. If the mean value is not zero, there is a bias in the estimation of the variable. If the standard deviation is more(less) than one, there is an underestimation(overestimation) of the error on the fitted parameter. A pool distribution was constructed for each variable using the results of the converged Toys. All these distributions are as expected. Four of them are shown in figure 8.4.

8.6 $\sin 2\beta$ sensitivity

In figure 8.5 the Toy Monte Carlo simulated mass and lifetime are shown superimposed to their *pdf* distributions. In the same figure the asymmetry plot is shown. It is clear from the plot that the fit with the minimal *nll* method can correctly extract 16 free parameters, obtaining *pdf* that describes the simulated data. In particular the asymmetry, which depends only on $\sin 2\beta$, is well determined. The results of one fit are shown in table 8.1. The *nll* as a function of $\sin 2\beta$ is shown

Fixed parameters			
Parameter	**Initial value**		
n	7		
ω_{tag}	0.3908		
Δm	0.502		
σ_{ts1}	$3.32 \cdot 10^{-2}$		
σ_{ts2}	$7.21 \cdot 10^{-2}$		
w_{ts}	0.867		
s_{ml}	$-1.43 \cdot 10^{-3}$		
τ_{t2}	1.35		
Free parameters			
Parameter	**Initial value**	**Fitted value**	**Glob. corr.**
$\frac{S}{S+B}$	$4.10 \cdot 10^{-2}$	$4.24 \pm 0.021 \cdot 10^{-2}$	0.54
f_p	0.64	0.64 ± 0.0008	0.86
m_B	5279.2	5279.0 ± 0.10	0.22
σ_{m_B}	18.4	18.3 ± 0.10	0.51
a	1.85	1.85 ± 0.07	0.55
τ	1.536	1.533 ± 0.006	0.21
$\sin 2\beta$	0.700	0.730 ± 0.031	0.03
m_c	5159.4	5159.8 ± 0.025	0.11
p	-16.5	-16.6 ± 0.21	0.44
w_{ml}	0.84	0.84 ± 0.0016	0.59
τ_{t1}	0.38	0.38 ± 0.003	0.84
w_{tl}	0.41	0.41 ± 0.0023	0.81
s_{mp}	$-1.052 \cdot 10^{-3}$	$-1.056 \pm 0.003 \cdot 10^{-3}$	0.20
σ_{tp1}	$4.56 \cdot 10^{-2}$	$4.55 \pm 0.004 \cdot 10^{-2}$	0.45
σ_{tp2}	$8.20 \cdot 10^{-1}$	$8.36 \pm 0.13 \cdot 10^{-1}$	0.83
w_{tp}	0.926	0.924 ± 0.0013	0.90

Table 8.1: Parameters of a fit with their initial and final values and their global correlation coefficients.

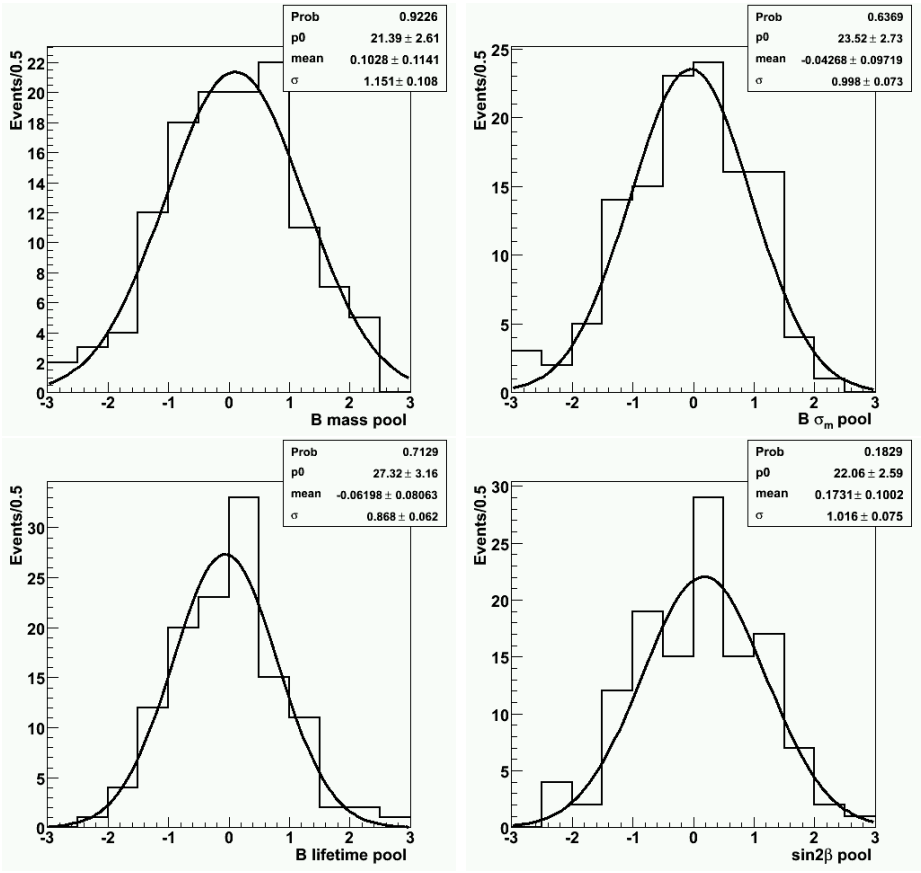

Figure 8.4: Pool distributions of m_b, σ_{m_B}, τ and $\sin 2\beta$.

in figure 8.5. The expected error on $\sin 2\beta$ after one year (10^7 s) of data taking at LHCb is 0.031.

The Toy Monte Carlo fit was repeated with different amount of background. In particular B/S and the f_p were varied. The results of these fits are shown in table 8.2. The fit with no background was performed only on the time distribution, since in this case there is no need to separate signal from background. From the results shown in table 8.2 two things are clear: the main contribution to $\sigma_{\sin 2\beta}$ is due to statistics of the signal, and the prompt background does not deteriorate significantly the $\sin 2\beta$ measurement. Reducing the amount of background by 1/4, 10% better sensitivity is achieved; if we use a number of signal events four time larger we reduce $\sigma_{\sin 2\beta}$ by $1/\sqrt{4}$ (50%). Thus an increase of $\varepsilon_{sel.}$ is more effective than a reduction of B/S. If we reduce the amount of prompt background, maintaining B/S constant, (i.e. increasing the fraction of long–life background)

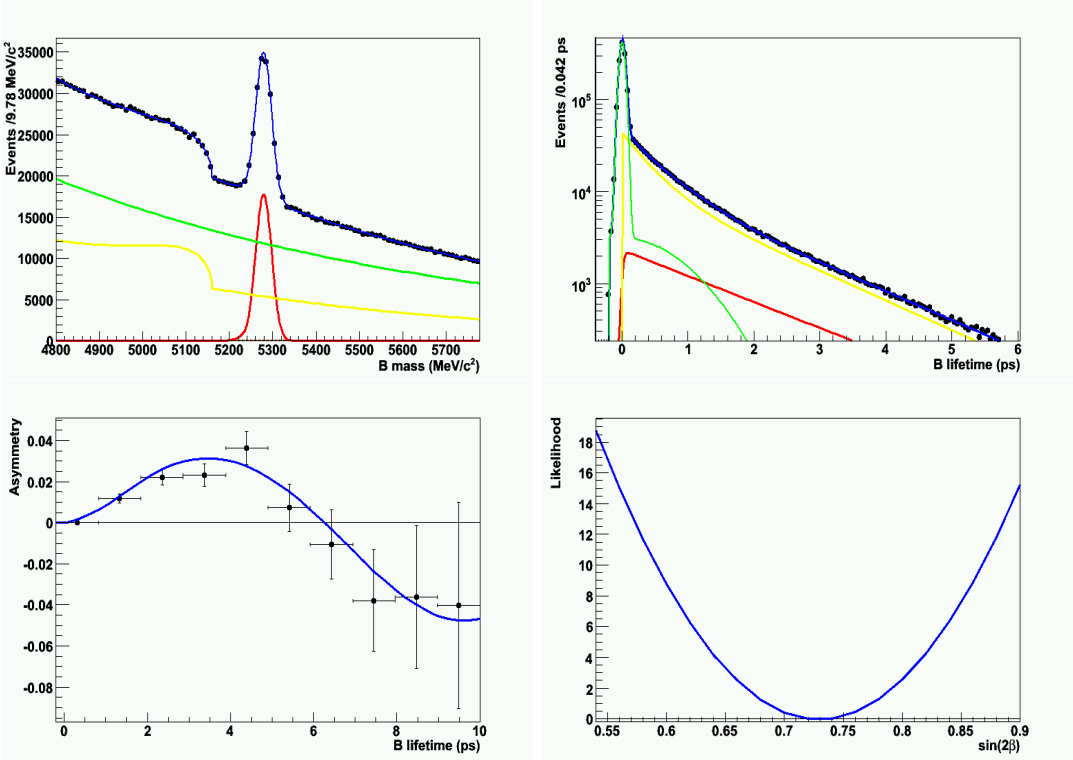

Figure 8.5: Top left: the mass distribution. Top right: the lifetime distribution in logarithmic scale. For each of these two plot the black dots are the Toy Monte Carlo simulated values, the lines are the fitted *pdf*. The line in blue is the total *pdf*, in red is the signal *pdf*, in green is the prompt background, in yellow is the long–life background. Bottom left: the asymmetry distribution. The black dots are the Toy Monte Carlo simulated values, the blue line is the fitted curve. Bottom right: *nll* scan in function of $\sin 2\beta$.

a larger $\sigma_{\sin 2\beta}$ is obtained. This means, as expected, that the effect of prompt background is neglegible with respect to the long–life one.

Background amount	$\sigma_{\sin 2\beta}$
Baseline $\frac{B}{S} = 13.1$ $f_p = 0.64$	0.031
No background	0.026
$\frac{3}{2}$ $\frac{B}{S}$	0.033
$\frac{1}{2}$ $\frac{B}{S}$	0.029
$\frac{1}{4}$ $\frac{B}{S}$	0.028
$\frac{1}{8}$ $\frac{B}{S}$	0.027
$\frac{1}{16}$ $\frac{B}{S}$	0.027
$\frac{3}{2}$ f_p	0.028
$\frac{1}{2}$ f_p	0.034
$\frac{1}{4}$ f_p	0.035

Table 8.2: $\sigma_{\sin 2\beta}$ results for different amount of B/S and f_p.

Conclusions

The selection of $B_d^0 \to J/\psi(\mu^+\mu^-)K_S^0(\pi^+\pi^-)$ (golden channel) and $B_d^0 \to J/\psi(\mu^+\mu^-)K^{*0}(K^+\pi^-)$ (control channel) were tuned using the LHCb simulation. Data selection algorithms considered the kinematics and topological features of B_d^0 meson decays to achieve the goal of selecting the largest number of signal events, while keeping the background level at a low value. A particular care was put in reducing the background from b events. After the selection an annual signal yield of 144568 events for the golden channel is expected and 626931 events for the control channel. The total (B/S) is $\sim$3 for the golden channel and $\sim$7 for the control channel. The selection was performed taking care not to introduce a bias in the lifetime distribution of the B meson. The selection cuts applied are as similar as possible for the two channels.

A comparison between the tagging variables of both channels was made and it was demonstrated that they agree inside the errors. The mistag fraction measured from a division of the simulated sample in five tagging categories is: $(37.4 \pm 0.3)\%$ for the golden channel, and $(37.9 \pm 0.1)\%$ for the control channel. It was shown how the mistag does not depent on the p_T of the reconstructed B.

In order to measure from a fit to data $\sin 2\beta$ after one year of data taking, a *pdf* model for the lifetime and mass distribution of the B mason was constructed. The $B_d^0 \to J/\psi(\mu^+\mu^-)K_S^0(\pi^+\pi^-)$ selected events of both signal and background were considered. To estimate the LHCb sensitivity a number of events equal to the expected selected and tagged signal events in one year of data taking was simulated using this *pdf*. From the *pdf* and the simulated events a likelihood fit was performed to extract $\sin 2\beta$. The sensitivity to $\sin 2\beta$ obtained is $\sigma_{\sin 2\beta} = 0.031$. The precision $\sigma_{\sin 2\beta}$ improves more increasing the statistics of signal events that decreasing the background level. Moreover it was shown that the effect of prompt background is negligible.

My work was performed in the frame of the Flavour Tagging and Proper Time

& Mixing working group. I studied the golden channel in all its aspect, and tuned the selection of its control channel. The results of my work will be used by the LHCb HLT group to optimize the trigger of signal events. The results of this work were presented by me at the meetings of 18 July 2007[46], 12 December 2007[47], and 21 February 2008[48].

Bibliography

[1] D. H. Perkins, *Introduction to high energy physics*, Cambridge University Press (2000), 4^{th} edition.

[2] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5** (1967), 32

[3] A. D. Sakharov, Sov. Phys. JETP Lett. **5** (1967), 24

[4] G. C. Branco *et al.*, *CP Violation*, Oxford University Press (1999)

[5] C. S. Wu *et al.*, Phys. Rev. **105** (1957), 1413

[6] J.H. Christenson *et al.*, Phys. Rev. Lett. **13** (1964), 138

[7] NA 48 Collaboration, Fanti, V. *et al.*, Phys. Lett. **465B** (1999), 335;
K TeV Collaboration, Alavi-Harati, A. *et al.*, Phys. Rev. Lett. **83** (1999), 22

[8] B. Aubert *et al.*, Nucl. Instr. and Meth., **A479** (2002), 1

[9] A. Abashian *et al.*, Nucl. Instr. and Meth. **A479** (2002), 117

[10] M. Kobayashi *et al.*, Prog. Theor. Phys. **49** (1973), 652

[11] L. Wolfenstein, Phys. Rev. Lett. **51** (1983), 1945

[12] LHC Study Group, *LHC Conceptual design report (Yellow Book)*, (CERN/AC/95–05 (LHC)), 1995.

[13] ATLAS Collaboration, *ATLAS Technical Proposal*, CERN/LHCC/94–43 (1994)

[14] CMS Collaboration, *CMS Technical Proposal*, CERN/LHCC/94–38 (1994)

[15] ALICE Collaboration, *ALICE Technical Proposal*,
CERN/LHCC/95–71 (1995)

[16] TOTEM Collaboration, *TOTEM Technical Proposal*,
CERN/LHCC/99–7 (1999)

[17] LHCB Collaboration, *LHCb Technical Proposal*,
CERN/LHCC/98–4 (1998)

[18] LHCB Collaboration, *LHCb Reoptimized Detector Design and Performance TDR*,
CERN/LHCC/2003–030 (2003)

[19] M. Needham and T. Ruf, *Estimation of the material budget of the LHCb detector*,
Note LHCb 2007–025 (2005)

[20] The LHCb Collaboration, *LHCb VELO TDR*,
CERN/LHCC/2001–011 (2001)

[21] The LHCb Collaboration, *LHCb Magnet TDR*,
CERN/LHCC/2000–007 (2000)

[22] The LHCb Collaboration, *LHCb Outer Tracker TDR*,
CERN/LHCC/2001–024 (2001)

[23] The LHCb Collaboration, *LHCb Inner Tracker TDR*,
CERN/LHCC/2002–029 (2002)

[24] The LHCb Collaboration, *LHCb RICH TDR*,
CERN/LHCC/2000–037 (2000)

[25] The LHCb RICH Group, *LHCb RICH 1 Engineering Design Review Report*,
LHCb–EDR–2004–121 (2004)

[26] The LHCb RICH Group, *LHCb RICH 2 Engineering Design Review Report*,
LHCb–EDR–2002–009 (2002)

[27] R. Forty *et al.*, *RICH pattern recognition*,
LHCb/98–040 (1998)

[28] The LHCb Collaboration, *LHCb Calorimeters TDR*, CERN/LHCC/2000–036 (2000)

[29] The LHCb Collaboration, *LHCb Muon System TDR*, CERN/LHCC/2001–010 (2001)

[30] The LHCb Collaboration, *Addendum to the LHCb Muon System TDR*, CERN/LHCC/2003–002 (2003)

[31] The LHCb Collaboration, *Second Addendum to the LHCb Muon System TDR*, CERN/LHCC/2005–012 (2005)

[32] The LHCb Collaboration, *LHCb Trigger System TDR*, CERN/LHCC/2003–031 (2003)

[33] E. Rodrigues, *The LHCb Trigger System*, LHCb, 2006-065 (2006)

[34] W.-M.Yao et al.(Particle Data Group), J. Phys. G 33, 1 (2006)

[35] The LHCb Collaboration, *LHCb Computing TDR*, CERN/LHCC/2005–019 (2005)

[36] B. Aubert *et al.* BABAR Collaboration. Phys. Rev. D 71, 032005 (2005)

[37] R. Itoh *et al.* BELLE Collaboration. Phys. Rev. Lett. 95, 091601 (2005)

[38] P. Krokovny *et al.* BELLE Collaboration. Phys. Rev. Lett. 97, 081801 (2006)

[39] T. Sjostrand *et al.*, *High-Energy physics event generation with Pythia*, Comp.Phys.Comm., 135 (2001) 238.

[40] D. Lange et al., *The EvtGen particle decay simulation package*, Nucl.Inst.Meth.A 462 (2001) 152.

[41] René Brun and Fons Rademakers, *ROOT: An Object-Oriented Data Analysis Framework*, http://root.cern.ch/

[42] S. Amato *et al.*, *Update to the LHCb sensitivity to* $\sin 2\beta$ *from the CP-asymmetry in* $B^0 \to J\psi(\mu\mu)K^0_S$ *decays*,
LHCb, 2007-045 (2007)

[43] M. Calvi, O. Leroy and M. Musy, *Flavour Tagging Algorithms and Performances in LHCb*,
LHCb, 2007-058 (2007)

[44] D. Kirkby W. Verkerke, *The ROOFIT Toolkit for Data Modelling*,
http://roofit.sourceforge.net/

[45] Fred James, *MINUIT*,
http://wwwinfo.cern.ch/asdoc/minuit/

[46] http://indico.cern.ch/contributionDisplay.py?contribId=3&confId=9609

[47] http://indico.cern.ch/contributionDisplay.py?contribId=5&confId=19350

[48] http://indico.cern.ch/contributionDisplay.py?contribId=3&confId=26053

List of Figures

List of Tables

Acknowledgements

I would like to thank Marta Calvi for her help and supervision in my thesis work. Many thanks to teach me all that I know in the field of data analysis and to give me the possibility to work as a member of the LHCb experiment. Thank you to Clara Matteuzzi who supervised me and was patient enough to correct my very bad English.

A special thank to Marco Musy. He was always ready to debug my codes even when he had no time. He was also able to answer to all kind of questions. My gratitude goes also to Tito Bellunato and Davide Luigi Perego. I would like especially to thank Davide that taught me many triks to survive as a thesis student.

I spent four months working at CERN in the RICH group, during this time I worked also on my thesis. I would like to thank the great Olav Ullaland for giving me the opportunity to work at CERN in a fantastic team. He taught me lots of things on CERN and LHCb. He also gave me the possibility to built a little detector with my hands. Above all he made those four mounths unforgettable with his irony. And how to forget Erich Albrecht with his Wednesdays passed with him, Olav and other fantastic people. A special thank to Gaia Lanfranchi who helped me in my four mounths period of study at CERN. She worked with me in the selections and taught me the working method and some pretious trick. I thank also Stefania Vecchi for the precious help on the fit program.

I would like to thank all the friends that I met in these five years at the Bicocca university, among them I would like to cite Henry, Giovanni and Marco. A special thank goes to my dear friend Damiano Tommasini.

My gratitude goes to my family. I thank my syster Lucrezia Mangiafave who corrected the first, unreadable, version of this thesis, and I thank my mother and my father that believed in me.